Polskie Towarzystwo Informatyczne

PTI

Bezpieczeństwo i efektywność systemów informatycznych

Redakcja naukowa:
Kapczyński Adrian
Trawka Janusz

Bezpieczeństwo i efektywność systemów informatycznych

Redakcja naukowa:

Kapczyński Adrian
Trawka Janusz

Wydawnictwo PTI Oddział Górnośląski
Katowice 2009

Recenzenci:
Prof. dr hab. inż. Andrzej Grzywak
Prof. dr hab. inż. Jerzy Klamka

Redakcja techniczna:
Mgr inż. Arkadiusz Banasik
Mgr inż. Anna Dudek
Mgr inż. Piotr Janke
dr inż. Adrian Kapczyński

Projekt okładki:
Jacek Uroda

ISBN: 978-83-60810-28-6

Polskie Towarzystwo Informatyczne
Oddział Górnośląski
40-040 Katowice, ul. J. Lompy 2/10
Tel. +48 (32) 253 61 09
www.pti.katowice.pl
e-mail: biuro@pti.katowice.pl

Wydrukowano w Pracowni Komputerowej Jacka Skalmierskiego.
Nakład – 100 egz.

SPIS TREŚCI

CZĘŚĆ 1. BEZPIECZEŃSTWO INFORMACJI

CZĘŚĆ 2. EFEKTYWNOŚĆ SYSTEMÓW INFORMATYCZNYCH

Przedmowa

Niniejsza monografia prezentuje nowe wyniki prac badawczych z zakresu bezpieczeństwa oraz efektywności systemów informatycznych.

Pierwsza część opracowania jest poświęcona zagadnieniom bezpieczeństwa informacji i składa się z dziewięciu rozdziałów.

W drugiej części opracowania zamieszczono rezultaty rozważań z zakresu efektywności systemów informatycznych i składa się z dwunastu rozdziałów.

Trzecia część książki dotyczy zastosowań informatycznych i składa się z dziewięciu rozdziałów.

W ostatniej części, która składa się z trzech rozdziałów przedstawiono pozostałe, interdyscyplinarne zagadnienia.

Redaktorzy składają serdeczne podziękowania wszystkim autorom, których dorobek umożliwił powstanie niniejszej monografii.

Adrian Kapczyński

Janusz Trawka

Część 1.

Bezpieczeństwo informacji

Rozdział 1

Bezpieczeństwo systemów informatycznych – teraźniejszość i przyszłość

Katarzyna Oryszczak
Akademia Ekonomiczna im. K. Adamieckiego w Katowicach

badanie.bezpieczeństwo.si@gmail.com

Streszczenie

W rozdziale przedstawiono wyniki badań pilotażowych przeprowadzonych w Internecie wśród indywidualnych użytkowników systemów informatycznych, jak i pracowników przedsiębiorstw prowadzących działalność w różnych branżach, w tym w branży informatycznej. Celem opracowania jest analiza i ocena obecnej sytuacji bezpieczeństwa systemów informatycznych w opinii respondentów. Autor podjął próbę odpowiedzi m.in. na pytania: „Z jakim naruszeniem bezpieczeństwa systemów informatycznych spotykają się użytkownicy w gospodarstwach domowych i przedsiębiorstwach?", „Z jakimi zagrożeniami bezpieczeństwa systemów informatycznych mają do czynienia użytkownicy?", „Jakie sposoby zabezpieczeń wykorzystują użytkownicy w celu poprawy bezpieczeństwa systemów informatycznych?" Oprócz tego Autor na podstawie przeprowadzonych badań dokonał rozpoznania kierunku, w jakim będzie się rozwijać bezpieczeństwo systemów informatycznych.

„Jedyną osobą odpowiedzialną za twoje bezpieczeństwo jest ta, którą widzisz w lustrze"
/Bill Jeans/

1. Wprowadzenie

Ciągły i efektywny dostęp do informacji - danych biznesowych jest kluczowym elementem sprawnego funkcjonowania przedsiębiorstwa na rynku. Wymiana

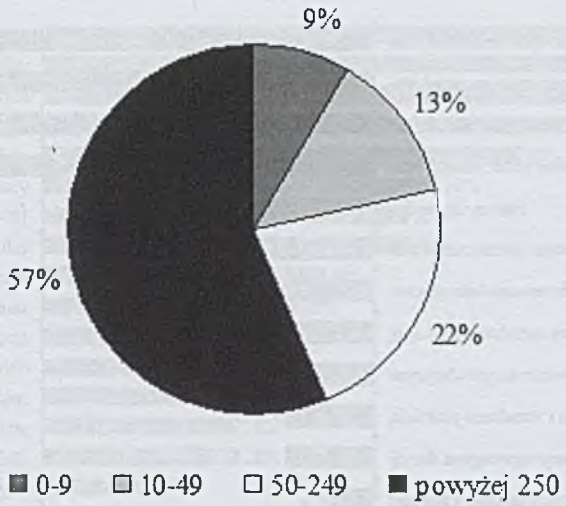
informacji i przetwarzanie danych odbywa się również pomiędzy użytkownikami gospodarstw domowych. Rosnące poczucie bezpieczeństwa przechowywania danych i informacji w systemach, liczba pojawiających się nowych zagrożeń bezpieczeństwa systemów informatycznych, a w związku z tym rozwój środków ochrony bezpieczeństwa, których jest wciąż zbyt mało w stosunku do zagrożeń zrodził potrzebę przeprowadzenia badań wśród dwóch grup docelowych: pracowników przedsiębiorstw i użytkowników indywidualnych w gospodarstwach domowych. Dostępne są na rynku raporty i wyniki badań odnoszące się do sytuacji w przedsiębiorstwach prowadzących działalność głównie w branży informatycznej[1], istnieje jednak luka informacyjna dotycząca stanu bezpieczeństwa systemów informatycznych w przedsiębiorstwach, które funkcjonują także w innych branżach. Brakuje również wyników badań, które przedstawiałyby sytuację bezpieczeństwa systemów informatycznych w gospodarstwach domowych.

2. Charakterystyka i przebieg badań internetowych

Badanie bezpieczeństwa systemów informatycznych zostało zrealizowane za pośrednictwem Internetu i przy użyciu narzędzia, jakim jest kwestionariusz on-line zamieszczonym w domenie www.webankieta.pl . Został przygotowany kwestionariusz dla przedsiębiorstw oraz na potrzeby gospodarstwa domowego, które korzystają z systemów informatycznych i mają dostęp do Internetu. Pytania w kwestionariuszach mają charakter pytań zamkniętych i otwartych. Trzeba podkreślić, iż pytania zawarte w kwestionariuszu on-line mogły sprawić problem respondentom ze względu na poruszane w nim kwestie, które nie należą do najłatwiejszych, co z resztą było zabiegiem zamierzonym. Badanie ma charakter badań pilotażowych.

W badaniu wzięli udział pracownicy przedsiębiorstw prowadzących działalność m. in. w bankowości, administracji (samorządowa, publiczna), w opiece zdrowotnej, w branży motoryzacyjnej, chemicznej, handlowej, informatycznej (w tym sprzedaż internetowa, IT, serwery, rozrywka on-line), ubezpieczeniowej, projektowej (m.in. budownictwo), FMCG, naukowo-edukacyjnej oraz przemysłowej (w tym metalowej). Większość pracowników reprezentuje przedsiębiorstwa duże zatrudniające powyżej 250 pracowników (57%) oraz średnie - od 50 do 249 pracowników (22%) – rys. 1. 13% stanowią pracownicy mniejszych przedsiębiorstw, które liczą od 10 do 49 pracowników oraz mikroprzedsiębiorstwa - zatrudniające do 9 pracowników (9%).

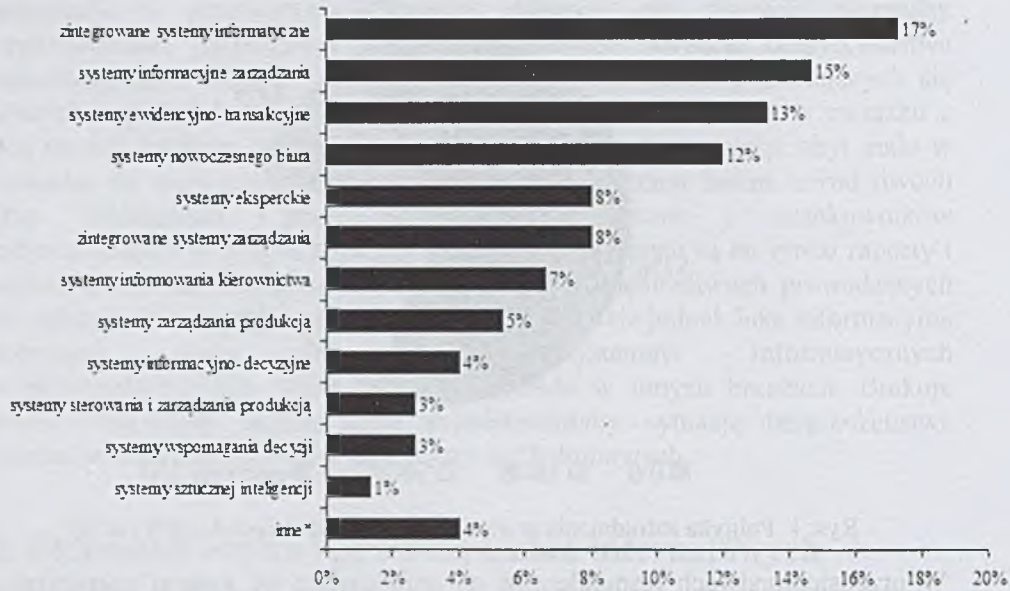
[1] Na przykład raporty opracowane przez firmy Symantec, Oracle



Rys. 1. Polityka zatrudnienia w przedsiębiorstwach respondentów (w %)

W przedsiębiorstwach respondentów na ogół stosuje się system operacyjny z rodziny systemów Microsoft Windows (Windows XP, Windows Professional XP, Windows 2000, Windows Vista, Windows Serwer 2003) – blisko 73%. Niewiele przedsiębiorstw wykorzystuje w swojej działalności system operacyjny Linux (prawie 15%) i inne (12%)[1]. Blisko 20% badanych przedsiębiorstw w działalności gospodarczej wykorzystuje zintegrowane systemy informatyczne. Z systemów informacyjnych zarządzania korzysta 15% przedsiębiorstw, systemy ewidencyjno-transakcyjne posiada 13% przedsiębiorstw. W opinii respondentów powyżej 10% przedsiębiorstw używa systemy nowoczesnego biura – rys. 2.

[1] Respondenci wskazali jako „inne": Solaris, Mac os, Debian, Simik, Novell

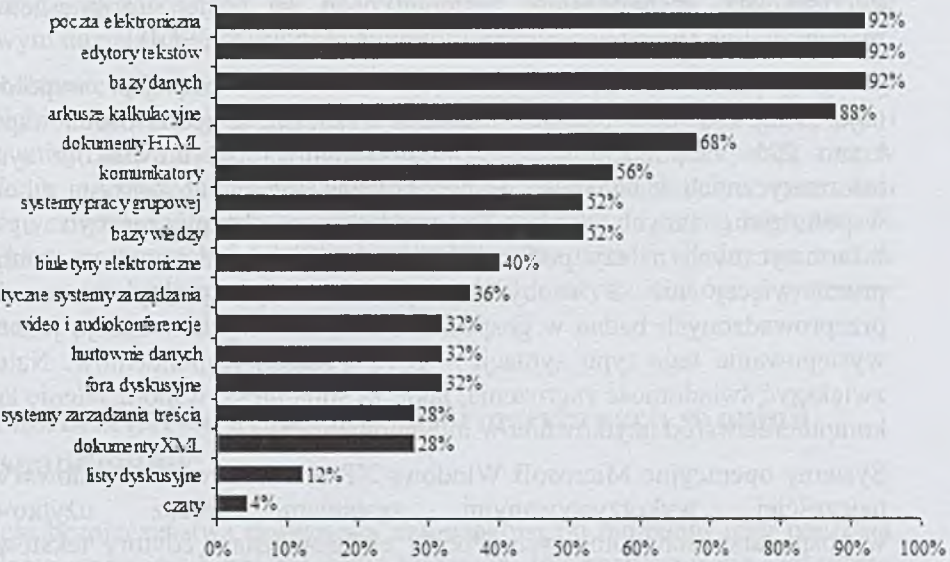


**nie wiem, specjalistyczne systemy obliczeniowe np. ClusterX*

Rys. 2. Systemy informatyczne wykorzystywane w działalności przedsiębiorstw respondentów (w %)

W większości badanych przedsiębiorstw pracownicy równorzędnie korzystają z technologii informatycznych i systemów informatycznych, takich jak poczta elektroniczna, edytory tekstów i bazy danych (po 92% wskazań). Prawie 90% respondentów pracuje przy użyciu arkuszy kalkulacyjnych, a blisko 70% badanych posługuje się dokumentami HTML – rys. 3. Respondenci deklarują także, że podczas pracy korzystają z komunikatorów (około 60% wskazań), systemów pracy grupowej oraz baz wiedzy (powyżej 50% wskazań). Aż 32% respondentów wskazuje, iż używa fora dyskusyjne w pracy.

W opinii pracowników badanych przedsiębiorstw 45% użytkowników komputerów nie współdzieli kont. W pozostałych przypadkach konta na komputerze współdzieli dwóch pracowników (23% wskazań), zdaniem 18% respondentów od trzech do czterech pracowników korzysta ze wspólnego konta. Nawet pięciu i więcej pracowników współdzieli konta w przedsiębiorstwach – aż 14% wskazań. Współdzielenie konta jest źródłem zwiększonego ryzyka dla bezpieczeństwa systemów informatycznych. Sytuacja, w której na stanowisku pracuje więcej niż jeden współwłaściciel konta może doprowadzić do braku kontroli przez przełożonych, a także braku odpowiedzialności pracownika za bezpieczeństwo systemów informatycznych oraz spadku efektywności i organizacji pracy. Niestety, jak wynika z przeprowadzonych badań nie wszystkie przedsiębiorstwa, zwłaszcza zatrudniające większą liczbę pracowników mają tego świadomość.



Rys. 3. Technologie informatyczne i systemy informatyczne stosowane w badanych przedsiębiorstwach (w %)

W badaniu internetowym użytkowników indywidualnych – użytkowników systemów informatycznych gospodarstwach domowych przeważali mężczyźni (53%), pozostałe 47% stanowiły kobiety. Najmłodszy respondent był w wieku 16 lat, a najstarszy miał 54 lata. Najliczniejszą grupą respondentów biorących udział w badaniu była grupa osób w wieku od 26 lat do 35 lat (60%), a o ponad połowę mniej było respondentów w wieku do 25 lat. Aż 74% użytkowników to osoby z wykształceniem wyższym, średnim wykształceniem legitymuje się blisko 25% respondentów. Zaledwie 1% badanych deklaruje ukończenie szkoły podstawowej. Użytkownicy korzystający z systemów informatycznych w gospodarstwach domowych pochodzą głównie z miejscowości województwa śląskiego (Katowice, Gliwice, Bytom, Będzin, Chorzów, Sosnowiec, Ruda Śląska, Rybnik, Dąbrowa Górnicza, Tychy), ale także z miast Warszawa, Szczecin, Płock, Słupsk, Wodzisław Śląski, Olsztyn i Kraków. Wśród użytkowników indywidualnych w badaniu wzięli udział również mieszkańcy Londynu i Duesseldorfu.

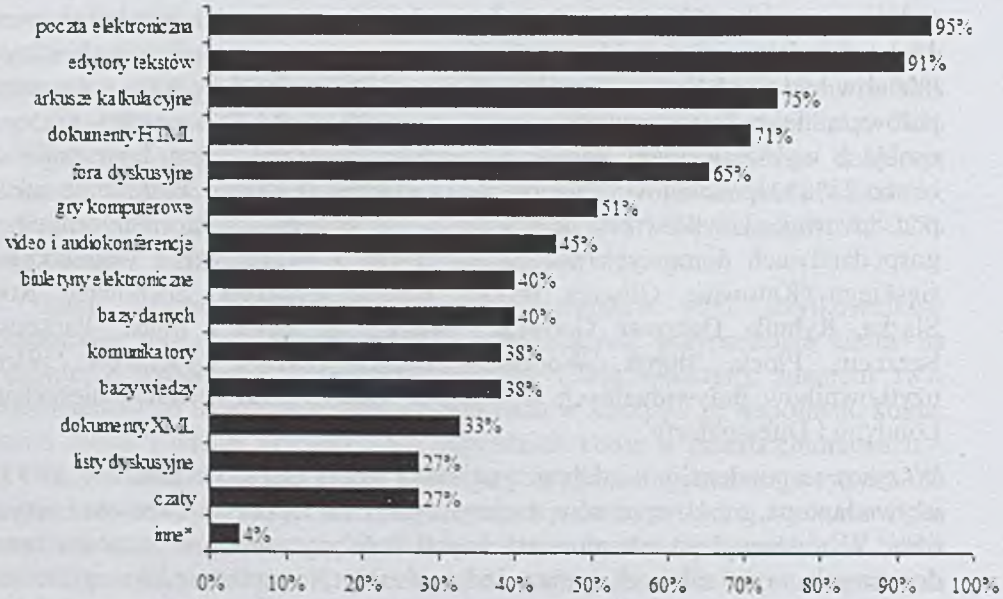
Wszyscy respondenci posiadają przynajmniej jeden telefon komórkowy, aż 93% używa laptopa, a z komputerów stacjonarnych z dostępem do Internetu korzysta 85%. Wśród urządzeń teleinformatycznych znajdujących się w gospodarstwach domowych wymienianych przez indywidualnych użytkowników systemów informatycznych są także drukarki, skanery, telefon stacjonarny oraz kamera.

Blisko 90% użytkowników systemów informatycznych w gospodarstwach domowych łączy się z Internetem za pośrednictwem stałego łącza, to jest telewizji kablowej, sieci osiedlowej, Neostrady. Z sieci bezprzewodowej

korzysta 24% respondentów, najmniej osób ma dostęp do Internetu przez modem/dial up (linia telefoniczna TP SA, Dialog, ISDN, SDI).

W gospodarstwach domowych konto na komputerze jest współdzielone najczęściej z dwoma osobami (aż 42% wskazań). Z jedną osobą współdzieli konto 22% respondentów. 20% indywidualnych użytkowników systemów informatycznych w badanych gospodarstwach domowych korzysta z konta bez współudziału innych osób. Ze względu na bezpieczeństwo systemów informatycznych należy podkreślić, że współdzielenie konta na komputerze przez więcej niż 3 osoby nie powinno być praktykowane. Wyniki przeprowadzonych badań w gospodarstwach domowych wskazują jednakże na występowanie tego typu sytuacji – 16% wskazań respondentów. Należałoby zwiększyć świadomość zagrożenia, jakie za sobą niesie współdzielenie konta na komputerze wśród użytkowników indywidualnych.

Systemy operacyjne Microsoft Windows XP oraz Microsoft Windows Vista są najczęściej wykorzystywanymi systemami przez użytkowników w gospodarstwach domowych. Poczta elektroniczna i edytory tekstów (95% oraz 91% wskazań) to główne technologie informatyczne, jakimi posługują się w gospodarstwach domowych respondenci. Powyżej 70% użytkowników pracuje na komputerze w domu w arkuszu kalkulacyjnym, a także ma do czynienia z dokumentami HTML – rys. 4.



**inne (C++/media odtwarzacze, programy multimedialne do odtwarzania muzyki i filmów)*

Rys. 4. Technologie i systemy informatyczne wykorzystywane w gospodarstwach domowych przez respondentów (w %)

Na forach dyskusyjnych za pośrednictwem komputera w gospodarstwie domowym udziela się 65% użytkowników.

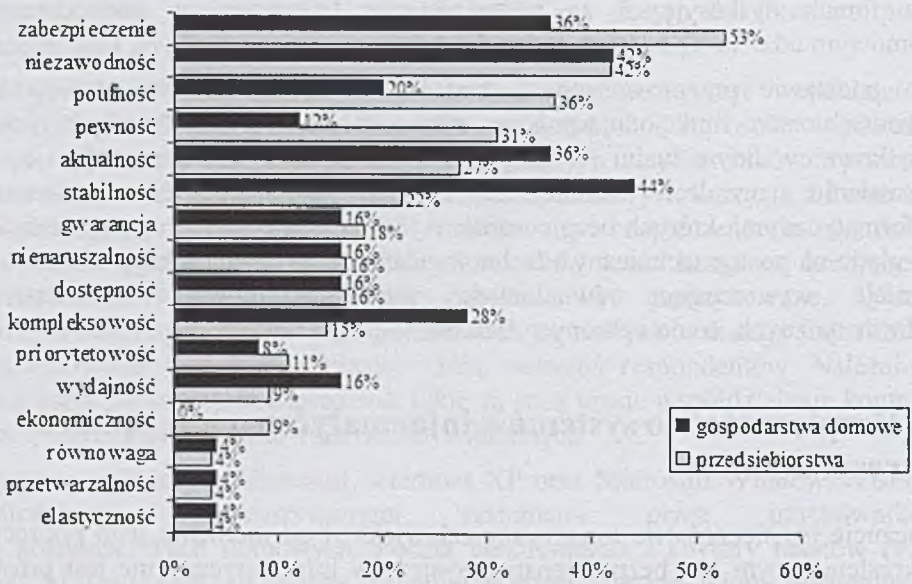
Na podstawie przeprowadzonych badań wynika, iż zarówno pracownicy przedsiębiorstw funkcjonujących na rynku w odmiennych branżach, jak i użytkownicy indywidualni (w gospodarstwach domowych) korzystają i są w posiadaniu urządzeń teleinformatycznych wspomaganych systemami informatycznymi, których bezpieczeństwo jest narażone na szereg zagrożeń. Ze względu na postęp techniczny i technologiczny oraz rozwój wiedzy nie zawsze istnieje wystarczająca świadomość wśród użytkowników systemów informatycznych, co do ochrony systemów.

3. Bezpieczeństwo systemów informatycznych w opinii respondentów

Poczucie bezpieczeństwa towarzyszy człowiekowi od momentu jego poczęcia. Określenie czym jest bezpieczeństwo systemów informatyczny nie jest proste, a w literaturze można spotykać się z różnym podejściem. Nie zmienia to faktu, iż trzeba mieć zawsze świadomość, że bezpieczeństwa nie da się osiągnąć w stu procentach, będzie to bowiem obszar podatny na ciągłe zagrożenie.

W dobie cybercywilizacji należałoby odpowiedzieć na pytania „Co oznacza bezpieczeństwo dziś?" i „Czy podejście do bezpieczeństwa jest takie samo jak przed kilku laty, czy zmienia się?" Niewątpliwie problem bezpieczeństwa nasila się wraz z postępem technicznym, naukowym, ale nie tylko. Wiąże się przede wszystkim z pojawianiem się nowych zagrożeń, których tempo występowania jest zaskakujące. Potrzeba wzmożonej ochrony systemów informatycznych rodzi zmiany w podejściu do kwestii bezpieczeństwa.

Na podstawie opinii pracowników przedsiębiorstw oraz użytkowników systemów informatycznych gospodarstw domowych w badaniu podjęto próbę określenia cech charakteryzujących bezpieczeństwo systemów informatycznych. Zdaniem 53% respondentów – użytkowników indywidualnych, cechą najbardziej charakteryzującą bezpieczeństwo to zabezpieczenie. Podobnie uważają pracownicy polskich przedsiębiorstw (48%). Niezawodność (42%), poufność (36%) i pewność (31%) są cechami wymienianymi kolejno przez użytkowników gospodarstw domowych. Z kolei w przypadku pracowników przedsiębiorstw cechami charakterystycznymi dla bezpieczeństwa wskazywanymi częściej to stabilność (44%), niezawodność i aktualność (36%) – rys. 5.



Rys.5 Cechy charakteryzujące bezpieczeństwo w opinii respondentów (w %)

Zadziwiające jest to, że respondenci przypisując bezpieczeństwu różne cechy dokonali wyboru tych, które z bezpieczeństwem systemów informatycznych nie mają wiele wspólnego, na przykład aktualność, kompleksowość, priorytetowość.

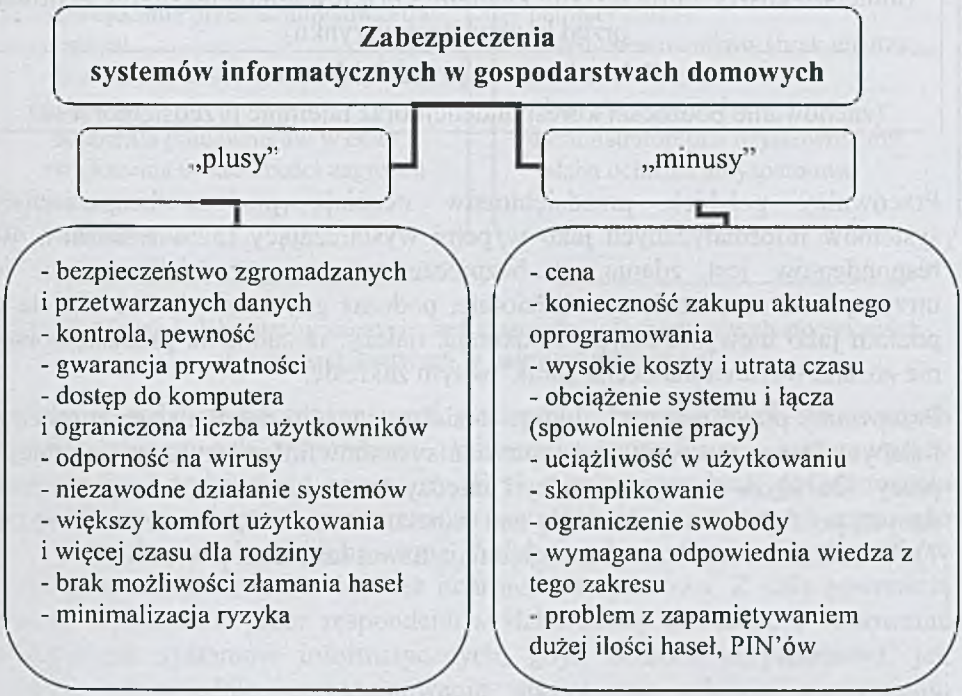
Respondenci reprezentujący przedsiębiorstwa uważają, że bezpieczeństwo systemów informatycznych ma obecnie strategiczne znaczenie dla przedsiębiorstwa. Natomiast użytkownicy indywidualni wyrażają przekonanie, że trzeba zabezpieczać systemy informatyczne, a za główne powody podają ochronę danych, informacji i prywatności – tab. 1.

Tab. 1. Powody zabezpieczania systemów informatycznych w gospodarstwach domowych w opinii respondentów

zachowanie prywatności
ochrona danych/bezpieczeństwo informacji (przed dostępem innych, obcych użytkowników)
dostęp do Internetu (korzystanie z usług bankowości elektronicznej)
ochrona sprzętu (przed zniszczeniem, zepsuciem, awarią, atakami z zewnątrz)
ochrona przed zagrożeniami (wirusami, robakami, dialerami)

Respondenci użytkujący systemy informatyczne w gospodarstwach domowych zostali poproszeni o zestawienie plusów i minusów zabezpieczeń systemów.

Bezpieczeństwo związane z gromadzeniem i przetwarzaniem informacji, jak również kontrola, pewność oraz gwarancja prywatności to najczęściej pojawiające się pozytywne wskazania dla zabezpieczeń systemów informatycznych (rys. 6). Istotnym plusem przy zabezpieczaniu systemów dla użytkowników indywidualnych są odporność na wirusy oraz zwiększony komfort pracy, co nie pozostaje bez wpływu na wolny czas dla rodziny, który zostaje dzięki temu wydłużony. Największym minusem stosowania zabezpieczeń systemów informatycznych dla respondentów jest cena oprogramowania, czyli poniesione koszty z tego tytułu i utrata czasu. Za mniej pozytywne uznają także użytkowanie, które ich zdaniem jest uciążliwe i skomplikowane.



Rys. 6. Pozytywne i negatywne strony zabezpieczeń systemów informatycznych stosowanych w gospodarstwach domowych w opinii respondentów

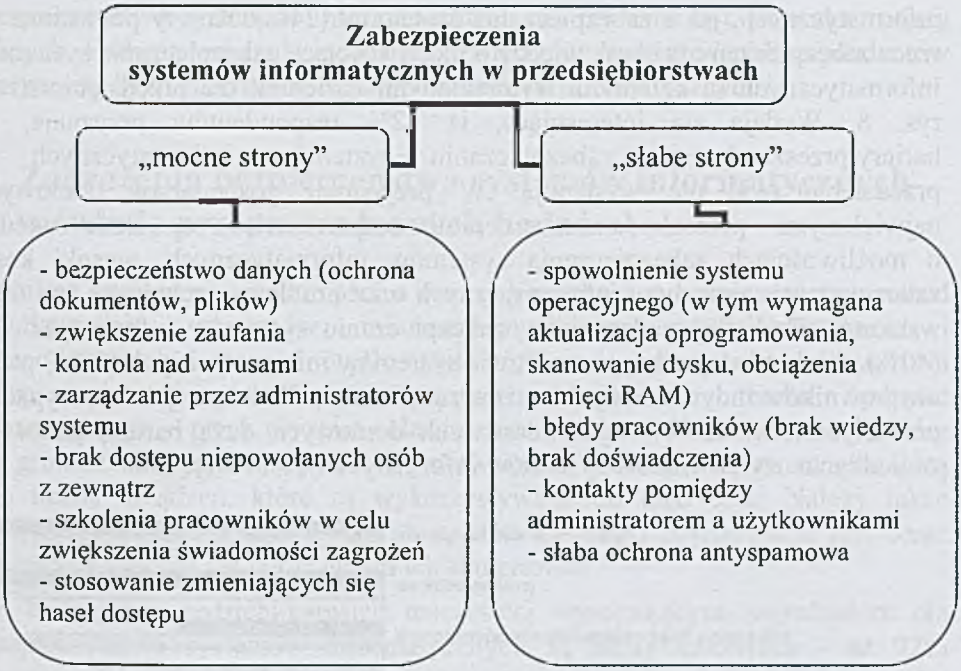
Zasadność stosowania zabezpieczeń systemów informatycznych w przedsiębiorstwach dostrzegają również pracownicy. Do najważniejszych zaliczają ochronę poufnych danych przedsiębiorstwa przed kradzieżą, skasowaniem, atakiem z zewnątrz (tab. 2.). Korzystne rozwiązanie w zabezpieczaniu systemów informatycznych upatrują między innymi w ograniczeniu dostępu osób trzecich do informacji, podkreślają także aspekty ekonomiczne oraz wpływ na dobro przedsiębiorstwa.

Tab. 2. Powody zabezpieczania systemów informatycznych w przedsiębiorstwach w opinii respondentów

ochrona danych/własności intelektualnej (przed kradzieżą, utratą danych bieżących, archiwalnych, przed posłuchaniami, szpiegowaniem i przejęciem)
ograniczenie dostępu osób niepowołanych (zagrożenie integracją osób trzecich, manipulacja z zewnątrz)
minimalizacja ryzyka
aspekty ekonomiczne (mniejsze koszty, obrót dużymi kwotami i przelewami pieniężnymi, ochrona przed konkurencją na rynku)
dobro i interes przedsiębiorstwa (zachowanie poufności korespondencji oraz tajemnic przedsiębiorstwa)

Pracownicy polskich przedsiębiorstw oceniają poziom bezpieczeństwa systemów informatycznych jako w pełni wystarczający (52% wskazań). 40% respondentów jest zdania, iż bezpieczeństwo w przedsiębiorstwach jest utrzymywane na przeciętnym poziomie, podczas gdy zaledwie 8% określa ten poziom jako niewystarczający.Podkreślić należy, iż żadnemu przedsiębiorstwu nie została wystawiona ocena „brak" w tym zakresie.

Pracownicy przedsiębiorstw, biorąc udział w badaniu dokonali oceny mocnych i słabych stron stosowania zabezpieczeń systemów informatycznych w miejscu pracy. Do stron mocnych zaliczyli między innymi ochronę i bezpieczeństwo danych poufnych firmy, kontrolę nad wirusami oraz zwiększenie zaufania (rys. 7). W ocenie mocnych stron uwzględniają nawet szkolenia pracowników.

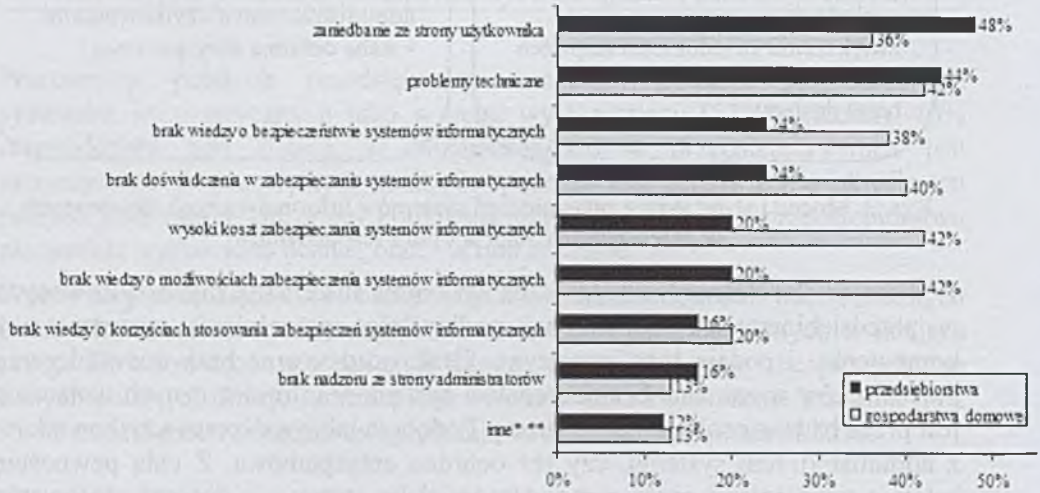


Rys. 7. Mocne i słabe strony zabezpieczeń systemów informatycznych stosowanych w przedsiębiorstwach w opinii respondentów

Wyliczając słabe strony zabezpieczeń systemów informatycznych stosowanych w przedsiębiorstwach respondenci podkreślają spowolnienia systemu/pracy komputerów i podają jego przyczyny. Brak wiedzy oraz brak doświadczenia pracowników w zakresie bezpieczeństwa systemów informatycznych uznawane jest przez badane osoby za słabą stronę. Podobnie jak współpraca użytkowników z administratorem systemu, czy też ochrona antyspamowa. Z całą pewnością ostatnie wymienione przez respondentów słabe strony nie dotyczą stosowania zabezpieczeń systemów informatycznych, gdyż ochrona antyspamowa, jak również odpowiednie oprogramowanie służyć ma właśnie zapewnieniu odpowiedniego poziomu bezpieczeństwa systemów informatycznych. Błędów pracowników wynikających z niewiedzy na temat bezpieczeństwa także nie można zaliczyć do słabych stron zastosowania zabezpieczeń. Podobnie ma się rzecz, jeśli chodzi o kontakty z administratorem. Na podstawie udzielonych odpowiedzi można sądzić, iż respondenci nie do końca rozumieją i są świadomi stosowania zabezpieczeń systemów informatycznych w miejscu pracy.

Obie grupy respondentów wskazują na występowanie barier lub przeszkód przy zabezpieczaniu systemów informatycznych. Największymi barierami, z jakimi spotykają się pracownicy w badanych przedsiębiorstwach są zaniedbania ze strony pracownika/użytkownika (48% wskazań) oraz problemy techniczne (44%). Inną dostrzeganą przeszkodą w przedsiębiorstwach jest brak wiedzy pracowników i to zarówno z zakresu bezpieczeństwa systemów

informatycznych, jak i zabezpieczania systemów (24%). Koszty poniesione na rzecz zabezpieczeń oraz brak wiedzy o możliwościach zabezpieczania systemów informatycznych są kolejnymi wymienianymi barierami dla przedsiębiorstw – rys. 8. Wydaje się interesujące, iż 12% respondentów przyznaje, że bariery/przeszkody w zabezpieczaniu systemów informatycznych w przedsiębiorstwie nie występują. W przypadku gospodarstw domowych największymi przeszkodami w opinii respondentów są: brak wiedzy o możliwościach zabezpieczenia systemów informatycznych, wysoki koszt zabezpieczania systemów informatycznych oraz problemy techniczne (po 42% wskazań). Brak doświadczenia w zabezpieczaniu systemów informatycznych (40%) i brak wiedzy o bezpieczeństwie systemów informatycznych (38%) przez użytkowników indywidualnych są znaczącą barierą. Podobnie jak w przypadku przedsiębiorstw, tak i w gospodarstwach domowych dużą barierą pozostają zaniedbania użytkowników systemów informatycznych (36%).



*inne * (właściwie nie ma barier, raczej nie ma takich barier)*
*** (brak czasu, czasem trzeba zmienić swoje nawyki)*

Rys. 8. Bariery i przeszkody zabezpieczeń systemów informatycznych występujące w przedsiębiorstwach i gospodarstwach domowych w opinii respondentów (w%)

Podejście do kwestii bezpieczeństwa systemów informatycznych w gospodarstwach domowych i przedsiębiorstwach jest jednoznaczne, wyniki przeprowadzonych badań pokazują, że problem ten jest istotny. Kradzież, utrat danych, dostęp osób niepowołanych, zniszczenie chociażby przez wirusy, zmienia wąskie myślenie o bezpieczeństwie systemów informatycznych. Dzisiaj o bezpieczeństwie mówi się w kontekście teleinformatyki, której postęp i rozwój wywołuje określone skutki dla większości prowadzonej działalności gospodarczej i nie pozostaje bez wpływu na funkcjonowanie gospodarstw domowych. Jednocześnie niewiedza użytkowników odpowiedzialnych za

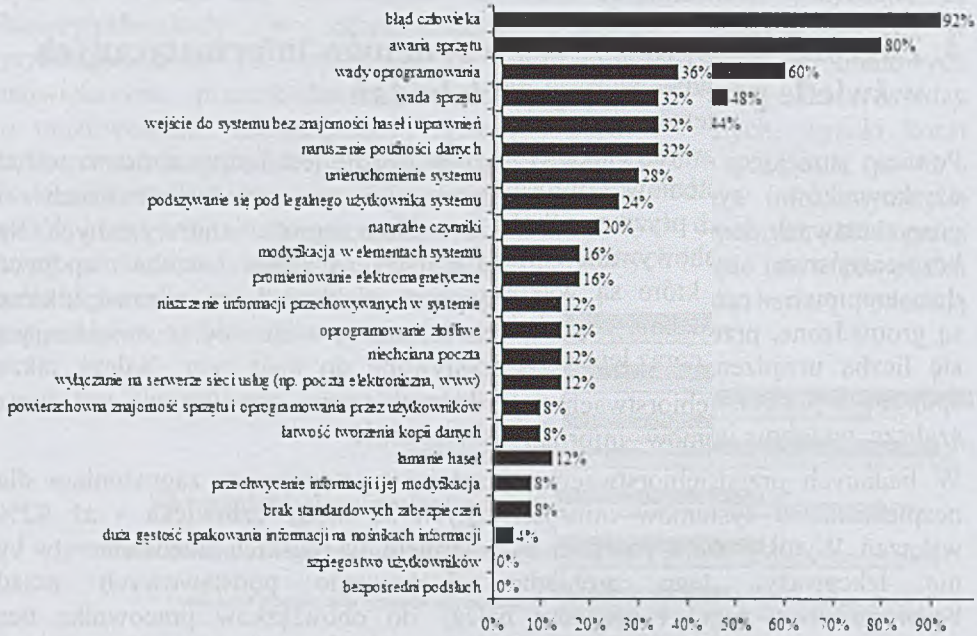
bezpieczeństwo systemów informatycznych nie powinna być powodem utraty danych. Badania wskazują jednak na pewien brak wiedzy wśród respondentów dotyczącej bezpieczeństwa i ochrony systemów informatycznych.

4. Zagrożenia bezpieczeństwa systemów informatycznych w świetle przeprowadzonych badań

Pomimo istniejącej świadomości o tym, jak ważne jest bezpieczeństwo wśród użytkowników systemów informatycznych w przedsiębiorstwach i gospodarstwach domowych ciągle rośnie problem zagrożenia utraty danych. Na bezpieczeństwo systemów informatycznych obecnie trzeba spojrzeć dwustopniowo – przez pryzmat ochrony coraz większej ilości informacji, które są gromadzone, przechowywane i przetwarzane, co wiąże się ze zwiększającą się liczbą urządzeń, które są wykorzystywane do tego celu. Należy także spojrzeć z punktu widzenia zagrożeń, których tempo powstawania jest coraz szybsze, podobnie jak coraz lepsza ich skuteczność.

W badanych przedsiębiorstwach najczęściej występującym zagrożeniem dla bezpieczeństwa systemów informatycznych są błędy człowieka - aż 92% wskazań. Wynik badania powinien być alarmem dla polskich przedsiębiorstw by nie lekceważyć tego problemu. Zachowanie podstawowych zasad bezpieczeństwa, reguł i procedur należy do obowiązków pracownika bez względu na zajmowane w hierarchii przedsiębiorstwa stanowisko. Sprzęt komputerowy, czy inne urządzenia teleinformatyczne wspomagane przez systemy informatyczne bywają zawodne, o czym deklarują respondenci - 80% wskazań dla awarii sprzętu, nie zmienia to jednak faktu, że to użytkownicy posługują się sprzętem i nim zarządzają – rys. 9 Innym dostrzeganym przez pracowników zagrożeniem jest wada oprogramowania (60%). Można przypuszczać, że jest to raczej wynik niezadowolenia z użytkowania oprogramowania a niżeli zagrożenie dla bezpieczeństwa systemów informatycznych. Bardziej zaskakujące jest to, iż respondenci nie postrzegają łatwości tworzenia kopii danych (12% wskazań), braku standardowych zabezpieczeń w przedsiębiorstwie (zaledwie 8% wskazań), czy też przechwycenie i modyfikowanie informacji (8%) jako większe, a nawet potencjalne źródło zagrożenia dla systemów informatycznych, a wskazują na przykład na naturalne czynniki – 32 %. Jest to wynik braku świadomości, co jest większym zagrożeniem dla systemów informatycznych. A przecież każdy z pracowników posługuje się informacją, korzysta z niej, przetwarza na określone potrzeby, dlaczego więc zapomina, że sposób, w jaki użytkuje dane – zasoby przedsiębiorstwa, jest narażony na zniszczenie, kradzież lub utratę? Co więcej, za pośrednictwem pamięci zewnętrznych (pamięć flash, USB) dane są przenoszone poza obszar przedsiębiorstwa, a pracownicy opuszczający miejsce pracy nie są poddawani kontroli. Wszystko to przy braku odpowiednich zabezpieczeń systemów informatycznych stanowi dla nich

zagrożenie. Przedsiębiorstwa powinny kierować się nie tylko wygodą w użytkowaniu systemów informatycznych, ale większą uwagę skupić na ochronie zasobów informacyjnych i wdrażać konkretne systemy zabezpieczające.



Rys. 9. Rodzaje zagrożeń bezpieczeństwa systemów informatycznych występujących w przedsiębiorstwach w opinii respondentów (w%)

W gospodarstwach domowych najczęściej spotykanymi zagrożeniami systemów informatycznych zdaniem respondentów są niechciana poczta (85%), awaria sprzętu (76%), wada sprzętu oraz złośliwe oprogramowanie (53%) – rys. 10. Użytkownicy indywidualni, podobnie jak pracownicy przedsiębiorstw nie zwracają uwagi na istotniejsze źródła zagrożeń systemów informatycznych.

Rys. 10. Rodzaje zagrożeń bezpieczeństwa systemów informatycznych występujących w gospodarstwach domowych w opinii respondentów (w%)

Bardzo często w gospodarstwie domowym użytkownicy systemów informatycznych spotykają się ze spamem, wirusami, atakami koni trojańskich i robaków. Sporadycznie mają do czynienia z zagrożeniami dotyczącymi sprzętu i zasilania, jak zaniki i spadki napięcia oraz zakłócenia elektromagnetyczne. Rzadko występują także zagrożenia związane z atakami na www, skanowaniem firewalla i skanowaniem hosta.

Respondenci reprezentujący świat biznesowy również wskazują na bardzo częste narażenie ze strony ataków wirusów, robaków oraz spamu. W sporadycznych przypadkach zagrożenie występuje po stronie spadków i zaników napięć.

Jak wynika z przeprowadzonych badań na zagrożenia i ataki z zewnątrz narażeni są nie tylko użytkownicy systemów informatycznych w gospodarstwach domowych, ale i przedsiębiorstwa – nie ma tutaj wyjątku od reguły.

5. Środki bezpieczeństwa systemów informatycznych

Ze względu na ogromne ilości przetwarzanych informacji do skutecznej ochrony zasobów przedsiębiorstwa niezbędnym jest stosowanie odpowiednich środków bezpieczeństwa systemów informatycznych. Rosnące zapotrzebowanie na korzystanie z danych poszerza w szybkim tempie także bazy gromadzonych informacji, a co za tym idzie – rośnie zapotrzebowanie na ochronę i środki bezpieczeństwa.

W badanych przedsiębiorstwach najczęściej stosowanym środkiem bezpieczeństwa systemów informatycznych są identyfikacja i uwierzytelnienie użytkownika (92% wskazań). Na stosowanie w przedsiębiorstwie haseł

zabezpieczających systemy informatyczne wskazuje aż 88% respondentów. Ochrona antywirusowa i zapory ogniowe oraz autoryzacja są wykorzystywane do ochrony systemów informatycznych w przedsiębiorstwie zdaniem 80% pracowników. Kompresję plików danych stosuje zaledwie 44% badanych przedsiębiorstw. W opinii tylko 8% pracowników w badanych przedsiębiorstwach realizowany jest podpis cyfrowy. Co ciekawe zdaniem 24% respondentów przedsiębiorstwa wymieniają klucze szyfrujące informacje (rys. 11).

inne (brak prawa administratora na komputerach użytkowników)*

Rys. 11. Środki bezpieczeństwa systemów informatycznych najczęściej stosowane w przedsiębiorstwach w opinii respondentów (w %)

Z kolei w gospodarstwach domowych użytkownicy indywidualni wykorzystują głównie takie środki bezpieczeństwa systemów informatycznych, jak ochrona antywirusowa i zapory ogniowe (84% wskazań), hasła (80%), identyfikację i uwierzytelnienie użytkownika (71%) – rys. 12. Użytkownicy chroniąc przetwarzane zasoby informacji w komputerach domowych w większym stopniu przywiązują wagę do środka, jakim są programy archiwizująco-pakujące do kompresji plików (31%).

Rys. 12. Środki bezpieczeństwa systemów informatycznych najczęściej stosowane w gospodarstwach domowych w opinii respondentów (w %)

Użytkownicy indywidualni dokonują także częściej backupu (tworzenie kopii zapasowej) za pomocą umieszczenia plików danych na nośnikach: płyty CD, pendrive (80%). Innym sposobem jest umieszczanie danych w tak zwanej przestrzeni wirtualnej, czyli serwery ftp i serwisy archiwizujące – tak postępuje 27% respondentów.

Interesujący jest także sposób tworzenia oraz zapisywania haseł dostępu do kont pocztowych systemów innych programów, z których korzystają użytkownicy na komputerach w gospodarstwie domowym. 55% respondentów używa do tworzenia haseł co najmniej ciągu 8 znaków, a cyfr i znaków specjalnych – 33%. W tworzonych hasłach użytkownicy umieszczają i duże i małe litery (aż 25% wskazań). Aż 22% badanych użytkowników tworzy hasła posługując się datami urodzenia, imionami i nazwiskami bliskich osób. Co najciekawsze, respondenci przyznają, że podczas tworzenia haseł do kont pocztowych i innych programów dokonują losowości znaków – tab. 3.

Tab. 3. Sposoby tworzenia haseł do kont pocztowych i innych programów w opinii respondentów (w%)

Wyszczególnienie	Użytkownicy w gospodarstwach domowych	Pracownicy przedsiębiorstw
ciągu co najmniej 8 znaków	55	57
cyfry i znaki specjalne	33	48
małych i dużych liter	25	48
nazwisk i imion bliskich osób	22	13
dat urodzenia, innych	22	13
losowość znaków	20	17
ciągu poniżej 6 znaków	13	17
elementy adresów	5	9
ciągu kolejnych znaków z klawiatury	2	9
nazwy komputera	2	4
numerów telefonów, samochodu, dowodu osobistego itp.	2	-

W przypadku pracowników przedsiębiorstw sytuacja ma się podobnie, tworzenie haseł do kont pocztowych odbywa się przy wykorzystaniu ciągu co najmniej 8 znaków (prawie 60% wskazań; tab. 3) oraz cyfr i znaków specjalnych oraz małych i dużych cyfr (48%).

Hasła dostępu do kont i innych programów pracownicy przedsiębiorstw przechowują w postaci zaszyfrowanej przy użyciu kluczy – 22% wskazań. Nikt nie zadeklarował, iż hasła zapisuje na dysku komputera, natomiast 17% respondentów zapisuje hasła na kartce i chowa do szuflady lub zapisuje w telefonie komórkowym. Tego typu postępowanie nie jest możliwe do przyjęcia, gdyż ułatwia dostęp osobom nieupoważnionym do informacji, której jedynym „właścicielem" powinien być użytkownik, nie wspominając już o zwiększeniu ryzyka zagrożenia. Innym praktykowanym przez pracowników przedsiębiorstw sposobem na zachowanie utworzonego hasła jest zapamiętywanie haseł „we własnej pamięci".

Sposób przechowywania haseł wśród użytkowników w gospodarstwach domowych nie jest najlepszym rozwiązaniem, ponieważ respondenci hasła zapisują na kartce i chowają do szuflady lub pod klawiaturę. Skuteczność tego sposobu uzależniona jest oczywiście od sytuacji, czyli od tego czy użytkownik ma dostęp do własnego komputera, z którego korzysta tylko on sam.

Wykorzystanie systemów informatycznych, jakimi posługują się, czy to pracownicy w przedsiębiorstwach, czy też przez indywidualni użytkownicy w gospodarstwie domowym wiąże się z ochroną informacji, która wymaga zastosowania odpowiednich środków zabezpieczających. Ciągły rozwój zagrożeń, z jakimi zmagają się przedsiębiorstwa i użytkownicy indywidualni

wymusza stosowanie coraz większej liczby środków, co nie pozostaje bez wpływu na koszty. Przedsiębiorstwa korzystają z usług zewnętrznych – wyspecjalizowanych dostawców środków zapewniających bezpieczeństwo systemów informatycznych (ponad 60% wskazań) lub kupują gotowe produkty w sklepie komputerowym (30%). W przedsiębiorstwach środkami bezpieczeństwa systemów dysponują administratorzy, którzy konfigurują i monitorują pracę serwerów, zajmują się oprogramowaniem i są odpowiedzialni za monitoring stanowiska komputerowego każdego pracownika. Niestety pośród dobrych przykładów sposobów zaopatrywania się przedsiębiorstw w środki ochrony systemów informatycznych znajdują się i takie, które wskazują na zakup programów bezpłatnie dołączanych na płytach do czasopism (4% wskazań).

Wśród użytkowników systemów informatycznych w gospodarstwach domowych przeważają sposoby zaopatrywania się w środki za pośrednictwem bezpłatnych programów załączanych do czasopism (20% wskazań). O ile w tym przypadku jest to sposób dopuszczalny ze względu na możliwość poznania i wypróbowania produktu przed podjęciem decyzji o zakupie, tak dla przedsiębiorstw jest to rozwiązanie niekorzystne. Oprócz tego popularnością wśród gospodarstw domowych cieszy się sposób zaopatrywania w darmowe oprogramowanie za pośrednictwem Internetu (pobieranie wersji trial lub całości, skanowanie online). Najczęściej wymienianymi przez użytkowników programami antywirusowymi są programy: Avast, Kaspersky Trial, AVG, Panda i Comodo Internet Security.

Z przeprowadzonych badań wynika, iż poziom stosowania przez przedsiębiorstwa środków w celu ochrony systemów informatycznych znajduje się na dobrym poziomie, a ich liczba i różnorodność także wskazują na wysoką świadomość pracowników, w tym i pracodawców, co do zapewnienia bezpieczeństwa informacji. Przy czym, przy bliższej analizie wyników badań trzeba wziąć pod uwagę, iż stan – obraz przedstawiany przez pracowników przedsiębiorstw nie jest obiektywny. Wynika on zapewne z niewiedzy o możliwościach ochrony systemów informatycznych, jak również zastosowaniu poszczególnych środków bezpieczeństwa.

6. Przyszłość bezpieczeństwa systemów informatycznych – wyniki badań

Wzrost zagrożenia i ochrony bezpieczeństwa systemów infromatycznych rodzi szereg pytań, między innymi: „Jaka przyszłość czeka systemy informatyczne zarówno w przedsiębiorstwach, jak i gospodarstwach domowych?", „W jakim kierunku bezpieczeństwo systemów informatycznych będzie się rozwijać?", „Czy postęp w rozwoju środków zabezpieczających systemy informatyczne zdoła powstrzymać rozwój zagrożeń?"

Pracownicy polskich przedsiębiorstw uważają, że w ciągu najbliższych 2-3 lat będzie konieczna ochrona bezpieczeństwa systemów informatycznych. W uzasadnieniu respondenci wskazali potencjalne przyczyny: rozwój systemów szpiegowskich oraz zwiększenie liczby hakerów, rozwój technologii, pojawianie się nowych systemów. Wzrost zagrożeń pochodzących z Internetu, ciągły wzrost i zapotrzebowanie na dane informacyjne, wzrost znaczenia systemów informatycznych oraz edukacja i zdobywanie umiejętności - to kolejne przesłanki, które zdaniem pracowników należy mieć na uwadze. Co najistotniejsze dla samych przedsiębiorstw – „nie mogą sobie pozwolić na błędy, ponieważ w przeciwnym wypadku przestaną istnieć na rynku". W opinii respondentów reprezentujących przedsiębiorstwa w okresie najbliższych 2-3 lat na zwiększenie bezpieczeństwa systemów informatycznych będą miały wpływ elementy otoczenia przedsiębiorstwa. Do elementów tych pracownicy zaliczają między innymi współpracę z kontrahentami z branży informatycznej, pozyskanie lepszej kadry specjalistów z zakresu bezpieczeństwa oraz rozwój oprogramowania antywirusowego.

Użytkownicy korzystający z systemów informatycznych w gospodarstwach domowych są zgodni, iż w przyszłości wzrośnie zapotrzebowanie na zabezpieczanie systemów informatycznych ze względu na postęp techniki, rozwój nowych technologii informatycznych, wzrost zagrożeń. Zdaniem indywidualnych użytkowników zwiększy się liczba stanowisk komputerowych, w gospodarstwach domowych pojawiać się będzie coraz więcej sprzętu teleinformatycznego, a wraz z nim przybędzie więcej osób z niego korzystających. Systemy informatyczne, które obecnie są mniej dostępne staną się bardziej upowszechnione, a informatyzacja obejmie nowe dziedziny życia. Respondenci zauważają, iż zapotrzebowanie na ochronę bezpieczeństwa będzie dotyczyło także społeczeństwa - zagrożenia będą miały wpływ w większym zakresie na dzieci, na przykład ze względu na zwiększenie liczby stron pornograficznych, czy też pojawienie się coraz bardziej ekspansywnych reklam. Wraz ze wzrostem postępu nauki wzrastać będzie świadomość bezpieczeństwa i zagrożeń wśród użytkowników domowych, co znajdzie przełożenie na zwiększenie zapotrzebowania na wiedzę w tym zakresie oraz na środki ochrony bezpieczeństwa systemów. Przy rozwoju techniki, towarzyszyć będzie rozwój różnych form i odmian wirusów, powstaną nowsze sposoby włamywania – hacking, phishing, snifing. Powstanie także więcej sieci internetowych, wzrośnie zapotrzebowanie na elektroniczną formę komunikacji.

7. Podsumowanie

Podejście do bezpieczeństwa systemów informatycznych wśród pracowników przedsiębiorstw i użytkowników gospodarstw domowych jest podobne, a mianowicie istnieje przekonanie o ważności tego problemu. Co prawda respondenci nie do końca wykazali się wiedzą czym jest bezpieczeństwo

systemów informatycznych, ale w gruncie rzeczy nie zawsze chodzi o to by wiedzieć „czym to jest", ale „jak to chronić". W obu przypadkach świadomość zagrożeń bezpieczeństwa systemów informatycznych nie różni się znacząco. Pracownicy przedsiębiorstw w mniejszym stopniu dostrzegają większe źródła zagrożeń, jakie występują w ich środowisku pracy. Nie mają wątpliwości, że znaczenie ochrony bezpieczeństwa jest strategicznym punktem w działalności przedsiębiorstwa i wyrażają w większości przekonanie, że za bezpieczeństwo systemów informatycznych odpowiedzialni są wszyscy pracownicy firmy (68% wskazań). Nie mniej jednak uważają także, że za bezpieczeństwo systemów odpowiedzialni są tylko informatycy i administratorzy systemów – 26% wskazań.

Z przeprowadzonych badań wynika, iż wiedza respondentów o bezpieczeństwie systemów informatycznych ciągle jest jeszcze niewystarczająca. Z jednej strony stosują środki bezpieczeństwa systemów informatycznych, potrafią nazwać zagrożenia, z jakimi spotykają się na co dzień, mają też świadomość skutków, jakie wywołują, lecz nie do końca wiedzą na czym polega zastosowanie środków ochrony bezpieczeństwa systemów informatycznych. Co więcej, brak wiedzy z zakresu bezpieczeństwa respondenci podkreślają i wykazują jako przyczynę zagrożenia dla systemów informatycznych. Polskie przedsiębiorstwa mają pozytywny stosunek do ochrony bezpieczeństwa systemów informatycznych, ale sytuacja, jaką przedstawiają pracownicy trzeba uznać za subiektywną. Z przeprowadzonych badań wynika, iż większą wagę do bezpieczeństwa informacji przywiązują użytkownicy indywidualni w gospodarstwach domowych, niż pracownicy przedsiębiorstw. Być może jest to efekt podejścia do odpowiedzialności za dane informacyjne i jednocześnie sprzęt, jakim dysponuje użytkownik, mając na względzie swoje dobro. Przyszłość bezpieczeństwa systemów informatycznych respondenci upatrują głównie w rozwoju techniki i technologii informatycznych, są zgodni, iż wraz ze wzrostem zagrożeń będzie rosło zapotrzebowanie na środki ochrony bezpieczeństwa. Tym samym nie jest możliwe by całkowicie wyeliminować zagrożenia, jakie się pojawiają. Znaczenie bezpieczeństwa systemów informatycznych będzie rosło, a za jego ochronę będzie zawsze odpowiedzialny człowiek.

LITERATURA

Rozdział powstał wyłącznie w oparciu o własne opracowania Autorki.

Rozdział 2

Prawne aspekty bezpieczeństwa systemów teleinformatycznych i ochrony infrastruktury krytycznej w służbach żeglugi powietrznej

Sebastian Burgemejster
CISA, CGAP, CCSA

Polska Agencja Żeglugi Powietrznej

Streszczenie

W pracy scharakteryzowano wpływ działalności służb żeglugi powietrznej na bezpieczeństwo w ruchu lotniczym oraz dokonano opisu stosowanych rozwiązań prawnych w zapewnieniu bezpieczeństwa systemów teleinformatycznych i ochrony infrastruktury krytycznej w służbach żeglugi powietrznej.

1. Wstęp

Transport lotniczy stanowi jedno z najważniejszych ogniw krajowego i międzynarodowego transportu. Jest to zarówno najszybszy, jak i najbezpieczniejszy sposób na przemieszczenie ludzi oraz towarów w dowolne miejsce na Ziemi. Pomimo zajmowania pierwszego miejsca w zakresie bezpieczeństwa, to właśnie katastrofy lotnicze poprzez nieuchronność śmierci, olbrzymie straty materialne oraz finansowe pojawiają się zawsze w czołówkach mediów. Dlatego też branża lotnicza podlega stałej kontroli pod względem zapewnienia odpowiednich środków zabezpieczeń przeciwko aktom terroru, jak i awariami urządzeń lub błędem człowieka.

Dla przeciętnego obywatela ruch lotniczy składa się z przewoźników lotniczych oraz portów lotniczych. Jedynie wprawny obserwator dostrzega rolę, jaką w bezpieczeństwie, ciągłości oraz płynności komunikacji lotniczej spełniają służby ruchu lotniczego (Air Navigation Service Provider - ANSP). Służby żeglugi powietrznej składają się z kilku podstawowych obszarów działania, którymi m.in. są: kontrola ruchu lotniczego, służba informacji powietrznej, służba

alarmowa, planowanie przepływu ruchu lotniczego oraz koordynacja zajętości przestrzeni powietrznej. Wszystkie w/w zadania wymagają pełnego wykorzystania potencjału technicznego, zespołu wysoko wykwalifikowanych specjalistów oraz bardzo szczegółowych procedur postępowania. Z uwagi na wysoce złożony charakter działań oraz istotny wpływ na bezpieczeństwo publiczne ANSP poddawane są bardzo rygorystycznym regulacjom oraz procesowi nieustannej certyfikacji/recertyfikacji świadczonych usług. Dodatkowo poza prowadzeniem działań w zakresie cywilnego zarządzania przestrzenią powietrzną wchodzącego w zakres bezpieczeństwa publicznego, ANSP prowadzą współpracę z organami wojskowymi w zakresie bezpieczeństwa państwa zapewniając infrastrukturę oraz wyspecjalizowane służby mogące zostać wykorzystane jako krytyczny zasób w trakcie konfliktu zbrojnego. [14-16]

2. Funkcjonowanie służb żeglugi powietrznej

Podstawowymi zadaniami ANSP jest zarządzanie przestrzenią powietrzną oraz zapewnienie wysokiego poziomu bezpieczeństwa statków powietrznych. W ramach tych działań służby ruchu lotniczego realizują następujące zadania [14-16]:

- Zarządzanie przestrzenią powietrzną (Air Space Management - ASM) – jest to projektowanie, wdrażanie, zarządzanie i alokacja przestrzeni powietrznej dla wszystkich użytkowników tej przestrzeni,
- Służby Ruchu Lotniczego (Air Traffic Services - ATS):

  Głównymi zadaniami służb ruchu lotniczego jest: zapobieganie zderzeniu się statków powietrznych, usprawnianie, zarządzanie i utrzymywanie uporządkowanego przepływu ruchu lotniczego, zbieranie informacji o zagrożeniu statków powietrznych i dostarczanie ich organom odpowiedzialnym za uruchomienie systemu ratownictwa lotniczego oraz współdziałanie z w/w organami w czasie prowadzenia akcji ratowniczych oraz dostarczanie informacji użytkownikom przestrzeni powietrznej o warunkach w niej panujących.
- Służba Informacji Lotniczej (Air Information Service - AIS) – jej zadaniem jest rozpowszechnianie informacji ważnych dla zachowania efektywności
  i bezpieczeństwa żeglugi powietrznej.

Sprawne i bezpieczne zarządzanie przestrzenią powietrzną ma kluczowe znaczenie dla zwiększenia pojemności systemu służb ruchu lotniczego, dla optymalnego zaspokojenia różnorodnych potrzeb użytkowników oraz dla osiągnięcia najbardziej elastycznego wykorzystania przestrzeni powietrznej.

W/w służby wspierane są przez grono specjalistów m.in. w zakresie bieżącej obsługi, konserwacji etc., urządzeń zapewniających nawigację, dozorowanie

oraz łączność. Dodatkowo do obowiązków ANSP należy zapewnienie funkcjonowania służb odpowiedzialnych za ochronę oraz bezpieczeństwo ruchu lotniczego.

Normy prawne narzucają na ANSP działające w ramach Wspólnoty Europejskiej wymóg certyfikacji działalności zgodnie z Rozporządzeniem (WE) Nr 2096/2005 Parlamentu Europejskiego i Rady z dnia 20 grudnia 2005 r. W Załączniku nr 1 do przedmiotowego rozporządzenia znajdują się ogólne wymogi dotyczące zapewnienia służb żeglugi powietrznej. Głównymi obszarami certyfikacji są [16]:

- Potencjał i kompetencje techniczne i operacyjne,
- Struktura organizacyjna i zarządzanie,
- Zarządzanie bezpieczeństwem i jakością,
- Ochrona,
- Zasoby ludzkie,
- Siła finansowa,
- Odpowiedzialność i zakres ubezpieczenia,
- Jakość służb oraz
- Wymogi dotyczące sprawozdawczości.

Włączenie funkcji bezpieczeństwa oraz ochrony w proces certyfikacyjny/re-certyfikacyjny wskazuje, iż te obszary należą do kluczowych w funkcjonowaniu ANSP. W związku z czym trzeba podkreślić, iż szczegółowa realizacja wymogów z tym związanych stanowi istotną kwestię dla władzy lotniczej.

3. Funkcjonowanie służb żeglugi powietrznej

Transport lotniczy jest narażony na wielorakie zagrożenia, które w przypadku ich materializacji mogą skutkować katastrofą lotniczą. Katastrofa lotnicza w swojej charakterystyce jest przyczyną śmierci większości jej uczestników i zniszczenia dobra wartego wiele milionów dolarów. Skutki finansowe i reputacyjne dla podmiotów związanych z takim zdarzeniem, w tym dla całej branży są bardzo dotkliwe. Dla Organizacji Międzynarodowego Lotnictwa Cywilnego (International Civil Aviation Organization - ICAO) oraz wszystkich podmiotów związanych z ruchem lotniczym jednym z najważniejszych aspektów jest zapewnienie odpowiedniego poziomu bezpieczeństwa, gdyż w ostateczności może się to przełożyć na końcowy wynik finansowy.

W transporcie lotniczym występuje bardzo wiele rodzajów zagrożeń, aczkolwiek można wyodrębnić dwie główne kategorie [11-13, 22-24]:

- zagrożenia spowodowane świadomym działaniem człowieka zmierzające do zakłócenia, uszkodzenia, zniszczenia urządzeń, informacji oraz statku powietrznego,

- nieświadome działanie człowieka, błąd, usterki oprogramowania lub urządzenia oraz działaniami przyrody (np. huragany, burze, powodzie).

W celu zapewnienia odpowiedniego rozróżnienia sposobu zapobiegania dla pierwszej kategorii w lotnictwie cywilnym przyjęto nazwę „security" – „ochrona", natomiast dla drugiej „safety" – „bezpieczeństwo".

Trzeba zaznaczyć, iż swój sukces branża lotnicza jako najbezpieczniejsza na świecie (jako środek transportu) zawdzięcza bardzo rygorystycznym rozwiązaniom prawnym. Zarówno w sferze zapewnienia ochrony przed aktami bezprawnej ingerencji, jak również przed błędem człowieka, usterkami urządzeń oraz błędami oprogramowania. Niestety bardzo często nowe rozwiązania podyktowane są katastrofami lotniczymi lub aktami terroru, które zostały spowodowane lukami lub brakami w dotychczasowo stosowanych rozwiązaniach.

4. Ochrona infrastruktury krytycznej ANSP

Ochrona lotnictwa cywilnego, w tym infrastruktury krytycznej służb żeglugi powietrznej wiąże się z zagrożeniem terrorystycznym oraz definicją aktu bezprawnej ingerencji. [1-13]

Aktem bezprawnej ingerencji w lotnictwie cywilnym nazywamy:

- użycie przemocy przeciwko osobie znajdującej się na pokładzie statku powietrznego będącego w trakcie lotu, jeżeli akt ten może zagrozić bezpieczeństwu tego statku,
- zniszczenie statku powietrznego znajdującego się w trakcie lotu lub spowodowanie jego uszkodzeń, które uniemożliwiają lot lub mogą stanowić zagrożenie dla bezpieczeństwa tego statku w trakcie lotu,
- umieszczenie na pokładzie statku powietrznego urządzenia lub substancji, które mogą zniszczyć statek powietrzny lub spowodować jego uszkodzenia, mogące uniemożliwić jego lot lub stanowić zagrożenie dla bezpieczeństwa tego statku powietrznego w trakcie lotu,
- porwanie statku powietrznego z załogą i pasażerami na pokładzie lub bez nich lub innego aparatu latającego w celu użycia ich jako środków ataku terrorystycznego z powietrza,
- zniszczenie lub uszkodzenie urządzeń naziemnych lub pokładowych, zakłócenie ich działania, w przypadku gdy stanowi to zagrożenie dla bezpieczeństwa statku powietrznego.

Poniżej przedstawiono wybrane akty terrorystyczne przeciwko lotnictwu cywilnemu:

- 1972 Tel Aviv;
- 1976 wybuch bomby Middle East Airlines;

- 1978 Air Rhodesia zestrzelony rakietami ziemia- powietrze;
- 1985 atak terrorystyczny na a/c Air India;
- 1988 Lockerbie;
- 1990 próba zawładnięcia podczas lądowania Xiamen;
- 1996 B767 Etiopian Airways;
- 11 września 2001 WTC.

Pierwszymi rozwiązaniami prawnymi regulującymi aspekt zapewnienia ochrony w ruchu lotniczym była Konwencja o międzynarodowym lotnictwie cywilnym sporządzona w Chicago w dniu 7 grudnia 1944 roku, podczas której uregulowano sprawy związane z całokształtem międzynarodowego lotnictwa cywilnego.

Pierwszym aktem prawnym opisującym akt bezprawnej ingerencji oraz wskazanie, iż jest to przestępstwo była Konwencja o zwalczaniu bezprawnych czynów skierowanych przeciwko bezpieczeństwu lotnictwa cywilnego, sporządzona w Montrealu w dniu 23 września 1971 roku.

Po wydarzeniach w Lockerbie w 1988 roku, została wydana Konwencja w sprawie znakowania plastycznych materiałów wybuchowych w celu ich wykrycia, sporządzona w Montrealu w dniu 1 marca 1991 roku. Zadaniem tego dokumentu było stworzenie systemu znakowania plastycznych materiałów wybuchowych w celu uzyskania możliwości ich rozpoznawania, wykrycia oraz śledzenia dróg transportu i kolportowania.

Obecnie poza w/w dokumentami w zakresie ochrony lotnictwa cywilnego, z uwzględnieniem ANSP funkcjonują również następujące rozwiązania prawne:

- Aneks 17 z 22 marca 1974 r. do Konwencji o międzynarodowym lotnictwie cywilnym („Ochrona Międzynarodowego Lotnictwa Cywilnego przed aktami bezprawnej ingerencji") sporządzony 7 grudnia 1944 r. w Chicago.

W/w rozwiązania narzucają na ANSP stosowanie odpowiednich środków ochrony dla infrastruktury krytycznej (urządzenia nawigacyjne, telekomunikacyjne, dozorowe, wieże kontroli lotów, meteorologiczne oraz urządzenia prądotwórcze), w tym m.in. [1-13]:

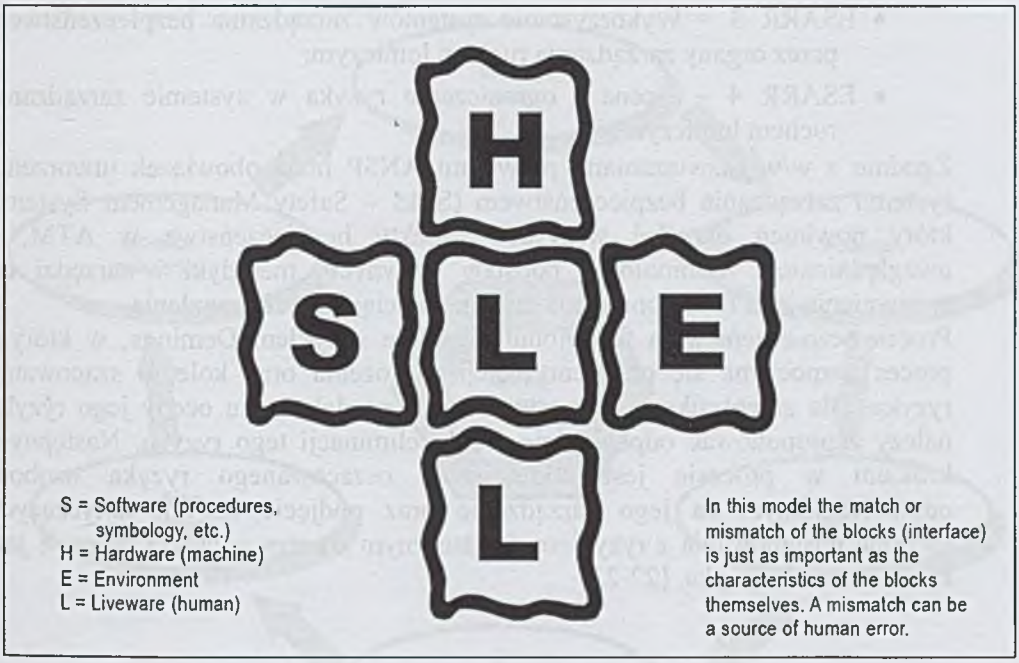
- utworzenie Programu ochrony zawierającego charakterystykę obiektów i urządzeń podlegających ochronie, utworzenie procedur i instrukcji dotyczących ochrony i kontroli dostępu, procedur na wypadek działań kryzysowych oraz działań zwiększających efektywność programu;
- wdrożenie programu szkolenia z zakresu ochrony lotnictwa cywilnego zawierającego szkolenia ze świadomości ochrony lotnictwa cywilnego dla wszystkich pracowników ANSP, szkolenia modułowe dla osób związanych z ochroną ANSP oraz szkolenia specjalistyczne, w tym również powtarzalność szkoleń w cyklu minimum 2u letnim,

- wdrożenie programu kontroli jakości w zakresie ochrony lotnictwa cywilnego zawierającego metodykę oraz zakres prowadzonych audytów kontroli jakości z zakresu ochrony lotnictwa cywilnego. W ramach prowadzonych audytów wyróżnia się: audyt ochrony, oceny ochrony, badanie ochrony, inspekcję ochrony, test ochrony, przegląd ochrony oraz ćwiczenie ochrony. Dla każdego z w/w rodzajów audytu został określony tryb oraz zakres prowadzenia badania.

Infrastruktura krytyczna ANSP chroniona jest w sposób wieloetapowy. Pierwszy etap stanowią środki ochrony fizycznej, w tym bariery, takie jak ogrodzenie o odpowiedniej strukturze i wysokości (najczęściej zwieńczone CONCERTINĄ), bramy oraz furty. Wszystkie strefy oraz obszary znajdujące się w bezpośrednim sąsiedztwie ogrodzenia są poddawane nadzorowi patrolowemu, systemowi telewizji dozorowej oraz innym metodom monitoringu i sygnalizacji włamania i napadu. Dodatkowo wprowadzone restrykcyjne i szczegółowe procedury kontroli dostępu do poszczególnych stref, w odniesieniu do osób bez widocznych kart identyfikacyjnych oraz w stosunku do osób udających się do stref, do których nie są upoważnione. Istnieją również szczegółowe procedury postępowania na wypadek wystąpienia sytuacji awaryjnych oraz sytuacji podwyższonego ryzyka. Wszyscy pracownicy ANSP przechodzą specjalistyczne szkolenia odnawiane w cyklach minimum 2u letnich informujące o zagrożeniach, sposobach ich minimalizacji oraz o zachowaniu w sytuacjach kryzysowych. Skuteczność zastosowanych metod i środków weryfikowana jest za pomocą audytów kontroli jakości w zakresie ochrony lotnictwa cywilnego (wewnętrznych i zewnętrznych). [1-13]

5. Bezpieczeństwo systemów teleinformatycznych ANSP

Bezpieczeństwo w realizacji zadań służb żeglugi powietrznej stanowi jeden z najistotniejszych filarów ich funkcjonowania. W skład kompleksowego systemu bezpieczeństwa wchodzi zapewnienie odpowiednich procesów dla kontrolowania i minimalizacji ryzyka wszystkich elementów ATM (Air Traffic Management - Zbiór naziemnych i pokładowych funkcji wymaganych do zapewnienia bezpiecznego i wydajnego przemieszczania się statków powietrznych podczas wszystkich faz lotu), w których skład wchodzą: ludzie, urządzenia, oprogramowanie oraz środowisko funkcjonowania. [17-25]
Połączenie w/w czynników i ich wspólne oddziaływanie tworzą tzw. model SHEL/SHELL (Liveware (L), Hardware (H), Software (S), Environment (E)). Model ten wskazuje wpływ poszczególnych czynników i ich połączenie na możliwość wystąpienia ludzkiego błędu, który może powodować niebezpieczeństwo w ruchu lotniczym. [23]



Rys. 1. Model SHELL [23]

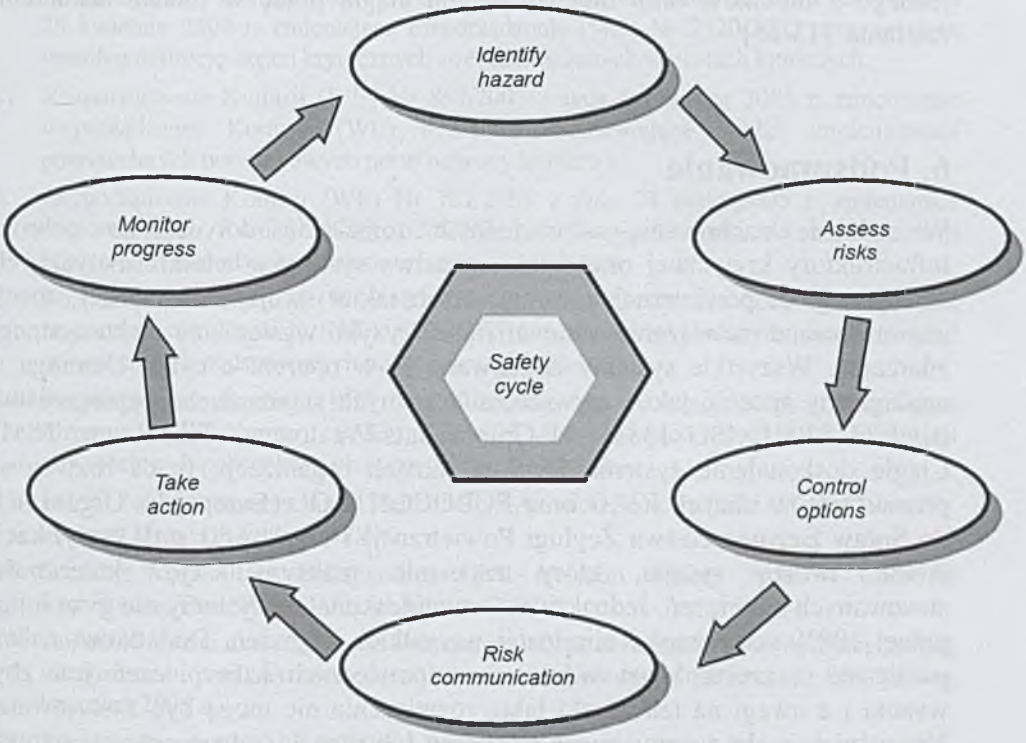
Prawne aspekty bezpieczeństwa systemów teleinformatycznych w ATM mają swoje źródło m.in. w poniższych aktach prawnych:

- Rozporządzenie (WE) NR 550/2004 Parlamentu Europejskiego i Rady Unii Europejskiej z dnia 10 marca 2004 r. w sprawie zapewnienia służb nawigacji lotniczej w Jednolitej Europejskiej Przestrzeni Powietrznej;
- Rozporządzenie Komisji (WE) Nr 2096/2005 z dnia 20 grudnia 2005 r. ustanawiające wspólne wymogi dotyczące zapewnienia służb żeglugi powietrznej;
- Rozporządzenie Komisji (WE) Nr 1315/2007 z dnia 8 listopada 2007 r. w sprawie nadzoru nad bezpieczeństwem w zarządzaniu ruchem lotniczym oraz zmieniające Rozporządzenie (WE) Nr 2096/2005;
- DOC. 4444 ICAO;
- EUROCONTROL Air Traffic Management (ATM) Strategy for the years 2000+;
- IEC 61508 Bezpieczeństwo funkcjonalne elektrycznych/elektronicznych/ programowalnych elektronicznych systemów wiążących się z bezpieczeństwem;
- ESARR 2 – Składanie meldunków oraz rozpatrywanie nieprawidłowości w ruchu lotniczym;

- ESARR 3 – Wykorzystanie systemów zarządzania bezpieczeństwem przez organy zarządzania ruchem lotniczym;
- ESARR 4 – Ocena i ograniczenie ryzyka w systemie zarządzania ruchem lotniczym.

Zgodnie z w/w rozwiązaniami prawnymi ANSP mają obowiązek utworzenia systemu zarządzania bezpieczeństwem (SMS – Safety Management System), który powinien określać wszystkie aspekty bezpieczeństwa w ATM, z uwzględnieniem terminologii, podstaw prawnych, metodyki i narzędzi do zapewnienia jego funkcjonowania oraz metod ciągłego doskonalenia.

Proces bezpieczeństwem funkcjonuje zgodnie z cyklem Deminga, w którym proces rozpoczyna się od identyfikacji zagrożenia oraz kolejno szacowanie ryzyka. Dla zidentyfikowanego zagrożenia i po dokonaniu oceny jego ryzyka należy zaproponować odpowiednie środki eliminacji tego ryzyka. Następnym krokiem w procesie jest komunikacja oszacowanego ryzyka osobom odpowiedzialnym za jego zarządzanie oraz podjęcie decyzji dotyczących sposobu postępowania z ryzykiem. Ostatecznym etapem w całym procesie jest monitorowanie ryzyka. [22-23]



Rys. 2. Cykl bezpieczeństwa [23]

W ramach SMS analogicznie systemu ochrony zalicza się: szkolenia personelu związanego z bezpieczeństwem, których zadaniem jest kształtowanie odpowiedniej kultury bezpieczeństwa w organizacji oraz prowadzenie wewnętrznego audytu bezpieczeństwa, który weryfikuje skuteczność oraz prawidłowość zastosowanych rozwiązań w ramach całego systemu.

Zgodnie z wymaganiami prawnymi wszystkie nowe systemy ATM/CNS (w tym wszystkie urządzenia) oraz wszystkie wprowadzane w nich zmiany muszą zostać poddane analizie pod względem bezpieczeństwa, w trakcie której zarówno producent, jak i ANSP zobowiązani są do szczegółowej analizy zagrożeń oraz ryzyka. W celu dopuszczenia omawianego wyżej systemu do użytkowania ANSP ma obowiązek złożenia pełnej dokumentacji wykonawczej oraz dokumentacji przeprowadzonych analiz, do ponownej weryfikacji przez władzę lotniczą, która wydaje decyzję o dopuszczeniu i certyfikacji takiego systemu (urządzenia).

Należy zaznaczyć, iż stosowane rozwiązania w swoim zakresie poza wysokimi współczynnikami niezawodności charakteryzują się najczęściej pełną redundancją, brakiem wspólnych punktów awarii oraz rozwiązaniami zapewniającymi zachowanie głównych funkcjonalności w przypadku degradacji

jednego z modułów oraz zapewniającymi ciągłą pracę w trakcie zaburzenia zasilania. [13-25]

6. Podsumowanie

W zakresie zachowania odpowiednich rozwiązań dotyczących ochrony infrastruktury krytycznej oraz bezpieczeństwa systemów teleinformatycznych, służby żeglugi powietrznej z uwagi na charakter swojej działalności stosują zaawansowane rozwiązania minimalizujące ryzyko wystąpienia niekorzystnego zdarzenia. Wszystkie systemy zbudowane są w oparciu o cyklu Deminga w analogiczny sposób, jak w powszechnie znanych standardach bezpieczeństwa (tj. ISO 27001, ISO 13335, IT-Grundschutz-Catalogues, TISM oraz TSM). Ciągłe doskonalenie systemu SMS w ramach organizacji, prace rozwojowe prowadzone w ramach ICAO oraz EUROCONTROL (Europejska Organizacja do Spraw Bezpieczeństwa Żeglugi Powietrznej) i ciągła oraz stała certyfikacja ANSP tworzą system, który zapewnia maksymalizację skuteczności stosowanych rozwiązań. Jednakże nawet najdoskonalsze systemy nie gwarantują pełnej 100% skuteczności eliminacji wszystkich zagrożeń. Dodatkowo należy pamiętać, iż często koszt wdrożenia odpowiednich zabezpieczeń jest zbyt wysoki i z uwagi na ten aspekt takie rozwiązania nie mogą być zastosowane. Najważniejszy, krytyczny czynnik sukcesu lub porażki całego procesu stanowi czynnik ludzki. Ponieważ na początku, w trakcie oraz na końcu każdego procesu znajduje się człowiek, jego ingerencja lub jej brak może ostatecznie przyczynić się do katastrofy, w wyniku której może stracić życie kilkaset istnień ludzkich. [26-30]

LITERATURA

1. Konwencja o zwalczaniu bezprawnych czynów skierowanych przeciwko bezpieczeństwu lotnictwa cywilnego - Montreal 1971 r. (Dz. U z 1976r., Nr 8, poz. 37 i 38 z późn. zm.).
2. Rozporządzenie (WE) Nr 2320/2002 Parlamentu Europejskiego i Rady z dnia 16 grudnia 2002 roku ustanawiające wspólne zasady bezpieczeństwa w lotnictwie cywilnym.
3. Rozporządzenie Komisji (WE) Nr 622/2003 z dnia 4 kwietnia 2003 r., ustanawiające środki w celu wprowadzenia w życie wspólnych podstawowych standardów dotyczących bezpieczeństwa lotnictwa cywilnego.
4. Rozporządzenie Komisji (WE) Nr 1217/2003 z dnia 4 lipca 2003 r. ustanawiające powszechne specyfikacje dla krajowych programów kontroli jakości w zakresie ochrony lotnictwa cywilnego.
5. Rozporządzenie Komisji (WE) Nr 68/2004 z dnia 14 stycznia 2004 r. zmieniające rozporządzenie 622/2003 ustanawiające środki do implementacji powszechnych podstawowych norm ochrony lotnictwa.

6. Rozporządzenie Komisji (WE) Nr 849/2004 Parlamentu Europejskiego i Rady z dnia 29 kwietnia 2004 r. zmieniające rozporządzenie (WE) Nr 2320/2002 ustanawiające wspólną definicję części krytycznych stref zastrzeżonych w portach lotniczych.
7. Rozporządzenie Komisji (WE) Nr 857/2005 z dnia 6 czerwca 2005 r. zmieniające rozporządzenie Komisji (WE) 622/2003 ustanawiające środki implementacji powszechnych podstawowych norm ochrony lotnictwa.
8. Rozporządzenie Komisji (WE) Nr 781/2005 z dnia 24 maja 2005 r. zmieniające rozporządzenie Komisji (WE) 622/2003 ustanawiające środki do implementacji powszechnych podstawowych norm ochrony lotnictwa.
9. Rozporządzenie Komisji (WE) Nr 65/2006 z dnia 13 stycznia 2006 r. zmieniające rozporządzenie Komisji (WE) 622/2003 ustanawiające środki do implementacji powszechnych podstawowych norm ochrony lotnictwa.
10. Rozporządzenie Komisji (WE) Nr 240/2006 z dnia 10 lutego 2006 r. zmieniające rozporządzenie Komisji (WE) 622/2003 ustanawiające środki implementacji powszechnych podstawowych norm ochrony lotnictwa.
11. ECAC Policy Statement in the Field of Civil Aviation Security (ECAC.CEAC Doc No. 30. (part II) – 11th Edition 2003 r.
12. Security Manual for Safeguarding Civil Aviation Against Acts of Unlawful Interference, Doc 8973 ICAO, 6th Edition 2002 r.
13. Annex 17 to the Convention on International Civil Aviation – Security Safeguarding International Civil Aviation Against Acts of Unlawful Interference, 8th Edition, kwiecień 2006 r.,
14. Rozporządzenie (WE) Nr 549/2004 Parlamentu Europejskiego i Rady z dnia 10 marca 2004 r. ustanawiające ramy tworzenia Jednolitej Europejskiej Przestrzeni Powietrznej (Dz. U. L 96/1 z 31.3.2004).
15. Rozporządzenie (WE) Nr 550/2004 Parlamentu Europejskiego i Rady z dnia 10 marca 2004 r. w sprawie zapewnienia służb nawigacji lotniczej w Jednolitej Europejskiej Przestrzeni Powietrznej (Dz. U. L 96/10 z 31.3.2004).
16. Rozporządzenie Komisji (WE) Nr 2096/2005 z dnia 20 grudnia 2005 r. ustanawiające wspólne wymogi dotyczące zapewnienia służb żeglugi powietrznej (Dz. U. L 335/13 z 21.12.2005).
17. Rozporządzenie Komisji (WE) Nr 1315/2007 z dnia 8 listopada 2007 r. w sprawie nadzoru nad bezpieczeństwem w zarządzaniu ruchem lotniczym oraz zmieniające Rozporządzenie (WE) Nr 2096/2005.
18. EUROCONTROL Air Traffic Management (ATM) Strategy for the years 2000+, 2003.
19. EUROCONTROL: ESARR 2 – Składanie meldunków oraz rozpatrywanie nieprawidłowości w ruchu lotniczym, EUROCONTROL, edycja 2.0, 03.11.2000 r.
20. EUROCONTROL: ESARR 3 – Wykorzystanie systemów zarządzania bezpieczeństwem przez organy zarządzania ruchem lotniczym, EUROCONTROL, edycja 1.0, 17.07.2000 r.
21. EUROCONTROL: ESARR 4 – Ocena i ograniczenie ryzyka w systemie zarządzania ruchem lotniczym, EUROCONTROL, edycja 1.0, 05.041.2001 r.
22. Safety Oversight Manual (Doc 9734) Second Edition 2006 .

23. Safety Management Manual (Doc 9859) First Edition 2006.
24. DOC. 4444 ICAO 15th Edition 2007.
25. IEC 61508 Bezpieczeństwo funkcjonalne elektrycznych/elektronicznych/ programowalnych elektronicznych systemów wiążących się z bezpieczeństwem.
26. ISO/IEC 27001:2005 Information technology – Security techniques – Information security management systems – Requirements.
27. ISO/IEC TR 13335-1:1996 Information technology – Guidelines for the management of IT Security – Part 1: Concepts and models for IT Security.
28. IT-Grundschutz Catalogues: Version 2005.
29. Total Information Security Management ver. 1.4.1.
30. Total Security Management ver. 1.0.

Rozdział 3

Wdrażanie systemu zarządzania bezpieczeństwem informacji w urzędach administracji publicznej

Beata Hysa
Politechnika Śląska, Wydział Organizacji i Zarządzania, Katedra Informatyki i Ekonometrii
bhysa@woiz.polsl.pl

Streszczenie

W rozdziale przedstawiono zagadnienie zarządzania bezpieczeństwem informacji w administracji publicznej. Na początku omówiono celowość wdrażania systemu bezpieczeństwa informacji w administracji publicznej oraz podano wynikające z tego korzyści. Następnie krótko opisano istniejącą sytuację wdrażania systemu zarządzania bezpieczeństwem informacji w urzędach administracji publicznej. Na koniec dokonano podsumowania.

1. Wzrost znaczenia zarządzania bezpieczeństwem informacji

Dlaczego bezpieczeństwo informacji staje się coraz bardziej popularne? Wymienić możemy trzy powody, dla których nastąpił istotny wzrost zainteresowania zarządzaniem bezpieczeństwem informacji:[1]

- wzrost znaczenia informacji w gospodarce,
- pogłębianie współpracy pomiędzy przedsiębiorstwami oraz instytucjami,
- rosnący poziom trudności zarządzania informacjami.

Wartość i znaczenie informacji we współczesnym świecie ciągle wzrasta a wraz z nim trzy podstawowe atrybuty: poufność, dostępność, integralność. Coraz większa liczba instytucji zauważa potrzebę zachowania poufności informacji w organizacji. Związane jest to z wejściem ustawy o ochronie danych osobowych,

oraz z pojawiającymi się przypadkami sprzedawania i wykorzystywania danych osobowych, zarówno przez legalnie działające firmy, jak i organizacje przestępcze. Współcześnie przedsiębiorcy są w stanie uszeregować informacje wykorzystywane w organizacji według poziomów poufności. Dostępność informacji nie jest zwykle postrzegana jako problem organizacyjny, natomiast brak dostępu do danych jest łatwo tłumaczony urlopem, brakiem prądu, komputerowym wirusem, zgubionym kluczem. Jednak włamania na strony rządowe w kwietniu 2008 roku pokazały, że brak dostępu do informacji może wpływać na wizerunek zarówno całej organizacji, jak i osoby nią zarządzającej.

Pogłębianie współpracy pomiędzy różnymi instytucjami, organizacje sieciowe, outsourcing, budowanie długotrwałych relacji zgodnie z zasadami zarządzania jakością i TQM, wprowadzanie międzyorganizacyjnych systemów informatycznych, powoduje konieczność badania nie tylko własnych zabezpieczeń, ale także bezpieczeństwa udostępnianych partnerom informacji. Z tego powodu system zabezpieczeń certyfikowanych przez niezależną, akredytowaną organizację, może być dobrym rozwiązaniem, pozwalającym potwierdzić właściwy poziom bezpieczeństwa informacji u klienta.

Trzecia istotna grupa przyczyn popularności Systemu Zarządzania Bezpieczeństwem Informacji (SZBI) to wzrost trudności zarządzania informacjami. Najczęściej w organizacji występują równolegle co najmniej dwa obiegi dokumentów – papierowy i elektroniczny. Dodatkowo szereg informacji przekazywanych jest ustnie bezpośrednio lub telefonicznie. Istniejące w przedsiębiorstwach służby informatyczne zajmują się wyłącznie funkcjonowaniem systemów informatycznych i ich bezpieczeństwem. Nadzorują działanie serwerów, funkcjonowanie sieci, instalują programy antywirusowe na stacjach roboczych. Jednak nie są zainteresowane pozatechnicznymi problemami obiegu informacji. Jeżeli przedsiębiorstwo wdrożyło system zarządzania jakością, to istnieją również służby odpowiedzialne za obieg dokumentów, ich aktualność i dostępność. Jednak zwykle pomijają one np. problemy poufności. Brakuje zatem spójnego podejścia do informacji i ich bezpieczeństwa w organizacji. Kompleksowa koncepcja zarządzania informacjami prezentowana przez SZBI stanowić może rozwiązanie tych problemów.

Zarządzanie bezpieczeństwem jest dziedziną z pogranicza informatyki, prawa, organizacji i zarządzania, która zajmuje się definiowaniem aspektów bezpieczeństwa dla organizacji i jej systemów informatycznych, jego osiąganiem i utrzymywaniem. Bezpieczeństwo informacji jest rozpatrywane w aspekcie technicznym, organizacyjnym, prawnym, ekonomicznym i społecznym.

Każda organizacja powinna stosować własną politykę bezpieczeństwa i ochrony danych, która określa jakie informacje, i w jaki sposób należy chronić. Polityka bezpieczeństwa jest fundamentem, na którym należy stworzyć ogólną strategię bezpieczeństwa działalności. Na jej podstawie można podejmować decyzje o

tym, ile i gdzie wydać na bezpieczeństwo. Zbyt często zdarza się bowiem, że organizacje wydają pieniądze na ochronę aktywów, które nie wymagają ochrony, oszczędzając natomiast w obszarach potrzebujących zabezpieczeń. Polityka bezpieczeństwa stanowi również bazę, dzięki której można określić procesy i procedury ochrony danych. Polityka bezpieczeństwa musi się rozwijać wraz z rozwojem organizacji i uwzględniać nowe systemy i aplikacje informatyczne oraz ciągle pojawiające się nowe zagrożenia. Przykładowy schemat struktury systemu zarządzania bezpieczeństwem informacji podzielony na obszary zabezpieczeń przedstawia rys. 1.

Rys. 1. Information Security Management
Źródło: Saint-Germain R., Information Security Management Best Practice Based on ISO/IEC 17799, The Information Management Journal, July/August 2005

2. Po co wdrażać SZBI w urzędach administracji publicznej?

Wdrożony System Zarządzania Bezpieczeństwem Informacji według normy PN-ISO/IEC-27001:2007 potwierdza, że organizacja stosuje niezbędne środki ostrożności, aby zabezpieczyć istotne informacje przed nieautoryzowanym dostępem lub zmianami. Stosowany, certyfikowany System Zarządzania Bezpieczeństwem Informacji (SZBI) pokazuje klientom, że zasoby danych

organizacji są odpowiednio chronione. Poza tym należy podkreślić, iż problematyka bezpieczeństwa informacji daleko wykracza poza strefę technologii informatycznych. Wraz z nowymi rozwiązaniami technicznymi, zapewniającymi łatwy i szybki dostęp do informacji, „przecieki" informacyjne stają się coraz częstszym zjawiskiem. Ryzyko utraty ważnych informacji zwiększa także zjawisko migracji pracowników między konkurującymi firmami. Wymaga to stworzenia odpowiedniego systemu ochrony informacji. Systemowe podejście do bezpieczeństwa informacji pozwala na racjonalne zarządzanie przepływem informacji z zachowaniem odpowiedniego poziomu bezpieczeństwa. Wdrożenie Systemu Zarządzania Bezpieczeństwem Informacji w organizacji publicznej i rezygnacja z działań doraźnych sprawia, że procesy wewnętrzne mogą stać się łatwiejsze do zarządzania, mierzenia i doskonalenia [3].

Ważnymi korzyściami z wdrożenia Systemu Zarządzania Bezpieczeństwem Informacji w urzędach administracji publicznej są:

- spełnienie wymagań ustawowych – ustawa o ochronie danych osobowych, ustawa o ochronie informacji niejawnych, ustawa o dostępie do informacji publicznej, ustawa o prawie autorskim i prawach pokrewnych,
- ochrona informacji znajdujących się w obiegu w ramach instytucji,
- zabezpieczenie informacji na wypadek katastrof lub awarii – zarządzanie ciągłością działania urzędu,
- uporządkowanie i klasyfikacja informacji przetwarzanych przez urząd,
- wzrost świadomości pracowników w zakresie bezpieczeństwa informacji,
- zapewnienie interesantów i zainteresowane instytucje, że ich dane są właściwie chronione,
- oszacowanie ryzyka związanego z utratą informacji,
- kompleksowe zarządzanie systemami informatycznymi i sieciami komputerowymi pod kątem bezpieczeństwa informacji.

Głównymi celami stawianymi przed Systemem Zarządzania Bezpieczeństwem Informacji (SZBI) w urzędzie administracji publicznej powinno być:

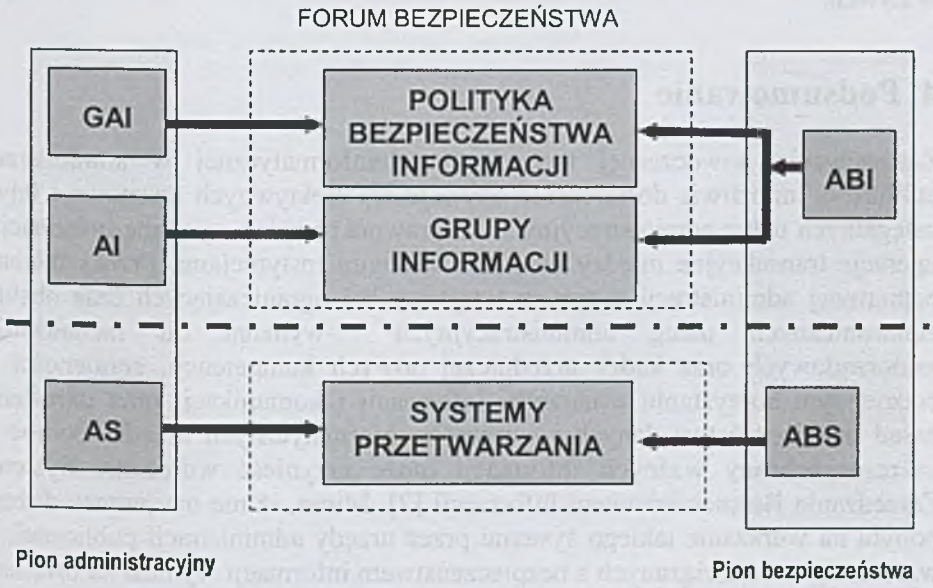
- zapewnienie zgodności z prawem obowiązującym na terytorium RP,
- ochrona systemów przetwarzania informacji przed nieuprawnionym dostępem bądź zniszczeniem,
- podnoszenie świadomości pracowników,
- zmniejszenie ryzyka utraty informacji,
- zaangażowanie wszystkich pracowników w ochronę informacji.

3. Urzędowe bezpieczeństwo – stan obecny

Różnego rodzaju badania i raporty pokazują, że bezpieczeństwo informacji w urzędach administracji publicznej jest na bardzo niskim poziomie. Na przykład ostatnie badanie na temat bezpieczeństwa informacji w administracji publicznej, przeprowadzone przez redakcję Computerworld pokazuje, iż wiele jeszcze musi się zmienić w świadomości urzędników odnośnie do zarządzania bezpieczeństwem informacji. Bowiem, oprócz środków technicznych dla zapewnienia bezpieczeństwa informacji ważne są również inne czynniki, takie jak organizacja, podział kompetencji, szkolenia pracowników. Okazało się na przykład, że wyznaczenie osobnej jednostki odpowiedzialnej za bezpieczeństwo teleinformatyczne, ma wpływ na poziom stosowanych zabezpieczeń technicznych. Najgorzej wypadły szkolenia pracowników gdyż w prawie jednej trzeciej ze zbadanych instytucji, nie przeprowadzono szkoleń z zakresu bezpieczeństwa informacji [4].

Nie mniej jednak należy wspomnieć o urzędach, które zauważają potrzebę właściwego zarządzania bezpieczeństwem informacji i opracowują odpowiednią strategię zarządzania bezpieczeństwem informacji lub też wdrażają SZBI zgodnie z normą ISO/IEC-27001.

Przykład odpowiedzialności za zarządzanie bezpieczeństwem informacji w Urzędzie Marszałkowskim Województwa Małopolskiego w Krakowie przedstawia rys.2.



Rys. 2. Schemat odpowiedzialności w SZBI Urzędu Marszałkowskiego Województwa Małopolskiego w Krakowie
Źródło: Polityka bezpieczeństwa systemu informacyjnego Urzędu Marszałkowskiego Województwa Małopolskiego w Krakowie

Nad przestrzeganiem postanowień Polityki Bezpieczeństwa Informacji i rozwojem Systemu Zarządzania Bezpieczeństwem w Urzędzie Marszałkowskim czuwa Forum Zarządzania Bezpieczeństwem. Forum Zarządzania bezpieczeństwem składa się z dwóch pionów: pionu administracyjnego (zarządzającego informacją), oraz pionu bezpieczeństwa (zarządzającego bezpieczeństwem informacji).

W ramach obu pionów wyróżnione zostały role [5]:

Na poziomie Polityki Bezpieczeństwa Informacji:

- Głównego Administratora Informacji (GAI),
- Administratora Bezpieczeństwa Informacji (ABI).

Na poziomie Grupy Informacji:

- Administratora Informacji – właściciela informacji (AI),
- Administratora Bezpieczeństwa Informacji (ABI).

Na poziomie Systemu Przetwarzania:

- Administratora Systemu (AS),
- Administratora Bezpieczeństwa Systemu (ABS).

Głównym Administratorem Informacji (GAI) jest Marszałek Województwa Małopolskiego.

W województwie śląskim certyfikat wdrożenia SZBI otrzymały: Urząd Miasta Bielsko-Biała, Starostwo powiatowe w Bielsku-Białej, Starostwo powiatowe w Żywcu.

4. Podsumowanie

Zastosowanie nowoczesnej technologii teleinformatycznej w administracji publicznej umożliwia dostarczanie obywatelom efektywnych kosztowo i łatwo osiągalnych usług administracyjnych. Usprawnia również wymianę informacji i operacje transakcyjne między urzędami i innymi instytucjami. Przekształcenie papierowej administracji w system wygodnych i ograniczających czas obsługi elektronicznych usług administracyjnych wymaga od menadżerów samorządowych oraz kadry urzędniczej nowych kompetencji, sprawności w codziennym korzystaniu z narzędzi informacji i komunikacji oraz określenia zasad bezpieczeństwa danych i systemów informatycznych urzędu. Pomoc w zakresie ochrony ważnych informacji może przynieść wdrożenie Systemu Zarządzania Bezpieczeństwem Informacji [2]. Mimo, iż nie ma jeszcze dużego popytu na wdrażanie takiego systemu przez urzędy administracji publicznej, to wzrost zagrożeń związanych z bezpieczeństwem informacji wymusi na urzędach większe zainteresowanie tym zagadnieniem.

LITERATURA

1. Wawak S.: Dlaczego bezpieczeństwo informacji jest tak popularne? Źródło:www.wawak.pl
2. Hysa B.: Wpływ bezpieczeństwa informacji na działalność urzędów administracji publicznej, Zeszyty Naukowe Politechniki Śląskiej – Organizacja i Zarządzanie z. 45, Gliwice 2008.
3. Wolniak R.: Tworzenie polityki bezpieczeństwa informacji w organizacjach publicznych, Zeszyty Naukowe Politechniki Śląskiej – Organizacja i Zarządzanie, Gliwice .
4. „Bezpieczeństwo informacji w administracji publicznej w Polsce", Badanie Computerworld w ramach przygotowań do V konferencji „Wolność i bezpieczeństwo", Wieliczka 2008. Źródło: WWW.publicstandard.pl
5. Polityka bezpieczeństwa systemu informacyjnego Urzędu Marszałkowskiego Województwa Małopolskiego w Krakowie, www.wrotamalopolski.pl
6. Hysa B.: Jakość informacji w zarządzaniu administracją publiczną [w] Technologie wiedzy w zarządzaniu publicznym '08 - Konwersja wiedzy. Praca zbiorowa pod redakcją naukową Gałuchowskiego J., Frączkiewicz-Wronki A., AE Katowice 2008.

Rozdział 4

Atakowanie systemów i sieci komputerowych – szkodliwe oprogramowanie z legalnych źródeł

Teresa Mendyk-Krajewska, Zygmunt Mazur
Politechnika Wrocławska, Instytut Informatyki
teresa.mendyk-krajewska@pwr.wroc.pl, zygmunt.mazur@pwr.wroc.pl

Streszczenie

Obserwowana od lat destrukcyjna i przestępcza działalność w sieciach komputerowych przybiera na sile wraz ze wzrostem atrakcyjności zasobów i usług internetowych. Do skutecznego przeprowadzania ataków sieciowych przyczyniają się wady instalowanego oprogramowania, zła konfiguracja systemów i niewłaściwe ich zabezpieczanie. Wraz z ciągłym udoskonalaniem zabezpieczeń oprogramowania, rozwijane są narzędzia do ich łamania. O ile dawniej do wyszukiwania luk systemowych i łamania zabezpieczeń potrzebna była zaawansowana wiedza, to obecnie wyzwanie takie może podjąć przeciętny użytkownik Internetu. Przeprowadzenie ataku umożliwiają dostępne w sieci, a także dołączane do popularnych czasopism, specjalizowane, rozbudowane funkcjonalnie narzędzia, a pomocy dostarcza odpowiednia literatura o charakterze instruktażowym. Wobec znaczącego wzrostu zagrożenia bezpieczeństwa sieciowego coraz częściej słyszy się o wymierzaniu kar producentom i dystrybutorom oprogramowania o szkodliwym działaniu. Ograniczenie jego rozpowszechniania nie usunie problemu, ale pozwoli znacznie ograniczyć skalę zjawiska.

1. Wstęp

Zjawisko zagrożenia bezpieczeństwa sieciowego niezmiennie wykazuje tendencję wzrostową. Ataki na systemy informatyczne podejmowane są z powodu dużej atrakcyjności zasobów sieciowych, zaś niewłaściwa konfiguracja, złe zabezpieczanie systemów oraz dostępność narzędzi hakerskich czyni z nich

łatwy cel. Tworzeniem niebezpiecznego oprogramowania zajmują się już nie tylko szukający wyzwania amatorzy, lecz zorganizowane grupy, których działalność często nosi znamiona przestępstwa. Dostępność specjalizowanych, rozbudowanych funkcjonalnie narzędzi pozwala podejmować ataki nawet osobom niedoświadczonym, a pełny sukces gwarantuje literatura stanowiąca swoisty instruktaż. Rosnąca w ostatnich latach liczba oprogramowania hakerskiego i jego łatwa dostępność nie pozostaje bez wpływu na bezpieczeństwo systemów i sieci komputerowych.

Z raportu należącej do FBI agencji National Infrastructure Protection Center zajmującej się problematyką bezpieczeństwa sieciowego opublikowanego w 2008 roku wynika, że hakerzy najczęściej nie stosują wyrafinowanych metod, lecz wykorzystują znane, nie załatane luki (*vulnerability*)[3] w systemach komputerowych, używając do włamań powszechnie dostępne narzędzia – nie tylko hakerskie, ale też przeznaczone dla administratorów. Wiele aplikacji można bowiem użyć zarówno w dobrych, jak i złych celach. Na przykład wszelkie narzędzia umożliwiające analizę przesyłanych w sieci pakietów pozwalają na szpiegowanie (np. podsłuchiwanie rozmów prowadzonych przy pomocy komunikatorów internetowych) oraz pozyskiwanie różnego typu informacji, jeśli nie zastosowano tzw. bezpiecznego połączenia, w którym strumień przesyłanych danych jest szyfrowany.

W ostatnich latach przedział czasu pomiędzy odkryciem wady oprogramowania umożliwiającej nieautoryzowany dostęp do systemu a opracowaniem wykorzystującego ją szkodliwego kodu stale się zmniejsza. Choć instytucje odpowiedzialne za aktualizację systemów starają się możliwie szybko publikować niezbędne nakładki programowe, użytkownicy nie nadążają z ich instalacją. Ponadto, zaatakowany już system najczęściej nie daje się automatycznie zaktualizować, o czym można przekonać się analizując informacje zebrane w dziennikach zdarzeń.

2. Problem luk w systemach oprogramowania

Według słownika Common Vulnerabilities and Exposures, tworzonego przez amerykańską grupę naukowo-rozwojową MITRE, należy rozgraniczać dwa pojęcia: luka i podatność [5]. I tak, lukę definiuje się jako stan systemu, który umożliwia atakującemu:

- wykonanie poleceń w imieniu innego użytkownika,

[3] Wady w zabezpieczeniach oprogramowania umożliwiające włamanie, czyli nieautoryzowany dostęp do zdalnego systemu przy pomocy specjalnie przygotowanego do tego celu kodu umożliwiającego bezprawne działania, a nawet przejęcie kontroli nad systemem (luki o znaczeniu krytycznym).

- zdobycie dostępu do danych z naruszeniem określonych ograniczeń dostępu,
- podszywanie się pod inną jednostkę,
- przeprowadzenie ataku DoS.

Jeśli atak jest możliwy dzięki słabej lub nieodpowiedniej polityce bezpieczeństwa, to ten stan systemu komputerowego należy określić jako podatność (lub wystawianie się na atak). Jest to stan, w którym:

- możliwe jest gromadzenie informacji,
- możliwe jest ukrywanie wykonywanych czynności,
- zachowywane są prawidłowe funkcje systemu, jednak można go zaatakować,
- istnieje problem z punktu widzenia polityki bezpieczeństwa.

Osoby zajmujące się wyszukiwaniem słabości systemów informatycznych to najczęściej crackerzy[4], hakerzy lub członkowie zespołu ds. bezpieczeństwa komputerowego. Różni je to, że crackerzy o znalezionych błędach informują zazwyczaj tylko swoją społeczność i wykorzystują wykryte luki do przeprowadzenia ataków, zaś pozostałe grupy powiadamiają o fakcie tylko zainteresowane instytucje (zajmujące się problematyką bezpieczeństwa czy twórców oprogramowania). Hakerzy po znalezieniu luki opracowują (i bardzo często upubliczniają) kod umożliwiający jej wykorzystanie do przeprowadzenia ataku (co zajmuje im coraz mniej czasu). Istnieje wiele exploitów[5] (a także wirusów i robaków) wykorzystujących powszechnie znane luki. Niektóre firmy (np. Microsoft) nawiązują współpracę z hakerami w celu upubliczniania informacji o nowych lukach dopiero w momencie, gdy zostaną opracowane i mogą być rozpowszechnione odpowiednie zabezpieczenia.

W marcu 2009 roku firma Secunia opublikowała raport, w którym między innymi przedstawiła dane liczbowe dotyczące luk w systemach komputerowych w latach 2003-2008 [3]. W 2008 roku wśród wszystkich luk najwięcej, bo aż 1814 pozwalało na uzyskanie dostępu do systemu operacyjnego (w roku 2003 było ich 1020). Drugie pod względem liczebności były luki umożliwiające przeprowadzenie ataku DoS (*Denial of Service*)[6], zaś kolejne pozycje to luki pozwalające na zwiększenie uprawnień i manipulowanie danymi. Znacząco spadła zaś liczba luk, w wyniku których atakujący mógł uzyskać dostęp do ważnych informacji. Jeszcze w 2003 roku firma Secunia informowała o 482 takich wadach, a w 2008 tylko o trzynastu przypadkach. Wybrane wyniki

---

[4] Cracker – powszechnie osoba łamiąca zabezpieczenia systemu oprogramowania przed jego nielegalnym modyfikowaniem bądź kopiowaniem.

[5] Specjalnie przygotowany kodu umożliwiający złamanie zabezpieczeń i w konsekwencji przejęcie kontroli nad systemem.

[6] Celem ataku jest blokada dostępności danych i usług.

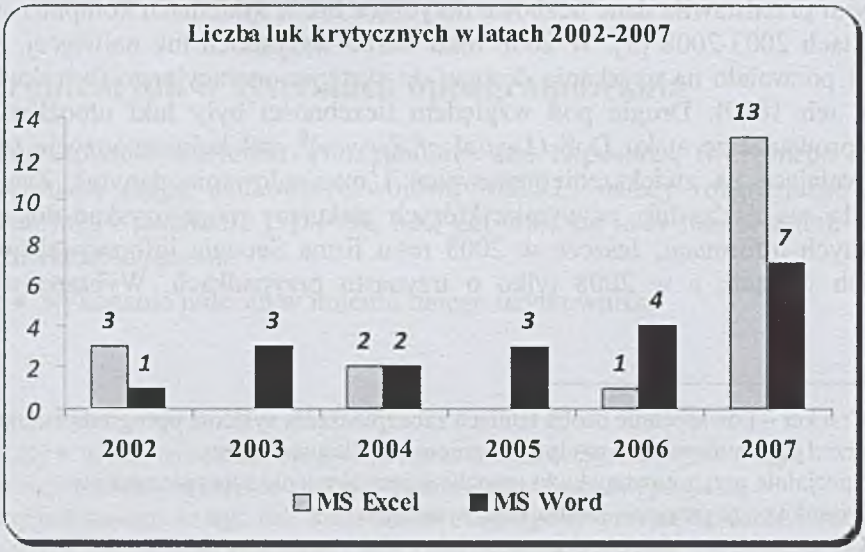
raportu zestawiono w tabeli 1 (w raporcie wszystkie badane wskaźniki zostały szczegółowo zdefiniowane).

Najwięcej luk wykrywa się w powszechnie stosowanym oprogramowaniu, zaś szczególnie ważną rolę odgrywają luki znajdujące się w systemach operacyjnych oraz przeglądarkach internetowych. Najbardziej niebezpieczne są luki, które umożliwiają atakującemu przejęcie kontroli nad systemem (luki krytyczne). Najczęściej exploity nie wykorzystują tylko jednej luki, ale posiadają całą ich bazę (np. dla konkretnego oprogramowania). Łatwo dostępne exploity mogą być wykorzystywane przez osoby nie mające elementarnej wiedzy z zakresu technik hakerskich, dlatego ataki tego typu są niezwykle popularne. Najgorszym z możliwych ataków jest tzw. zero-day attack wykorzystujący szkodliwy kod – publikowany, gdy nie ma jeszcze możliwości usunięcia luki w oprogramowaniu.

Tab. 1. Zestawienie liczby luk w systemach komputerowych latach 2003 i 2008

Wykorzystanie luki	Liczba luk	
	2003 r.	2008 r.
Uzyskanie dostępu do systemu operacyjnego	1020	1814
Przeprowadzenie ataku DoS	817	1538
Zwiększenie uprawnień	471	1040
Manipulowanie danymi	111	1162
Spoofing	45	883
Uzyskanie dostępu do ważnych informacji	482	13

Na rysunku 1 przedstawiono liczbę luk krytycznych w powszechnie używanych programach pakietu Microsoft Office: Excel i Word w latach 2002-2007.



Rys. 1. Liczba luk krytycznych w wybranym oprogramowaniu w latach 2002-2007 (na podstawie [6])

Przykładowo, w maju 2009 roku ujawniono krytyczne luki w serwerowym oprogramowaniu Microsoft Internet Information Services 6, w programach Adobe Reader 9.1 i Acrobat 9.1 (także w starszych wersjach). We wrześniu 2009 r. Laurent Gaffié odkrył krytyczne luki w systemach Microsoft Windows Vista i Windows 7.

Do oszacowania stanu zabezpieczeń systemu komputerowego (identyfikacji słabych punktów) wykorzystuje się specjalistyczne narzędzia, które umożliwiają kontrolę infrastruktury sieciowej dzięki rozpoznawaniu luk w systemach zabezpieczeń, nadawaniu im priorytetów oraz korygowaniu wykrytych wad (jak np. Symantec Vulnerability Assessment czy IBM Express Portal Vulnerability Assessment umożliwiający zdalną kontrolę stanu zabezpieczeń serwerów dostępnych z Internetu). Problemy pojawiają się, gdy baza skanera nie jest aktualna, lub gdy jej aktualizacja jest możliwa później, niż pojawił się exploit wykorzystujący daną lukę (co ma miejsce najczęściej). Po skutecznym włamaniu się do systemu, włamywacz dąży do poszerzenia swoich uprawnień (w przypadku systemu operacyjnego Windows – na poziomie Administrator lub SYSTEM), co też jest możliwe dzięki istnieniu odpowiednich narzędzi. Przykładem skutecznych narzędzi hakerskich pozwalających atakującemu zwiększyć uprawnienia są exploity z rodziny getadmin i inne, również stosujące technikę DLL injection[7].

Dotychczas nie ma jednolitego sposobu porównywania poziomu bezpieczeństwa aplikacji czy systemów o podobnej funkcjonalności. Liczba wykrytych błędów i luk nie zawsze jest dobrym miernikiem jakości systemu czy jego podatności na ataki.

3. Narzędzia do skanowania sieci

Do systematycznego przeprowadzania audytu poziomu ochrony systemu informatycznego dostarcza się wielu narzędzi o różnym stopniu zaawansowania – od prostych, realizujących określoną funkcję, do złożonych systemów wspomagających zarządzanie bezpieczeństwem. Współczesne skanery wykorzystują zarówno bazy zdefiniowanych luk i zagrożeń, jak i zaawansowane algorytmy działające w oparciu o sztuczną inteligencję (pozwalające na analizę systemu bezpieczeństwa pod kątem nie opisanych dotąd zagrożeń). Skaner pozwala na wstępie określić rodzaj testowanego systemu operacyjnego, a następnie wyszukuje znane dla niego zagrożenia (dla różnych systemów definiowane są inne bazy). Wyróżnia się skanery hostów (instalowane lokalnie

[7] Technika umożliwiająca uruchamianie kodu w ramach innego procesu (wymuszając dynamiczne dołączenie biblioteki).

w systemie) i skanery sieci – badające stan zabezpieczeń urządzeń sieciowych, głównie identyfikujące luki powiązane z usługami sieciowymi (jak np. SNMP czy SMTP).

Działanie skanera kończy wyświetlenie raportu o stanie bezpieczeństwa systemu (sieci), często z informacją o sposobie usunięcia zidentyfikowanych zagrożeń. Najczęściej podjęcie określonej akcji (jak np. usunięcie szkodliwego kodu, pobranie odpowiedniej nakładki systemowej, zamknięcie otwartego portu, zakończenie uruchomionego procesu czy usługi) wymaga decyzji administratora. Wszelkie działania bez wiedzy i zgody użytkownika rzadko są podejmowane.

Audyt zwykle obejmuje takie elementy jak:

- NetBIOS (*Network Basic Input/Output System*),
- HTTP (*Hypertext Transfer Protocol*), CGI (*Common Gateway Interface*),
- FTP (*File Transfer Protocol*),
- DNS (*Domain Name System*),
- POP (*Post Office Protocol*), SMTP (*Simple Mail Transfer Protocol*), LDAP (*Lightweight Directory Access Protocol*),
- TCP/IP (*Transmission Control Protocol/Internet Protocol*), UDP (*User Datagram Protocol*),
- serwery baz danych,
- serwery proxy,
- zapory sieciowe i rutery.

Ponadto sprawdzane są rejestry, konta użytkowników, systemy haseł oraz inne uruchamiane usługi, a także podatność na przeprowadzenie ataku DoS.

Niestety, to kierowane do administratorów oprogramowanie w rękach włamywacza może stanowić niebezpieczne narzędzie, pozwalające w pełni poznać atakowany system (czy sieć), jego słabości i luki w systemie zabezpieczeń.

Przykłady można by mnożyć. Jednym z takich narzędzi jest skaner Nessus firmy Tenable Network Security składający się z dwóch komponentów: klienta i serwera. To potężne narzędzie umożliwia w systemie Windows między innymi:

- sprawdzenie instalacji nakładek na znane luki,
- sprawdzenie konfiguracji kont użytkowników,
- testowanie oprogramowania (np. przeglądarki internetowej).

Jako kolejny przykład może posłużyć darmowy program Nmap, który można między innymi uruchomić w środowisku Windows, Linux, Unix, FreeBSD i MacOS. Program ten pozwala:

- określić wersję systemu operacyjnego,

- rozpoznać uruchomione usługi,
- rozpoznać rodzaj zapory sieciowej,
- sporządzić listę komputerów pracujących w danej sieci,
- sporządzić listę otwartych portów, z podaniem numeru, stanu (otwarty/zamknięty, filtrowany/nie), rodzaju używanego protokołu oraz nazwy i wersji programu,
- określić adres MAC skanowanego urządzenia.

Ponadto, dodatkowe opcje programu (po odpowiednim skonfigurowaniu) pozwalają na włączenie trybu omijania zapory sieciowej lub systemu wykrywania ataków.

Innymi przykładami narzędzi administracyjnych, które mogą być wykorzystane do zdobycia informacji o przedmiocie ataku są Microsoft Baseline Security Analyzer oraz LANguard firmy GFI. Pierwszy z programów, bezpłatny, skanuje system operacyjny i aplikacje firmy Microsoft. Między innymi kontroluje parametry konfiguracyjne i ustawienia mające znaczenie przy dostępie do systemu. Dostarcza informacji o:

- braku instalacji dostępnych poprawek systemowych,
- liczbie kont z uprawnieniami administratora,
- słabych hasłach dostępowych lub braku ich stosowania,
- aktywności konta „gość",
- usługach systemowych,
- udostępnianych zasobach.

Program LANguard służy między innymi do skanowania systemu w poszukiwaniu podatności na zagrożenia, do przeprowadzania audytu sieci, a także pozwala na instalację nakładek systemowych. Dodatkową jego opcję stanowi możliwość automatycznego usunięcia oprogramowania, które zostało uznane podczas skanowania za niepożądane (!)

Funkcjonalność popularnych bezpłatnych skanerów sieciowych opisano w tabeli 2. Wszystkie te narzędzia umieszczono na płycie dołączonej do jednego z czasopism z końcem 2008 roku, wobec czego pojawia się pytanie o cel tak powszechnego ich popularyzowania.

Tab. 2. Wybrane skanery sieciowe

Narzędzia	Funkcjonalność
Wireshark (dawny Ethereal)	Wszechstronny analizator ruchu w obrębie lokalnej infrastruktury sieciowej (może być wykorzystany do przechwytywania rozmów przy pomocy komunikatorów, np. Gadu-Gadu)
Advanced IP Scanner	Zaawansowane narzędzie pozwalające sprawdzić aktywność urządzeń pod adresami IP z określonego zakresu; m.in. możliwość ustalenia ich adresów MAC oraz informacji NetBios
LanSpy	Narzędzie pozwalające uzyskać nazwy domen, adresy MAC, informacje NetBios, dane o serwerach, użytkownikach, zasadach bezpieczeństwa, współdzielonych zasobach, sesjach, usługach itp.
Nessus	Umożliwia przeprowadzenie ok. 20 tys. różnych testów bezpieczeństwa
Nmap	Jeden z najlepszych bezpłatnych skanerów zabezpieczeń; rozpoznaje systemy operacyjne, ponad 450 usług i 144 protokoły IP, oraz obsługuje kilkanaście technik skanowania portów

Do wyszukiwania informacji o sieci Unix można na przykład wykorzystać narzędzie PScan. PScan skanuje pliki źródłowe zapisane w języku C w poszukiwaniu nieprawidłowego użycia funkcji w stylu printf, takich jak „sprintf(buffer, variable);" zamiast „sprintf(buffer,"%s", variable);". Ten rodzaj problemów jest źródłem wielu luk w bezpieczeństwie. Łatwym w użyciu narzędziem pozwalającym skanować całe bloki adresów IP i zdalnie atakować systemy, w których zostanie zidentyfikowana luka RPC, jest skaner Kaht II.

Do pozyskiwania informacji o sieciach bezprzewodowych również dostępnych jest wiele narzędzi. Jedne z nich (do wykrywania punktów dostępowych i analizowania ruchu sieciowego) wykorzystują tryb RFMON, dokonując jednocześnie przeskoków po kanałach DSSS[8] (*Direct Sequence Spread Spectrum*), inne umożliwiają skanowanie aktywne polegające na wysyłaniu ramek Probe Request i oczekiwaniu na odpowiedź w postaci ramek Probe Response, z których uzyskuje się wiele istotnych informacji (identyfikator sieci,

[8] Metoda DSSS (wykorzystywana między innymi w standardzie 802.11b) polega na wysyłaniu całego strumienia danych i braku stosowania polaryzacji fal nośnych.

numer kanału, dane dotyczące zabezpieczeń kryptograficznych itp.). Sieci zamknięte nie wysyłają odpowiedzi. Programy do prowadzenia tzw. wardrivingu[9] umożliwiają między innymi:

- analizowanie nagłówków pakietów oraz pól wektora inicjalizującego,
- sprawdzanie nazw sieci SSID/ESSID (*Service Set Identifier*, *Extended SSID*),
- sprawdzanie adresów MAC (*Media Access Control*),
- identyfikację używanych zabezpieczeń (WEP[10], WPA[11]) lub wykazanie ich braku,
- sprawdzenie zasięgu sygnału,
- uzyskanie informacji o używanych protokołach,
- zdobycie informacji o adresach IP.

Przykładami programów do prowadzenia wardrivingu są NetStumbler i MiniStumbler (dla środowiska Windows) oraz Kismet (dla systemu Linux i BSD).

Inne dostępne narzędzia (np. StumbVerter działający w oparciu o dane NetStumblera czy aplikacja GPSMap dla programu Kismet) pozwalają automatycznie sporządzić mapę lokalizacji wykrytych punktów dostępowych. Wiele z nich współpracuje z kartami bezprzewodowymi ustawionymi w tryb odbioru, działają w trybie tekstowym lub graficznym, niektóre monitorują ruch w czasie rzeczywistym. Wśród snifferów[12] dla sieci bezprzewodowych można wymienić: Mognet, tcpdump, airfart, THC-Wardrive, Wellenreiter oraz AiroPeek NX. Skanerami sieci bezprzewodowych identyfikującymi punkty dostępowe są też WifiScanner i Gtkskan.

4. Kompresja i ukrywanie kodu

Przykładem legalnego, a niosącego niebezpieczne konsekwencje oprogramowania jest oprogramowanie zabezpieczające jednej z chińskich firm, które zawiera tak ukryte katalogi, by były niewidoczne dla programów antywirusowych. Posłużono się w tym celu technikami stosowanymi przez

---

[9] Wardriving – proceder wykrywania sieci bezprzewodowych i zbierania o nich informacji podczas przemieszczania się z wykorzystaniem samochodu.

[10] *Wired Equivalent Privacy* – standard IEEE działający w warstwie łącza danych wprowadzający ochronę danych przed pasywnym podsłuchem.

[11] *Wi-Fi Protected Access* – standard IEEE kompatybilny z WEP wprowadzający mocniejsze mechanizmy zabezpieczeń; jego następcą jest WPA2 (brak kompatybilności z wcześniejszymi rozwiązaniami).

[12] Podstawowe narzędzie używane do prowadzenia podsłuchu.

rootkity[13]. Oprogramowanie jest fabrycznie instalowane w klipsach USB wykorzystujących technologie biometryczne. Niepokój budzi fakt, iż niewidoczne katalogi mogą zostać wykorzystane do ukrywania szkodliwego kodu [4].

Kryptografia pozwala na zabezpieczenie danych przed niepowołanym dostępem, jednak są one nadal widoczne w postaci kryptogramu, dlatego też do jego ukrycia (zamaskowania) wykorzystuje się dodatkowo metody steganografii.

Do szyfrowania stron WWW można wykorzystać dostępne narzędzia, takie jak HTML Cipher, HTML Guard, HTML Protector, HTML Guardian, Advanced HTML Encrypt and Password Protect, Web Page Protector, TagsLockPro czy bibliotekę mcrypt. Powstało również wiele narzędzi służących do celów steganograficznych, jak np. bezpłatny program S-Tools lub VSL (*Virtual Steganographic Laboratory*). Wyniki badania przeprowadzonego przez firmę McAfee wykazują rosnącą tendencję wykorzystania technologii ukrywania obecności kodu w szkodliwym oprogramowaniu. Obserwuje się także wzrost liczby tzw. potencjalnie niepożądanych kodów w oprogramowaniu komercyjnym. W ciągu ostatnich trzech lat liczba przypadków wykorzystania technologii ukrywania kodu wzrosła o ponad 600%. Według firmy McAfee ten nagły wzrost może być związany z działaniami obejmującymi sieciową współpracę wielu witryn WWW zawierających setki linii szkodliwego kodu dostępnego do rekompilacji, adaptacji oraz udoskonalania, wraz z binarnymi plikami wykonawczymi tego typu oprogramowania [8].

5. Odgadywanie i odczytywanie haseł

Użytkownikom dostarcza się również narzędzi, które pozwalają odtworzyć treść standardowo kodowanych plików z wykorzystaniem słabego, kilkuznakowego klucza. Programy te generują wszystkie możliwe ciągi znaków określonej długości, lub próbują szybciej odgadnąć klucz wykorzystując gotowe słowniki popularnych haseł. Do takich aplikacji należą programy ZIP Password Finger (dla plików .zip) oraz GG Tools 2.4 – do odzyskiwania haseł i poznania treści rozmów użytkowników komunikatora Gadu-Gadu.

Narzędzia do odgadywania haseł są zwykle wykorzystywane offline do przeszukiwania przechwyconych plików zawierających hasła dostępowe. Przykład może stanowić oprogramowanie L0phtcrack przeznaczone dla systemu operacyjnego Windows 2000. Zawiera ono również funkcję pozwalającą na podsłuch lokalnego segmentu sieci i przechwytywanie poszczególnych sesji logowania pomiędzy systemami Windows w celu wyodrębnienia wartości przydatnych do odtworzenia hasła, które są poddawane analizie przez główną

[13] Specjalizowane, pomocnicze we włamaniach narzędzie służące do ukrywania niebezpiecznych plików i procesów umożliwiających utrzymanie kontroli nad systemem operacyjnym.

część programu. Narzędzie łamie hasło metodą siłową (brute force)[14], a ponieważ stosowany przez Microsoft mechanizm jest słaby kryptograficznie oraz fragmentuje uzyskane skróty haseł, co stwarza możliwość równoległego ich łamania – efekty uzyskuje się w stosunkowo krótkim czasie [1]. Programiści z L0pht opracowali też sniffera, który przechwytuje skróty haseł systemu Windows w trakcie sesji logowania protokołu PPTP (*Point-to-Point Tunneling Protocol*), który jest wykorzystywany do tworzenia połączeń VPN (*Virtual Private Network*) w Internecie.

Od maja 2009 r. program L0phtcrack w. 6 dostępny jest dla systemów 64-bitowych (Vista, Windows 7 i Unix). Odpowiednikami tego programu są: Cain & Abel (bezpłatny program umożliwiający odzyskiwanie i łamanie haseł w systemie Windows oraz śledzenie pakietów sieciowych) oraz John the Ripper (jeden z najpopularniejszych programów do testowania i łamania haseł za pomocą ataku słownikowego lub metodą siłową). Wymienione programy nie mogą być wykorzystywane w celach niezgodnych z prawem [9].

Wspomniany już wielofunkcyjny program LANguard firmy GFI pozwala na sprawdzanie mocy haseł poprzez przeprowadzanie ataku słownikowego. Taki atak można wykorzystać do kontroli haseł używanych na przykład w protokole SNMP czy haseł dostępu do konta administratora SQL Serwera.

Dostępne są również różnego typu narzędzia, tzw. keyloggery (istnieją rozwiązania programowe i sprzętowe), pozwalające podglądać wprowadzane z klawiatury treści. Przykładem może być urządzenie firmy Thumbs Up o nazwie Keyshark wpinane pomiędzy klawiaturę a port USB komputera. Współpracuje ono z dowolną klawiaturą i każdym systemem operacyjnym, a jego zadaniem jest rejestracja w wewnętrznej pamięci wartości naciskanych przez użytkownika klawiszy. Producent zapewnia niewykrywalność urządzenia, jest ono reklamowane jako produkt do monitorowania działań pracowników czy użytkowników nieletnich, oraz jako narzędzie pomocne programistom, podczas gdy jest to po prostu sprzęt szpiegowski [4].

W przypadku podsłuchiwania sieci bezprzewodowych dostępne są także bardziej zaawansowane narzędzia, które pozwalają dodatkowo skutecznie łamać klucze kryptograficzne wykorzystywane w procesie szyfrowania. Wykorzystując słabości stosowanych mechanizmów zabezpieczeń umożliwiają szybkie i skuteczne ich przełamywanie. Są wśród nich zarówno narzędzia darmowe, jak i drogie programy komercyjne umożliwiające prowadzenie profesjonalnych rekonesansów.

---

[14] Efektywność metody zależy przede wszystkim od długości klucza kryptograficznego oraz wydajności systemu komputerowego.

6. Narzędzia do atakowania protokołów i algorytmów kryptograficznych

Systemy kryptograficzne mogą zapewniać poufność, integralność i autentyczność danych oraz umożliwić weryfikację źródła ich pochodzenia (np. autentyczność nadawcy). Zabezpieczają również przed kopiowaniem danych, czyniąc je bezużytecznymi dla nieuprawnionego odbiorcy. System kryptograficzny musi spełniać trzy podstawowe warunki:

- przekształcenia szyfrujące i deszyfrujące są efektywne dla wszystkich kluczy,
- użytkowanie systemu jest łatwe,
- jego bezpieczeństwo zależy od poufności kluczy a nie algorytmów.

Ponieważ żaden system nie jest całkowicie bezpieczny – każdy element systemu kryptograficznego (generator kluczy, protokół ich dystrybucji, klucze, szyfrator, deszyfrator itd.) może być przedmiotem ataku. Nawet jeśli zastosowano bardzo mocne mechanizmy gwarantujące wysoki poziom ochrony (np. algorytmy obliczeniowo bezpieczne[15]), istnieje ryzyko niewłaściwego ich stosowania lub popełnienia błędu. Użytkownicy wykorzystując dostępne systemy często nie przestrzegają podstawowych zaleceń, i tak na przykład używają krótkie i słabe kryptograficznie hasła, które poddają się atakowi słownikowemu przy pomocy dostępnych narzędzi, z kolei przy trudnych do zapamiętania hasłach – zapisują je. Z badań wynika, że zalecane najmocniejsze mechanizmy zabezpieczeń rzadko są stosowane. Korzystając z dostępnych narzędzi można przeprowadzić atak na większość używanych standardów kryptograficznych. W ostatnim okresie atrakcyjny i stosunkowo łatwy cel ataków stanowią sieci bezprzewodowe. Opracowano wiele narzędzi umożliwiających przeprowadzenie skutecznego ataku na ich zabezpieczenia (patrz tabela 3).

Tab. 3. Wybrane narzędzia do atakowania sieci bezprzewodowych

Narzędzia	Funkcjonalność

[15] Bezpieczeństwo obliczeniowe szyfru określane jest czasem koniecznym do złamania szyfru przy wykorzystaniu mocy aktualnie dostępnych komputerów.

Bwmachak, SMAC, wicontrol	Zmiana adresu MAC na adres zaufany
Aircrack, AirSnort, WepAttack, WepDecrypt, coWPAtty, KisMAC, wpa_crack	Przechwytywanie pakietów w celu złamania zabezpieczeń i uzyskania dostępu do sieci
Airpwn, File2air, void11, WEPWedgie	Tworzenie i wysyłanie spreparowanych ramek protokołu
Dsnif, Ettercap	Przekierowanie strumienia danych (atak Man-in-the-middle)
AirJack, Omerta, void11, aireplay	Zalanie stacji pakietami zawierającymi polecenie odłączenia od punktu dostępu

Pierwszy z proponowanych standardów bezpieczeństwa dla sieci bezprzewodowych – WEP okazał się niezwykle podatny na ataki. Można się o tym przekonać wykorzystując oprogramowanie Aircrack, zawierające trzy narzędzia służące do przeprowadzenia każdej z trzech faz ataku na klucz WEP:

- Airodump – sniffer do wykrywania sieci,
- Aireplay – narzędzie do wprowadzania pakietów (dla skrócenia czasu łamania),
- Aircrack-ng – narzędzie do łamania klucza (nowsza wersja: Aircrack-ptw).

Innym popularnym narzędziem do ataków na WEP (wykorzystującym słabość wektora inicjalizującego algorytmu szyfrującego RC4) jest AirSnort. Stanowi on zbiór skryptów i programów dla systemu Linux do przechwytywania i łamania pakietów sieci bezprzewodowych. Posiada też możliwość określania siły klucza WEP – standardu bezpieczeństwa dla sieci Wi-Fi, który (mimo słabości) jest nadal najczęściej wykorzystywany.

Kolejny przykład możliwości atakowania mechanizmów bezpieczeństwa sieci bezprzewodowych stanowi protokół LEAP (*Lightweight Extensible Authentication Protocol*) Cisco, którego działanie opiera się na wykorzystaniu zdalnego serwera RADIUS[16]. Do przeprowadzenia ataku na słabo zabezpieczone urządzenia z włączoną obsługą LEAP można użyć program Anwrap współpracujący z narzędziem ancontrol, lub program Asleap, których zadaniem jest przechwytywanie i deszyfrowanie słabych haseł protokołu z punktów dostępowych Cisco oraz kart bezprzewodowych. Program Asleap działając w trybie RFMON na bieżąco odczytuje ruch sieciowy i, integrując się z programem AirJack, odłącza uwierzytelnionych użytkowników od sieci, a następnie podsłuchuje i łamie hasła użytkowników, którzy ponownie próbują uwierzytelnić się w punkcie dostępowym.

[16] *Remote Authentication Dial In User Service* – serwer zdalnego uwierzytelniania realizujący daną usługę.

Przedmiotem ataku może też być popularny protokół uwierzytelniania użytkowników sieci lokalnych Kerberos, którego słabym punktem jest moment uwierzytelniania użytkownika w serwerze uwierzytelniającym (istnieje możliwość przechwycenia komunikatu o znanym formacie między serwerem a użytkownikiem). Protokół jest podatny na ataki na klucz długoterminowy[17] – są dostępne programy umożliwiające przeprowadzenie takiego ataku.

Atakuje się również tzw. bezpieczne kanały transmisyjne, czyli stosujące szyfrowanie przesyłanych danych, takie jak SSH (*Secure Shell*) czy SSL (*Secure Socket Layer*). Na przykład protokół SSH jest podatny na ataki typu Man-in-the-middle, bowiem możliwe jest użycie protokołu bez wcześniejszej weryfikacji klucza i serwera. Jest to możliwe, gdy klient zaakceptuje klucz podany przez atakującego jako klucz serwera (np. atakujący ma dostęp do klucza publicznego serwera lub klient nie ma informacji o aktualnym kluczu serwera) lub atakujący może próbować zamieniać pakiety w czasie transmisji już po ustaleniu sesji (przed tym ma zabezpieczać algorytm MAC). Zatem bezpieczeństwo protokołu SSH w znacznym stopniu zależy od bezpiecznej dystrybucji kluczy. SSH wykazuje też podatność na atak DoS, gdyż atakujący może zmusić serwer do wykonywania obciążających operacji ustanawiania połączenia i wymiany kluczy bez konieczności uwierzytelnienia.

W grudniu 2008 roku podano do publicznej wiadomości informację o odkryciu słabości w systemie certyfikatów protokołu SSL. Zostało bowiem udowodnione, że można stworzyć (podrobić) certyfikat dla dowolnej strony WWW i mimo fałszerstwa jest on akceptowany przez większość popularnych przeglądarek internetowych. Metoda ta, w której wykorzystuje się słabość funkcji haszującej MD5[18] może zostać wykorzystana na przykład do wystawiania certyfikatów fałszywym stronom bankowym. Już w 2004 roku chińscy naukowcy przedstawili metodę umożliwiającą wygenerowanie tego samego 128-bitowego skrótu MD5 dla dwóch różnych informacji. Firmy posługujące się algorytmem MD5 przy tworzeniu certyfikatów powinny zastąpić go mocniejszym mechanizmem, np. SHA-1, SHA-2, SHA-3 [4].

7. Inne niebezpieczne praktyki

Nowy problem stanowi sprzedaż miejsca reklamowego przez cieszące się zaufaniem korporacje stronom znanym z rozprzestrzeniania szkodliwego oprogramowania. Pod koniec 2008 roku analitycy bezpieczeństwa poinformowali o odkryciu w usłudze Google Ads. reklamy oprogramowania antywirusowego kierującej użytkowników na niebezpieczną stronę zawierającą

[17] Klucz kryptograficzny używany między elementami zabezpieczeń i centrum dystrybucji kluczy w procesie uwierzytelniania.

[18] *Message Digest Algorithm* – funkcja mieszająca opracowana przez Rona Rivesta dająca w wyniku 128-bitowy skrót.

fałszywe oprogramowanie antywirusowe. Przykładem jest oprogramowanie Antivirus XP 2008, które zostało sklasyfikowane jako jedno z najbardziej uciążliwych zagrożeń owego roku. Pojawia się na komputerze użytkownika w postaci wyskakującego okienka informując o odnalezieniu wirusów, proponuje darmowe skanowanie systemu, a następnie wyświetla listę rzekomo zidentyfikowanych zagrożeń wraz z ofertą ich usunięcia po uiszczeniu stosownej opłaty. Aplikacja jest trudna do usunięcia i potrafi pobierać aktualizacje umożliwiające jej ukrywanie się przed specjalistycznym oprogramowaniem antywirusowym. Nawet ostrożny użytkownik nie ustrzeże się przed takim zagrożeniem mając zaufanie do odnośników umieszczanych w boksie firmowanym przez Google. W odpowiedzi na stawiane zarzuty rzecznik firmy zapewnił, iż nie szczędzi ona wysiłków w zapewnieniu bezpieczeństwa swoim użytkownikom i reklamodawcom [4].

Dzięki powszechnemu udostępnianiu (np. na płytach CD dołączanych do specjalistycznych czasopism) gotowych „niebezpiecznych" narzędzi, podejmowanie różnych szkodliwych, sprzecznych z prawem działań staje się możliwe dla przeciętnego użytkownika Internetu. Na przykład do kopiowania płyt CD/DVD i uruchomienia ich obrazów w wirtualnych napędach dostarcza się takich narzędzi jak Daemon Tools, CDBumerXP Pro 4.1.2 czy DeepBumer 1.9. Najlepsze z nich są płatne – wśród nich: CloneCD 5.3.1.3 tworzący kopie dysków zarówno z muzyką jak i danymi, aplikacja Alcohol 120% 1.9.7 potrafiąca odczytać wszystkie najpopularniejsze obrazy płyt, czy Gamejack 5.0.4.4 kopiujący (po złamaniu zabezpieczeń) płyty zawierające gry komputerowe. Niestety, polskie prawo nadal podchodzi dość liberalnie do kwestii kopiowania utworów na własny użytek.

Do crackowania[19] programów udostępniane są niezbędne narzędzia programistyczne, jakimi są edytory szesnastkowe. Nieuprawnione modyfikowanie komercyjnych aplikacji oraz ich wykorzystywanie, to działania nielegalne, jednak w celu łamania zabezpieczeń oprogramowania można sięgnąć po takie (często darmowe) narzędzia jak: Cambiator 1.04, Hackman Hex Editor Lite 8.02, czy rozbudowane funkcjonalnie (wyposażone w klienta FTP) – UltraEdit 13.10.

W 2005 roku w polskim Internecie pojawił się bezpłatny program Krakersik służący do wyszukiwania numerów seryjnych i aplikacji łamiących zabezpieczenia. Aplikacja wzbudziła wprawdzie kontrowersje, ale i duże zainteresowanie – co przyczyniło się do opracowania jej kolejnych, udoskonalonych wersji. Autor programu na łamach jednego ze specjalistycznych czasopism oświadcza, że podczas tworzenia narzędzia to nie chęć niesienia pomocy w popełnianiu przestępstwa mu przyświecała, lecz potrzeba sprawdzenia opracowanego mechanizmu wyszukiwawczego (!).

---

[19] Usuwania zabezpieczeń z legalnego oprogramowania.

Adresy komputerów, które zostały zidentyfikowane jako prowadzące w sieci szkodliwą działalność (np. rozsyłające spam) są często blokowane (np. mogą być objęte blokadą dostępu do stron WWW). Dla potrzeb omijania takich problemów dostarczane są użytkownikom przeglądarki pozwalające ukrywać adres IP komputera – przykładem przeglądarka The Torpark 1.5.0.7 oparta na systemie Firefox Mozilli. Jeśli użytkownik chce korzystać z zasobów Internetu w trybie anonimowym[20], może sięgnąć do takich narzędzi jak przeglądarka Internet Explorer 8 beta 2 (opcja inPrivate Browsing), dodatek Mozilli Stealther 1.0.6 do Firefox'a 3.0, czy Google Chrome (tryb Incognito).

Istnieją nawet programy do różnego typu niepożądanych działań w serwisach społecznościowych. Na przykład do symulowania obecności użytkownika na stronie oGame.pl można użyć programu Ogamowiec 1.2. Inne programy służą do rozsyłania ogłoszeń i ofert. Przykładem automatycznego odwiedzania stron WWW jest popularna aplikacja Fotka Manager – kiedyś sprzedawana na aukcjach internetowych dla serwisu fotka.pl, której działanie może być traktowane jako rodzaj spamowania (obecnie darmowy program nie jest oficjalnie do zdobycia, ale pojawiają się jego pirackie kopie).

8. Podsumowanie

Wymienione programy wykorzystywane do atakowania systemów i sieci komputerowych stanowią jedynie wybrane przykłady popularnych narzędzi.

Wzrost aktywności przestępczej w sieciach komputerowych wiąże się z dostępnością szkodliwego oprogramowania, na które popyt nieustannie rośnie. Ogromny wybór narzędzi pozwalających skutecznie atakować systemy informatyczne powoduje konieczność ich reklamowania przez twórców (!), czego przykładem może być zachęcanie do kupna (np. koń trojański Limbo 2) poprzez gwarantowanie niewykrywalności oprogramowania przez powszechnie stosowane programy ochronne w odpowiednio długim okresie (w przeciwnym wypadku obietnica dostarczenia bezpłatnej udoskonalonej wersji).

Coraz częściej słyszy się głosy wskazujące na potrzebę ograniczenia dystrybucji szkodliwego oprogramowania, gdyż jego dostępność ułatwia przestępczą działalność w sieci. Przykładem jest zakaz wydany przez sąd na Florydzie dystrybucji keyloggerów o nazwie RemoteSpy przez firmę Cyberspy Software LLC. Ponadto wydano firmie nakaz wyłączenia wszelkich serwerów zbierających informacje przy ich pomocy. Pozew do sądu wniesiony został przez Federalną Komisję Handlu. W orzeczeniu sąd stwierdził, że z powodu praktyk pozwanej firmy konsumenci ponieśli straty materialne, podczas gdy ona sama czerpała finansowe korzyści [4].

---

[20] Bez pozostawiania śladów związanych z podejmowanymi działaniami, tj. bez zapisywania danych dotyczących otwieranych stron WWW.

Potencjalnie każdy komputer jest narażony na atak ponieważ włamywacze za pomocą posiadanych narzędzi skanują komputery podłączone do sieci globalnej w poszukiwaniu tych, na których luki nie zostały usunięte. Osiągnięcie stuprocentowego zabezpieczenia systemu nie jest możliwe, jednak dla maksymalnego bezpieczeństwa należy na bieżąco aktualizować system operacyjny i jego komponenty, oprogramowanie ochronne oraz wszelkie używane aplikacje użytkowe, a także używać najmocniejszych dostępnych zabezpieczeń i, co ważne, wprowadzić mechanizmy systemowo wymuszające ich stosowanie.

LITERATURA

1. McClure S., Scambray J., Kurtz G.: Hacking zdemaskowany. Bezpieczeństwo sieci – sekrety i rozwiązania. PWN, Warszawa 2006.
2. Internet. Agresja i ochrona. Tytuł oryginału: Maximum security: A Hacker's Guide. Robomatic, 1998 (Autor anonimowy).
3. Secunia Stay Secure. 2008 Report.
4. www.bezpieczenstwo.onet.pl
5. www.cve.mitre.org/about/terminology.html
6. www.sans.org/top20
7. www.sophos.com
8. www.telix.pl/artykul/rosnaca-liczba-technologii-ukrywania-kodu-we-wrogim-oprogramowaniu-3,13527.html
9. www.heise-online.pl/news/L0phtcrack-powrocil--/8836

Rozdział 5

Problem archiwizacji kluczy szyfrujących oraz metody dzielenia sekretu

Daniel Arendt
Politechnika Łódzka
arendt@zsk.p.lodz.pl
Krzysztof Lichy
Politechnika Łódzka
lichy@zsk.p.lodz.pl

Streszczenie

Dynamiczny rozwój elektronicznego przekazu informacji wymusił na użytkownikach konieczność stosowania technologii kluczy kryptograficznych. Są one używane nie tylko do szyfrowania danych ale także do autoryzacji oraz w coraz większym stopniu do podpisu elektronicznego. Wzrost popularności takich rozwiązań spowodował pojawienie się zupełnie nowego problemu jakim jest utrata klucza bądź trywialne zapomnienie przez użytkownika hasła. W rezultacie powodujące bezpowrotną utratę zaszyfrowanych danych. Z kolei fakt posiadania przez użytkownika kluczy ważnych w minionym okresie czasu wymusza konieczność ich przechowywania w celu umożliwienia dostępu do danych archiwalnych. Należy pamiętać, że przy zastosowaniu klucza do podpisu samo istnienie kopii klucza poza obszarem władzy użytkownika stawia pod znakiem zapytania niezaprzeczalność osobistego użycia podpisu. W pracy omówiono tę problematykę i typowe rozwiązania w szczególności oparte na segmentowaniu kluczy i przechowywaniu ich w postaci rozproszonej. Podano także opracowany algorytm dzielenia sekretu na wielu powierników. Odbudowa klucza następuje z użyciem operacji sumy logicznej bitów, w sposób zbliżony do algorytmu rekonstrukcji danych stosowanego w dyskach systemu RAID. Zaletą tej metody dzielenia sekretu jest możliwość zaplanowania zmiennej liczbę powierników klucza koniecznych do rekonstrukcji i wśród nich wskazać dwie kategorie: niezbędnych i opcjonalnych, co pozwoli na wprowadzenie bardziej złożonej polityki bezpieczeństwa.

1. Wstęp

W świetle obowiązujących rozwiązań prawnych podpis elektroniczny złożony z użyciem klucza może być równoważny klasycznemu podpisowi ręcznemu. Do tego celu wykorzystuje się parę kluczy niesymetrycznych z częścią prywatną, niejawną i częścią publiczną zazwyczaj wyposażone w certyfikat instytucji wydającej i włączone w hierarchiczną strukturę poświadczenia tożsamości PKI. Klucze takie znajdują zastosowanie głównie w trzech obszarach: szyfrowanie danych, podpis elektroniczny i autoryzacja dostępu do danych. Wymusza to przestrzeganie przez wydawców kluczy zdefiniowanych rygorów organizacyjno technicznych, które w istocie dotyczą staranności w używaniu środków elektronicznych tak, aby kradzież klucza nie była możliwa. Utrata klucza może spowodować czasami bezpowrotną utratę danych lub narazić użytkownika na poważne koszty. Naturalnym wydaje się zatem opracowanie i wdrożenie mechanizmu umożliwiającego kopiowane i przechowywane kluczy w bezpiecznym miejscu. Zauważyć jednak należy że każda kopia poza kontrolą użytkownika stwarza możliwość nadużyć. Istnieje szereg modelowych rozwiązań organizacyjnych i technicznych w tym zakresie. To zagadnienie stanowi tło pracy, choć zasadniczą uwagę poświęcono stronie technicznej awaryjnego odzyskiwania danych, którymi typowo są klucze szyfrujące, ale także często inne poufne dane. Skupiono się na schematach przechowywania kluczy w postaci segmentowanej, rozdzielonej na wielu powierników w taki sposób, że odtworzenie klucza, czy ogólniej dowolnego sekretu który możemy zapisać jako liczba binarna, jest możliwe tylko po zgromadzeniu progowej liczby segmentów. Od strony organizacyjnej oznacza to uzyskanie zgody wielu powierników na wykonanie takiej operacji. Zaprezentowane metody umożliwiają bezpieczne przechowywanie prywatnych danych użytkownika takich jak na przykład hasła, kody PIN lub inne poufne dane. Umożliwiają również przechowywanie kluczy szyfrujących w celu zabezpieczenia przed umyślnym lub niezamierzonym zniszczeniem. Rozwiązuje to również problem zagubieniem hasła. Opisane metody pozwalają również na przechowywanie danych autoryzujących dostęp w kilku magazynach w celu uniknięcia problemów awarii „*single point of failure*". Jednocześnie jest to rozwiązanie problemu archiwizacji i przechowywania dokumentów poufnych. Zasadniczymi zaletami opisanych rozwiązań jest replikacja danych co zabezpiecza przed techniczną awarią nośnika oraz rozproszenie znacznie utrudniające nieuprawnione pozyskanie danych.

2. Problem archiwizacji kluczy

Klucze cyfrowe mają trzy obszary zastosowań: szyfrowanie danych, podpis elektroniczny i autentykacja i autoryzacja dostępu[3]. W każdym z tych obszarów problem utraty klucza i konieczności jego odtworzenia ma inne znaczenie. Powszechnie wydaje się jeden klucz do wszystkich zastosowań. Zatem użytkownik ma w systemie PKI jeden klucz prywatny zwykle zabezpieczony dodatkowo hasłem i odpowiadający mu klucz publiczny powszechnie dostępny. Większa liczba kluczy, w istocie par kluczy, to niepotrzebna komplikacja.

Tab. 1. Przeznaczenie pary kluczy PKI, a potrzeba archiwizacji [4, 5]

Obszar zastosowań	znaczenie archiwizacji, tworzenia kopii
Szyfrowanie danych	często istnieje potrzeba
Autentyfikacja oparta na: podpisie lub słowie szyfrowanym	bez znaczenia, nie ma potrzeby
Podpis elektroniczny, niezaprzeczalność podpisu	Może podważyć wymaganie niezaprzeczalności

Istnieje sprzeczność pomiędzy potrzebą zachowania kopii klucza w bezpiecznym miejscu ze względu na dostęp do szyfrowanych danych, a negatywnym wpływem możliwości dostępu osób postronnych do kopii klucza używanego do podpisu. Większość typowych schematów organizacyjnych bezpieczeństwa powierza klucz opiece użytkownika i czyni go odpowiedzialnym za bezpieczne przechowywanie. Wtedy tylko można mówić o prawdziwej niezaprzeczalności podpisu. Jeśli użytkownik utraci klucz z jakiegokolwiek powodu dostaje nowy, a stary jest unieważniany od chwili zgłoszenia utraty. Sam użytkownik, dbając o bezpieczeństwo klucza może poddać go archiwizacji i na przykład przechowywać w postaci rozproszonej w kilku tylko jemu znanych miejscach. Klucz przeznaczony do autoryzacji dostępu powinien być wymieniony w razie utraty i tu jego archiwizacja nie ma znaczenia. Dla odmiany klucz używany do szyfrowania może spowodować nieodwracalną utratę danych. W tym obszarze zastosowań archiwizacja jest potrzebna. Tym bardziej potrzebna, że często:

- dane szyfrowane kluczem pracownika nie są jego własnością, ale własnością firmy
- klucze są okresowo wymieniane, istnieje odczytu dotarcia do danych szyfrowanych dawno

Złożoność problemu tworzenia kopii kluczy wymaga omówienia kilku przypadków. Nie będą tu przedstawiane problemy związane z dobrze znanym [12 ,13], kontrowersyjnym problemem wokół technologii Key Escrow System (KES), która pozwala instytucjom rządowym w „uzasadnionych" przypadkach uzyskiwać dostęp do danych poufnych.

2.1. Klucze dualne

Jak opisano powyżej sama potencjalna możliwość poznania, przez osoby trzecie poza właścicielem, klucza służącego do podpisu stawia pod znakiem zapytania najważniejszy atrybut podpisu, którym jest niezaprzeczalność, bo w istocie może zrobić to inna osoba. Z drugiej strony w zastosowaniu do szyfrowania, utracony klucz to utrata informacji. Przedsiębiorstwo zatrudniając pracownika i wydając mu klucz szyfrujący nie chce polegać na pamięci i dobrej woli pracownika, który może klucz zgubić, zapomnieć hasło, a zaszyfrowane dane często nie stanowią własności pracownika ale własność pracodawcy.

Dlatego dwie liczące się na rynku rozwiązań kryptograficznych firmy Verisign i Entrust wprowadziły, jako jeden ze schematów bezpieczeństwa, zestaw dwóch par kluczy tzw. „*dual key pairs*", aby rozwiązać ten problem tam, gdzie zachodzi potrzeba. Dwa klucze prywatne i dwa publiczne, razem cztery klucze są wydawane zamiast tradycyjnie jednej pary. Jedna para służy do podpisu, druga do szyfrowania. Tylko para, a właściwie klucz prywatny z pary przeznaczonej do podpisu mogą być archiwizowane. Co więcej szczegółowe procedury generowania kluczy wskazują, że para przeznaczona do podpisu jest generowana w siedzibie firmy i klucz prywatny jest zapisywany tylko jeden raz na jednym nośniku danych. Odmiennie para do szyfrowania jest generowana u wydawcy kluczy i zazwyczaj archiwizowana jeszcze przed wydaniem [5].

Ten schemat bezpieczeństwa wydaje się być najlepszym z punktu widzenia dbałości o zapewnienia niezagrożonego dostępu do danych firmy oraz z drugiej strony warunki zachowania niezaprzeczalności podpisu także są spełnione. Istnieje nieznacznie większa komplikacja techniczna związana z koniecznością obsługi większej liczby kluczy, ale ze strony użytkownika nie ma to znaczenia. Karta kryptograficzna lub podobny nośnik kluczy może magazynować wiele kluczy jednocześnie.

2.2. Przechowywanie kluczy aktywnych i archiwalnych

W przypadku instytucji funkcjonującej w wieloletnim okresie czasu, czyli niemal każdej, pojawia się problem archiwizacji zasobów dokumentów elektronicznych, poczty elektronicznej, dokumentacji działalności w tym dokumentacji finansowej oraz umów i kontraktów[11]. Wiele z nich ma dziś postać cyfrową szyfrowaną i podpisaną cyfrowo, w takiej zostało archiwizowane. Jak wrócić do tych danych po latach czy też jak zweryfikować podpisy elektroniczne złożone przed laty? Niezbędne jest przechowywanie

kluczy publicznych wydanych historycznie, aby sprawdzić podpisy. Tu wydaje się, że potrzeba archiwizacji kluczy prywatnych archiwalnych i nieaktywnych jest oczywista, w zestawieniu z kluczami, które są wciąż czynne. Nieaktywne klucze nie mogą zostać zastosowane do skutecznego złożenia podpisu. Ale warto pamiętać o możliwości umyślnego zniekształcenia przeszłości, przez wprowadzenie nowych, lub wymianę istniejących dokumentów historycznych. Dlatego w zasadzie nie ma różnicy w traktowaniu kluczy czynnych i nieczynnych. Choć powszechnie wydaje się, że unieważnione, przestarzałe klucze powinny być archiwizowane bezwarunkowo [11]. Jeśli już zostanie podjęta decyzja o archiwizacji kluczy to zachodzi potrzeba zdefiniowania schematu organizacyjno technicznego przechowywania kluczy. Tu jedną z ciekawych i skutecznych metod jest przechowywanie kluczy w rozproszeniu, szczególnie w takich sposób, że po rozproszeniu na n powierników trzeba liczby progowej t spośród nich aby klucz odtworzyć. Metoda ta może zostać zastosowana przez posiadacza klucza do ochrony klucza przed zgubieniem.

3. Rozproszenie danych poufnych

W kryptografii problem dzielenia sekretu, to jakakolwiek metoda podziału tajemnicy pośród grupę powierników na fragmenty w taki sposób, że każdy z powierników nie może poznać tajemnicy i do jej odtworzenia potrzeba złożenia fragmentów wszystkich lub tylko zaplanowanej liczby fragmentów[1, 2].

3.1. Dzielenie sekretu – schemat Shamir'a Definicja schematu Shamir'a (dzielenia sekretu)

Najczęściej problem definiuje się następująco. Dealer posiadający sekret i rozdziela go pośród n powierników. Do odtworzenia sekretu konieczne jest spotkanie co najmniej t powierników spośród n [1]. Taki schemat nazywamy schematem progowym i oznaczamy (t,n). Jeden z pierwszych historycznie schematów znany powszechnie jako *Shamir's secret sharing scheme* (SSSS) został opracowany jednocześnie i niezależnie przez Adiego Shamira [1] i Georga Blakleya [2] w 1979 roku. Schemat ten opiera się na użyciu wielomianu stopnia t 1, gdzie t to progowa liczba powierników niezbędna do rekonstrukcji sekretu, zaś s0 jest sekretem.

$$S(x) = s_0 + \sum_{i=1}^{t-1} a_i x_i \tag{1}$$

Współczynniki ai wielomianu są wybrane arbitralnie. Dealer generuje n>=t unikalnych par liczb [x, S(x)] i rozdziela powiernikom. Są to fragmenty sekretu. Po zebraniu t takich par i po zastosowaniu interpolacji wielomianowej

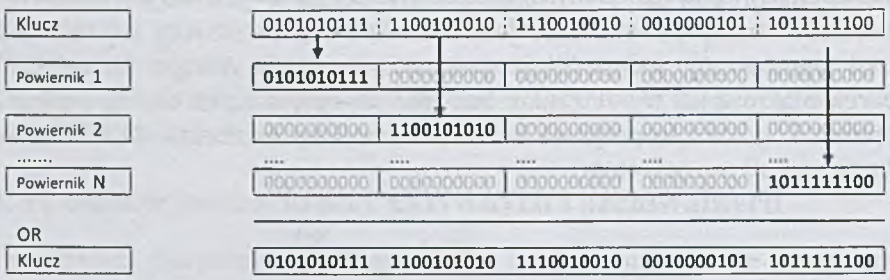
Lagrange'a można obliczyć sekretną liczbę s_0. Mniejsza liczba unikalnych fragmentów nie pozwoli poznać sekretu. Ta klasyczna metoda w której każdy z powierników jest równoprawny ma szereg wariantów otrzymywanych przy założeniu, że powiernicy otrzymują zmienną liczbę fragmentów oraz, że fragmenty nie są unikalne. Pozwala to tworzyć bardziej złożone schematy matematyczne i w ślad za nimi definiować schematy organizacyjne zapewniające zmienny poziom bezpieczeństwa zależnie od potrzeb.

3.2. Możliwość modyfikacji schematu SSSS i podobnych

Przedstawiony schemat i podobne wymagają dużej mocy obliczeniowej do odtworzenia sekretu szczególnie przy dużych liczbach progowych. Dodatkowo pojawia się problem reprezentacji w postaci binarnej i zaplanowania odpowiedniej długości liczb, na których prowadzimy obliczenia [7]. Choć schemat ten jest uznany i dobrze znany od niemal trzydziestu lat, to nadal, w tym współcześnie pojawiają się nowe schematy podziału sekretu [7, 8, 9] dopasowane do określonych modeli bezpieczeństwa choć często uproszczone czy zawężone do konkretnych liczb z pary (t,n).

3.3. Opracowany schemat dzielenia sekretu

Opracowano odmienny schemat rozdziału sekretu, a dalej jego rozszerzenie. Schemat ten wymaga niewielkiej mocy obliczeniowej szczególnie jeśli stosuje się go wielokrotnie do sekretów binarnych o podobnym rozmiarze. Niewielka moc obliczeniowa jest konieczna tak podczas rozdziału sekretu na fragmenty jak i podczas odtwarzania. Schemat opiera się na fragmentacji danych sekretu lub z drugiej strony na maskowaniu bitów w operacji bitowej AND. Odtworzenie klucza jest możliwe według schematu (t,n) przy zastosowaniu tylko operacji sumy logicznej bitów OR. Ideę objaśnia rysunek Rys. 1.

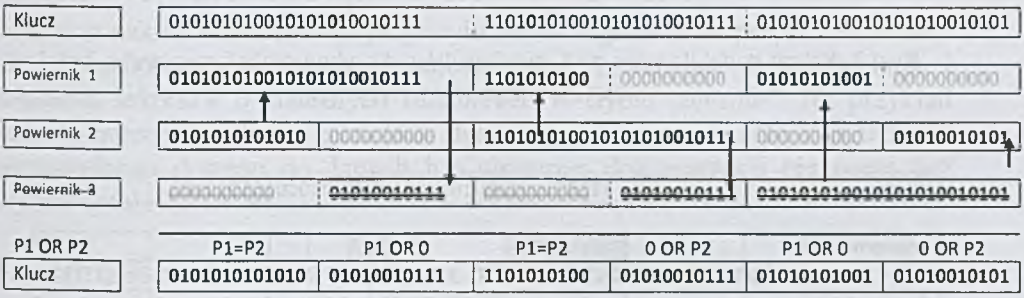


Rys. 1. Sekret podzielony na n=5 powierników do rekonstrukcji trzeba t=n, przypadek trywialny (5,5)

Przedstawiony na rysunku Rys.1. schemat (5,5) i ogólnie każdy w którym t=n są podobne do skrajnego przypadku (1,n), gdzie nie można mówić o podziale

sekretu. Zatem interesujące z praktycznego punktu widzenia są schematy w których liczba progowa t zawiera się w zakresie od 2 do n-1. Szczególnie interesujące są, patrząc na publikowane rozwiązania, te schematy, w których t=3 [7, 8].

Możemy teraz rozwijać ten podstawowy schemat tworząc schemat, w którym eliminujemy jednego z powierników zatem będzie to schemat (n-1,n). Algorytm postępowania przy budowie fragmentów jest następujący. Zaczynamy od podziału (n,n). Każdy z powierzanych fragmentów ma pozostawiony swój fragment. Eliminujemy powiernika i rozdzielamy jego fragment na n-1 fragmentów rozdzielanych pozostałym kolejno. Schemat ten ilustruje rysunek Rys.2 przygotowany przy założeniu, że n=3 i t=2



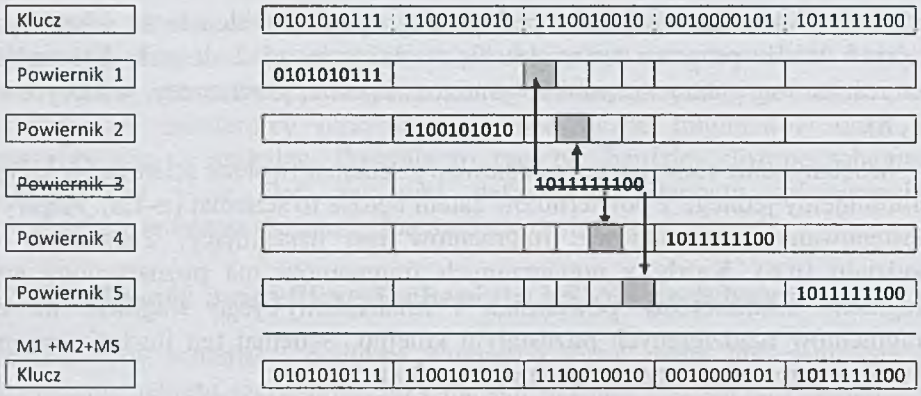
Rys. 2. Sekret podzielony na n=3 powierników do rekonstrukcji wystarcza t=n-1=2, schemat OR/AND (2,3)

Podany algorytm postępowania można powtarzać redukując kolejnego powiernika. Prześledźmy to w przypadku n=5 powierników pragnąc zbudować schemat (3,5)

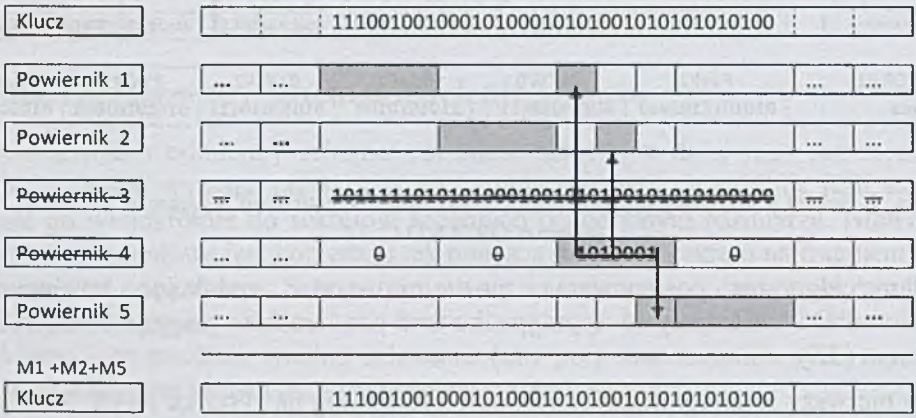
Etap pierwszy to zbudowanie schematu (5,5) według na rysunku Rys.1. Każdy z powierników otrzymuje fragment o długości 1/5 sekretu. Pozostałe bity są wyzerowane.

Etap drugi to zbudowanie schematu (n-1,n) czyli (4,5). Powtarzamy pięć razy algorytm podany na rysunku Rys.2 tym razem zastosowany do czterech pozostałych wykluczając kolejno powierników i rozdzielając ¼ fragmentu na pozostałych czterech, tak jak pokazano na rysunku Rys.3

Etap trzeci to zbudowanie fragmentów do zastosowania w schemacie docelowym (3,5). Postępujemy analogicznie wyłączając powierników parami.



Rys. 3. Sekret podzielony na n=5 powierników do rekonstrukcji wystarcza t=n-1=4, etap 2: eliminacja powiernika 3



Rys. 4. Sekret podzielony na n=5 powierników do rekonstrukcji wystarcza t=n-21=3, etap3: eliminacja powiernika4 po eliminacji powiernika3 w etapie 2

Kolejny i ostatni etap czwarty może doprowadzić do podziału sekretu według schematu (2,5). Procedura którą trzeba przeprowadzić to eliminacja kolejnego powiernika. Patrząc na przykład z rysunku Rys. 4 widać, że ten etap to działanie podobne do przedstawionego na rysunku Rys.2. Sekret przechowywany w trzech równych częściach przez trzech powierników, schemat (3,3), rozdzielić tak, aby możliwe było jego odtworzenie w schemacie (2,3).

Algorytm obliczeniowy konieczny do stworzenia schematu (t,n) jest łatwy do opracowania jako iteracyjny lub rekurencyjny. Sposób konstrukcji fragmentów gwarantuje, że przy zgromadzeniu tylko t-1 lub mniejszej liczby fragmentów odtwarzany sekret pozostanie niekompletny.

3.4. Nakład na obliczenia przy archiwizacji zbioru sekretów

Przedstawiony algorytm obliczeniowy konieczny do stworzenia schematu (t,n) jest łatwy do opracowania jako iteracyjny lub rekurencyjny. Jeśli przyjmiemy, że mamy do czynienia ze zbiorem sekretów bitów o stałej długości, to można zrealizować algorytm obliczania fragmentów przygotowując n liczb binarnych do maskowania bitów w obliczanych fragmentach sekretów. Wspomniane maski można wyznaczyć metodą obliczania fragmentów sekretów tylko jeden raz przyjmując, że sekretem jest liczba złożona z bitów równych jeden. Tak otrzymane maski należy w bitowej operacji logicznej AND nakładać na sekretne słowa binarne ze zbioru sekretów w celu wyznaczenia fragmentów przekazywanych powiernikom.

Ponieważ nakład obliczeniowy przy wyliczeniu fragmentów i przy ich składaniu jest niewielki, to przedstawiona metoda dobrze nadaje się do rozdzielania i składania sekretów o znacznych rozmiarach w trybie „*on-line*". Na przykład danych zapisywanych na zespołach dysków sieciowych w celu zapewnienia niezawodnego dostępu do danych lub obszernej dokumentacji niejawnej bez szyfrowania[10].

3.5. Długość sekretu, a liczby t i n, wydłużanie sekretu

Przyjmując, że mamy do czynienia z sekretem, który można przedstawić jako liczbę binarną o długości L bitów i po zastosowaniu schematu podziału sekretu (t,n), trzeba pamiętać, że warunkiem stosowania przedstawionego schematu jest podzielność liczby L przez n!/t!. Jeśli sekret jest krótszy to należy go wydłużyć do najbliższej takiej liczby większej lub równej iloczynowi n-t+1 kolejnych czynników n*(n-1)*....* (t+1)*t.

3.6. Warunki poznania sekretu w przy spotkaniu t-1 powierników

Z algorytmu konstrukcji fragmentów przez dealera wynika, że spotkanie mniejszej niż progowa liczby powierników nie umożliwi złożenia klucza. Jednak jeśli w polityce bezpieczeństwa zdefiniujemy małą progową liczbę t przy dużej liczbie powierników n, to spotkanie nawet mniejszej niż progowa liczby powierników może pozwolić złamać klucz algorytmem siłowym. Zobaczymy to analizując przypadek „zmowy" mniejszej niż progowa liczby powierników. Nawet bez zmowy, każdy z powierników zna pewną część klucza i poznanie pozostałej części wymaga zmniejszonego nakładu obliczeniowego. Jak wielka jest skala zagrożenia? Pokazano to na przykładzie klucza o długości 1024 bity i różnych wartości t i n w tabeli 2 przy założeniu, że zgromadzenie t-1 powierników postanowiło w wyniku zmowy naruszyć schemat bezpieczeństwa.

Tab. 2.

Liczba bitów do odgadnięcia przy połączeniu t-1 fragmentów ze schematu (t,n), klucz 1024 bity, w nawiasie rozmiar klucza wydłużonego

liczba progowa t	2	3	4	5
2 powierników	512 (1024)	x	x	x
3 powierników	171 (1026)	342 (1026)	x	x
4 powierników	43 (1032)	86 (1032)	256 (1024)	x
5 powierników	9 (1080)	18 (1080)	52 (1040)	205 (1025)

Patrząc na wyniki zgromadzone tabeli 2, obliczone dla klucza o typowej choć niewielkiej drogości widać, gdzie zachodzi duże niebezpieczeństwo złamania klucza. Klucz o projektowanej sile 1024 bitów w wyniku zastosowania schematu podziału na n=5 powierników, z których tylko t=2 wystarcza do odtworzenia klucza, ma taką konstrukcję że każdy z powierników ma zaledwie 9 bitów nieznanych. Jest to wynik z pewnością nie do przyjęcia. W tym samym wierszu liczba 52 oznacza, że przy schemacie (4,5) zmowa t-1, czyli trzech powierników wymaga złamania brakujących 52 bitów klucza.

3.7. Schemat mieszany, powiernicy konieczni i opcjonalni (k,t,n)

Przedstawiony schemat progowy konstrukcji dzielenia sekretu pozwala zbudować model bezpieczeństwa o charakterze hierarchicznym. W takim modelu można podzielić sekret na n powierników, przy czym do odtworzenia sekretu potrzeba wszystkich k spośród powierników grupy koniecznej oraz co najmniej t spośród grupy pozostałych przy podziale na n powierników, czyli schemat (k,t,n). Przy k=0 staje się wcześniej opisywanym klasycznym schematem progowym dzielenia sekretu (t,n). Tego rodzaju model otrzymuje się dzieląc cześć bitów sekretu w schemacie (k,k), a pozostałą cześć w schemacie (t,n-k). .

4. Podsumowanie

Przedstawiono problemy związane z archiwizacją kluczy PKI z punktu widzenia organizacyjnego i prawnego na tle rozwiązań technicznych. Dyskutowano odmienność zagadnienia archiwizacji zależnie od przeznaczenia klucza do szyfrowania lub podpisu. Jeśli już klucze są archiwizowane, to oczywiście

zostają odpowiednio silnie zabezpieczone przed niezamierzonym dostępem przez kolejne szyfrowania administratora archiwum. Zawsze istnieje groźba zdradzenia klucza przez administratora. Do wyobraźni ludzkiej znacznie lepiej przemawia zastosowanie metody podziału sekretu na wielu powierników niż jeden administrator. Do pozyskania hasła potrzeba wtedy zmowy kilku osób. Co więcej dane są bezpieczniejsze bo można zbudować schemat progowy gdzie rozdajemy sekret wielu powiernikom i bez obecności jednego czy kilku z nich nadal odbudowanie sekretu jest możliwe. Wydaje się, że metoda dzielenia sekretu należy do najciekawszych i dobrze osadzonych w tradycji. Przedstawiono nowy, prosty schemat rozdziału sekretu oparty na operacja bitowych, a dalej jego rozszerzenie. Schemat ten wymaga niewielkiej mocy obliczeniowej szczególnie jeśli stosuje się go wielokrotnie do sekretów binarnych o podobnym rozmiarze. Dzięki temu przedstawiona metoda dobrze nadaje się do rozdzielania i odtwarzania sekretów o znacznych rozmiarach w czasie rzeczywistym.

W oparciu o tę metodę przygotowany będzie system tworzenia kopii bezpieczeństwa danych na dyskach sieciowych w systemie wielu stacji sieciowych o małej niezawodności i zmiennym czasie dostępności[15]. Serwer archiwizacji powierzać będzie dane wielu powiernikom przy czym powiernicy przechowując dane nie będą mogli ich poznać.

LITERATURA

1. A. Shamir: How to share a secret, Communications of the ACM 22 (1979), 612-613.
2. G. Blakley. Safeguarding cryptographic keys. In Proc. of AFIPS National Computer Conference, 1979.
3. „Managed Public Key Infrastructure, Securing Your Business Applications", Verisign White Paper, dostępne:http://www.verisign.com.sg/guide/managedpki/whitepaper_managedPKI.pdf.
4. „Enterprise Key Management", Verisign White Paper, dostępne: http://www.verisign.com/static/005308.pdf.
5. „Key Update and the Complete Story on the Need for Two Key Pairs", Entrust 08/2000, dostępne: http://www.entrust.com/resources/pdf/2keypairs11.pdf
6. Asmuth Ch., Bloom J. A Modular Approach to Key Safeguarding. IEEE Transactions on Information Theory, vol. IT-29, No. 2, March 1983.
7. Jun Kurihara, i inni: A Fast (3,n)-Threshold Secret Sharing Scheme Using Exclusive-OR Operations, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences 2008 :str. 127-138.
8. Han-Yu Lin, Yi-Shiung Yeh: A Novel (t,n) Threshold Convertible Authenticated Encryption Scheme , Applied Mathematical Sciences, Vol. 2, 2008, no. 5, str. 249 – 254.
9. Yahya X ALSalqan: Cryptographic Key Recovery, 6th IEEE Workshop on Future Trends of Distributed Computing Systems (FTDCS '97)str. 34-37.

10. A.Brinkmann, S.Effert, F.Meyer: Dynamic and Redundant Data Placement, 27th International Conference on Distributed Computing Systems (ICDCS'07), str. 29-.
11. Eric K. Wang i inni: A Key-Recovery System for Long-term Encrypted Documents, 10th IEEE International Enterprise Distributed Object Computing Conference Workshops (EDOCW'06).
12. Yung-Cheng Lee, Chi-Sung Laih: On the Key Recovery of the Key Escrow System, Proceedings of 13th Annual Computer Security Applications Conference, 1997, str.216-220.
13. D.E. Denning: The US Key Escrow Encryption Technology, Computer Communications, Vol. 17,No.7, July 1994, pp.453-457.
14. D.Stinson; Ruizhong Wei: Bibliography on Secret Sharing Schemes, http://ccc.cs.lakeheadu.ca/bisss.html
15. D. Arendt, R. Krasiukianis, K.Lich: Archiwizacja kluczy szyfrujących i odmiana metody dzielenia sekretu, XVI Konferencja Sieci i Systemy Informatyczne. Teoria, Projekty, Wdrożenia, Łódź 2008.

Rozdział 6

Możliwości wykorzystania CAPTCHA do zabezpieczania stron internetowych

Iwona Iskierka
Politechnika Częstochowska
iwona.iskierka@el.pcz.czest.pl

Streszczenie

W pracy przedstawiono możliwości wykorzystania technologii CAPTCHA, polegającej na generowaniu obrazków z losowymi znakami bądź słowami umieszczonymi na wzorzystym tle, uniemożliwiającym odczytanie ich systemom rozpoznawania pisma (OCR), w procesie zabezpieczania stron internetowych. Podano także przykłady generowania CAPTCHA w technologii PHP. Mimo, że nie ma obecnie w pełni zabezpieczonych przed atakiem kodów tekstowych CAPTCHA, to wciąż jeszcze ten sposób weryfikacji użytkownika jest bardzo popularny.

1. Spam w polskim prawie

Dynamiczne strony internetowe dostarczają dla użytkowników aktualną informację, natomiast dla twórców takich witryn umożliwiają prostą i szybką ich obsługę. Dynamiczna strona internetowa jest najlepszym rozwiązaniem w sytuacji, gdy zawartość witryny musi być często aktualizowana. Statyczne strony WWW przestają satysfakcjonować zarówno odbiorców jak i tworzących witryny, ze względu na zaistniałą potrzebę wytworzenia interakcji pomiędzy użytkownikiem a stroną internetową. W związku z tym coraz bardziej dokuczliwe jest zjawisko spamu blogowego, które związane jest z automatycznym umieszczaniem przez spamerów losowych treści zawierających opisy lub linki do promowanych w ten sposób stron [6]. Istnieje kilka metod walki ze zjawiskiem spamu blogowego. Najpopularniejszymi metodami są: limity komentarzy, blokowanie wpisów zawierających określone słowa, odwrotny test Turinga, usuwanie linków z wpisów, podejścia dystrybuowane, monitorowanie kanałów RSS oraz specjalizowane narzędzia usuwające wpisy.

Dodatkowo używa się opcji przekierowań, by już umieszczone komentarze nie powodowały podwyższenia rankingu strony w wyszukiwarkach. Okazuje się, że najskuteczniejszym narzędziem jest odwrotny test Turinga, polegający na przedstawieniu zadania łatwego do wykonania przez człowieka, ale trudnego do automatyzacji przy użyciu algorytmów komputerowych. Oprócz prostych pytań, często stosuje się zniekształcone obrazki CAPTCHA zawierające tekst, który należy odczytać i wpisać.

Według słownika komputerowego spam to zawierająca niepożądaną treść (np. materiały reklamowe czy agitujące) wiadomość e-mail (lub artykuł) przesłana do wielu adresatów (lub grup dyskusyjnych) jednocześnie [7]. (Spam to także nazwa konserwy mięsnej popularnej w armii Stanów Zjednoczonych podczas drugiej wojny światowej.) W Polsce rozsyłanie spamu jest karalne. Według najnowszego raportu firmy Microsoft, liczba spamu wśród wysyłanych na świecie maili sięgnęła zawrotnych 97%. Raport firmy stwierdza, że tylko trzy na sto maili zawiera istotne informacje i nie zawiera niebezpiecznych załączników, spamu i phisingu. Aż 48% spamu zawiera reklamy produktów farmaceutycznych. Natomiast firma Sophos przedstawiła raport dotyczący pierwszych trzech miesięcy 2009 roku zawierający zestawienie 12 państw odpowiedzialnych za dystrybucję największej ilości spamu. Według wspomnianego raportu największa ilość spamu pochodzi ze Stanów Zjednoczonych - ponad 15% ogólnej liczby globalnego spamu [8]. Na drugim miejscu, notując jednocześnie największy wzrost, znalazła się Brazylia. W ciągu roku liczba niezamawianej korespondencji pochodzącej z tego kraju zwiększyła się z poziomu 4,3% do aż 10,2%. Wielka Brytania oraz Niemcy zredukowały w ostatnim czasie liczbę rozsyłanego spamu. Natomiast Polska jest na 10 pozycji z wynikiem 2,6% liczby globalnego spamu. Drastyczne przykłady spamu na stronach WWW to efekt działania automatów spamujących, ale winę ponoszą też administratorzy witryn. Na rysunku 1 przedstawiono etapy działania robotów spamujących.

W polskim prawie nie ma definicji spamu. Ustawa z dnia 18 lipca 2002 roku o świadczeniu usług drogą elektroniczną wprowadziła pojęcie informacji handlowej, stanowiąc, że jest to każda informacja przeznaczona do promowania towarów, usług lub wizerunku przedsiębiorcy [10]. Artykuł 10 tej ustawy zakazuje przesyłania niezamówionej informacji handlowej za pomocą środków komunikacji elektronicznej, w szczególności poczty elektronicznej. Informację handlową uważa się za zamówioną, jeżeli odbiorca wyraził zgodę na otrzymywanie takiej informacji, w szczególności zaś udostępnił w tym celu identyfikujący go adres elektroniczny. Według informacji Ministerstwa Spraw Wewnętrznych i Administracji, w ciągu ostatnich trzech lat zostało wszczęte tylko kilkadziesiąt postępowań przeciwko spamerom. Wydaje się więc, że obecnie obowiązujące regulacje prawne dotyczące spamu nie są wystarczające. W związku z zaistniałą sytuacją planowana jest nowelizacja ustawy Prawo telekomunikacyjne. W planowanej nowelizacji wprowadzona zostanie definicja spamu i uregulowane sprawy związane z odpowiedzialnością za spam.

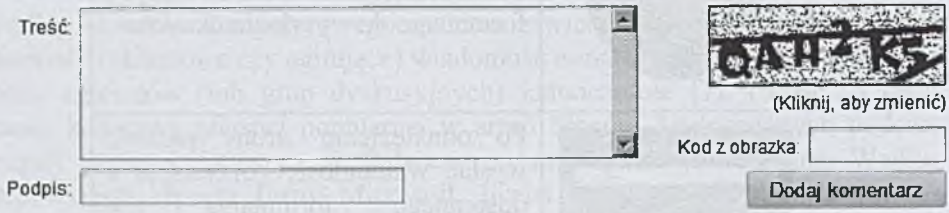
Odpowiedzialny za spam będzie nie tylko podmiot, który go rozsyła, ale również ten, który zleca jego przesyłanie lub odnosi z tego korzyści. Planuje się również kary pieniężne za przesyłanie spamu.

	Robot przegląda strony internetowe, poszukując znanych typów forów dyskusyjnych lub dokumentów zawierających formularze do wysyłania danych.
	Po odnalezieniu strony próbuje wysłać wiadomość, korzystając z dostępnego formularza lub znanego błędu w skrypcie strony.
	Potrafi także rejestrować użytkowników fałszywych i umieszczać w ich profilu niebezpieczne odnośniki. Nowe konto może być użyte do wysyłania wiadomości.
	Robot przystępuje do masowego wysyłania spamu. Ręczne usuwanie takich wiadomości nie zawsze skutkuje, bo będzie ponawiał próby. Konieczna jest zmiana zabezpieczeń strony internetowej.
	Po wykonaniu pracy robot dodaje do listy inne strony w tej samej domenie lub połączone odnośnikami z zaatakowaną stroną. Następny etap to próba zaspamowania kolejnej strony WWW.

Rys 1. Etapy działania robotów spamujących - opracowanie własne na podstawie [9]

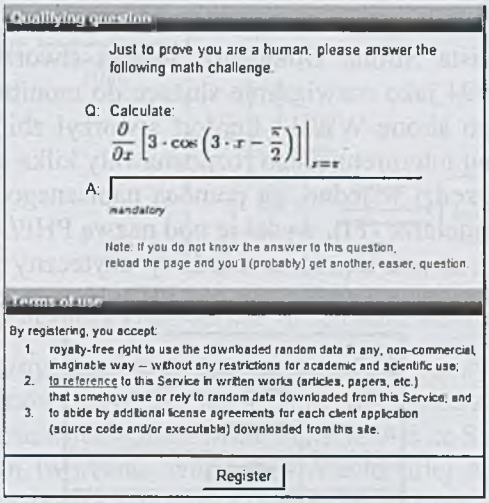
2. CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)

Nazwę CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) zaproponowali po raz pierwszy Luis von Ahn, Manuel Blum, Nicholas J. Hopper z Carnegie Mellon University [11].



Rys. 2. Przykład kodu CAPTCHA - opracowanie własne na podstawie [12].

CAPTCHA to kod weryfikacji człowieczeństwa. Z założenia CAPTCHA stanowić ma barierę dla botów, automatów spamujących i oprogramowania czytającego i łamiącego kody - OCR (ang. Optical Character Recognition). Kody typu CAPTCHA występują w formularzach rejestracyjnych, przy zakładaniu kont pocztowych, czy dodawaniu komentarzy do artykułu lub postów do bloga. Aby odróżnić prawdziwych internautów od programów spamujących, administratorzy stosują różne sposoby - najpopularniejszy z nich polega na przepisaniu tekstu z obrazka. Kod ma zwykle postać liter lub cyfr, które należy przepisać z obrazka. W związku z różnymi opiniami użytkowników dotyczącymi użyteczności kodów CAPTCHA, Agencja e-biznes Symetria [13] przeprowadziła badania dotyczące stosowania kodów CAPTCHA. Przedstawione wnioski z badania pozwalają uporządkować pojęcia w zakresie budowy użytecznych, przyjaznych dla użytkownika kodów CAPTCHA. Użyteczności kodów CAPTCHA została poświęcona praca autorstwa Jeffa Yan, Ahmada Ahmad [5]. Główne problemy użyteczności CAPTCHA to między innymi: niska rozpoznawalność przez użytkownika, skomplikowanie samego kodu - zadania matematyczne, itp., problemy z odczytaniem kodu CAPTCHA przez osoby niepełnosprawne: niewidzące, niedowidzące, niesłyszące, brak przycisku odśwież kod. Podstawowym aspektem bezpieczeństwa CAPTCHA jest celowe zaburzanie elementów kodu – distortion, w celu uniemożliwienia jego odczytania przez oprogramowanie OCR. Do takich zaburzeń zalicza się: przemieszczanie elementów w górę i w dół, obroty elementów, skalowanie elementów i gięcie elementów. Pojawiają się także pomysły na nietypowe CAPTCHA [12]. Korzystając z aktualnie niedostępnej usługi hotcaptcha.com należało z dziewięciu dostarczonych zdjęć osób wybrać osoby, które wyglądają na atrakcyjne. System hotcaptcha.com został w szybkim czasie złamany. System, w którym trzeba podać wynik wygenerowanego zadania matematycznego to następny przykład na nietypowe CAPTCHA. System taki wdrożono np. na stronie [14].



Rys 3. Przykład nietypowego kodu CAPTCHA [14]

3. Technologia PHP (Personal Home Page HyperText Preprocesor)

Dynamiczne strony internetowe mają bardzo duże możliwości ze względu na swoją elastyczność. Tworzone są na żądanie użytkownika w danej chwili. Według L. Ullman [4] mogą reagować na zmieniające się parametry (np. na porę dnia lub wersję przeglądarki internetowej), często wyposażone są w interfejs, za pośrednictwem którego administrator może zarządzać zawartością witryny. Istotne jest, że dysponują „pamięcią", dzięki czemu użytkownicy mogą rejestrować się w systemie, logować się do niego i np. dokonywać zakupów. Dynamiczne strony internetowe są dużo łatwiejsze do utrzymania, uaktualniania i rozwijania.

Techniki używane do tworzenia stron internetowych mogą być stosowane zarówno po stronie serwera (technologie Server-side), jak i po stronie klienta (client-side). Znacznie większe możliwości związane z projektowaniem stron dynamicznych dają programy i skrypty działające po stronie serwera. W tej sytuacji serwer odpowiada za wygenerowanie witryny HTML w zależności od potrzeb klienta i wysłanie jej do przeglądarki. Zawartość wygenerowanej witryny zależy więc od wyników działania skryptów. Standardem, który definiuje metodę porozumiewania się serwera WWW ze skryptami, jest Common Gateway Interface (CGI). Skrypty CGI można pisać w językach wysokiego poziomu (np.Perl, C++ czy BASIC), a także w językach osadzanych w treści pliku HTML (PHP, ASP) [1]. Największą zaletą programów działających po stronie serwera jest to, że mogą one współpracować z innymi aplikacjami uruchomionymi na serwerze, przede wszystkim z bazami danych.

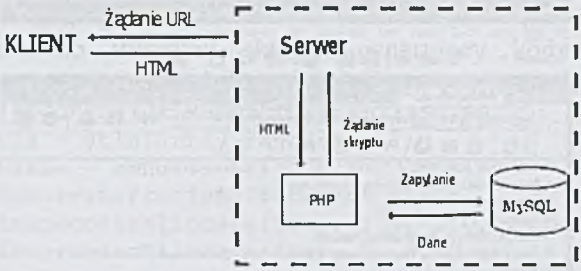
PHP (Personal Home Page HyperText Preprocesor) jest najszybciej rozwijającą się technologią tworzenia dynamicznych stron internetowych. Pierwsza wersja

PHP, rozpowszechniana pod nazwą PHP/FI (Personal Home Page/Forms Interpreter – Osobista Strona Domowa), została stworzona przez Rasmusa Lerdorfa w roku 1994 jako rozwiązanie służące do monitorowania liczby osób odwiedzających jego stronę WWW. Lerdorf stworzył zbiór narzędzi, które za pomocą mechanizmu interpretującego rozpoznawały kilka specjalnych makr. Po połączeniu tych narzędzi w jedno, za pomocą napisanego przez siebie pakietu interpretującego formularze (FI), wydał je pod nazwą PHP/FI [3].
W miarę jak PHP stawał się coraz bardziej użyteczny i powstawały nowe funkcje, jego oficjalną nazwę zmieniono na „PHP: Hypertext Preprocessor".

Rok	
1994	PHP/FI
1998	PHP 3.0
2000	PHP 4
↕	PHP 4.1, 4.2 oraz 4.3
2004	PHP 5

Rys. 4. Historia rozwoju języka PHP. Opracowanie własne na podstawie [4]

Zgodnie z definicją podaną na oficjalnej stronie internetowej tego języka, PHP jest to „język skryptowy osadzany w kodzie HTML". Oznacza to, że wyrażenia języka PHP mogą przeplatać się ze znacznikami HTML-a, co bardzo ułatwia tworzenie dynamicznych stron internetowych. Należy również pamiętać, że PHP jest językiem skryptowym, przetwarzanym i wykonywanym po stronie serwera, a nie typowym językiem programowania. Reaguje on jedynie na pewne zdarzenia, takie jak zatwierdzenie danych w formularzu lub przejście pod określony adres URL [4].
PHP jest przenośną technologią działającą po stronie serwera. Wszystkie operacje zdefiniowane w skryptach tego języka są wykonywane na serwerze. PHP może pracować na większości dostępnych systemów operacyjnych: Windows, Unix (w wielu odmianach) oraz Macintosh. Strony PHP składają się z trzech składników: tekstu, kodu HTML i skryptów PHP. Strony zawierające skrypty PHP mają inną strukturę niż te, które zawierają jedynie HTML i w celu poinformowania o tym analizatora PHP są zapisane na serwerze WWW w plikach z rozszerzeniem .php i na tej podstawie są rozpoznawane i wykonywane przez PHP zainstalowany na serwerze. Wyniki działania skryptu są przesyłane do przeglądarki w postaci HTML.
PHP umożliwia także współpracę z wieloma systemami relacyjnych baz danych (np. MySQL, Oracle, PostgreSQL, SQLite) oraz korzystanie z alternatywnych sposobów przechowywania danych - plików tekstowych i XML-owych. Należy zaznaczyć, że PHP jest produktem Open Source.

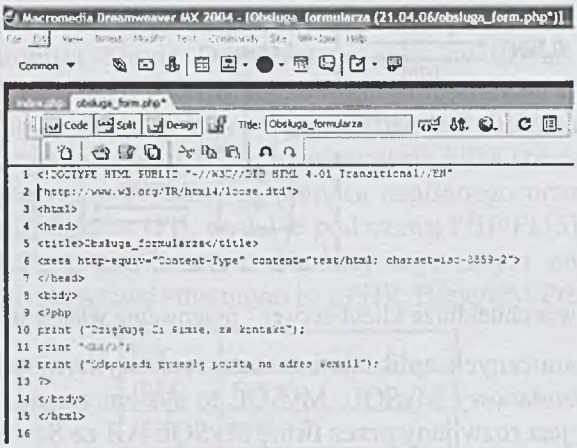


Rys. 5. PHP w architekturze klient-serwer - pracowanie własne na podstawie [4]

Większość dynamicznych aplikacji internetowych wykorzystuje zarówno PHP jak i system bazodanowy MySQL. MySQL to system zarządzania relacyjnymi bazami danych, jest rozwijany przez firmę MySQL AB ze Szwecji.

Podstawowym celem tworzenia witryny WWW opartej na bazie danych jest przechowywanie zawartości witryny w bazie danych tak, aby można ją było stamtąd pobierać dynamicznie, generując treści stron w czasie rzeczywistym za pomocą skryptów PHP.

Język PHP zyskał popularność dzięki prostej składni i olbrzymim możliwościom. Oferuje wysoką wydajność, stabilność i ścisłą integrację z systemami baz danych dostępnymi na rynku, zarówno komercyjnymi (np. Oracle, Sybase) jak i bezpłatnymi (MySQL, PostgreSQL). Wśród najczęstszych zastosowań języka PHP wymienia się: wykonywanie funkcji systemu: tworzenie, otwieranie, czytanie z, zapisywanie do i zamykanie plików w systemie; wykonywanie poleceń systemowych; tworzenie katalogów i modyfikowanie zezwoleń dostępu. Innym przykładem jest zbieranie danych z formularzy: zapisywanie danych do pliku, wysyłanie danych przez e-mail, zwracanie danych przetworzonych użytkownikowi, uzyskiwanie dostępu do baz danych i generowanie zawartości w czasie rzeczywistym lub tworzenie interfejsu WWW służącego do dodawania i modyfikowania elementów w bazie danych użytkownika oraz ich usuwania [3]. Inne zastosowania to tworzenie cookies, uzyskiwanie dostępu do ich zmiennych, rozpoczynanie sesji i używanie zmiennych i obiektów sesji, kodowanie danych, tworzenie obrazów na bieżąco. Bardzo ciekawą rzeczą jest dynamiczne tworzenie ilustracji przez skrypt działający na serwerze. Rysunki tworzone po stronie serwera mogą być sposobem prezentacji różnego rodzaju danych dynamicznych na stronie internetowej.



Rys. 6. Przykład skryptu PHP - wykorzystanie programu Dreamweaver

Aby wykorzystać możliwości tworzenia dynamicznych stron internetowych w technologii Personal Home Page, należy mieć dostęp do serwera WWW z obsługą skryptów PHP oraz do baz danych. Duże możliwości daje program Macromedia Dreamweaver jako aplikacja typu WYSIWYG obsługująca PHP. Można więc założyć konto (darmowe lub płatne) na serwerze w Internecie, udostępniającym powyższe usługi lub też zainstalować oprogramowanie na własnym komputerze. W takiej sytuacji warto wykorzystać pakiety, takie jak XAMPP, a także Krasnal, które zawierają zestaw wielu komponentów, począwszy od serwera Apache z modułami PHP i Perl, poprzez bazę danych MySQL aż po serwery FTP i SMTP [2]. Wspomniane pakiety zawierają również popularne narzędzie typu open source o nazwie phpMyAdmin – pozwala ono komunikować się z serwerem MySQL za pośrednictwem przeglądarki internetowej oraz program WebAlizer służący do analizy logów serwera i tworzenia na jej podstawie statystyk. Należy zaznaczyć, że pakiet phpMyAdmin jest doskonałym rozwiązaniem umożliwiającym administrowanie bazą MySQL.

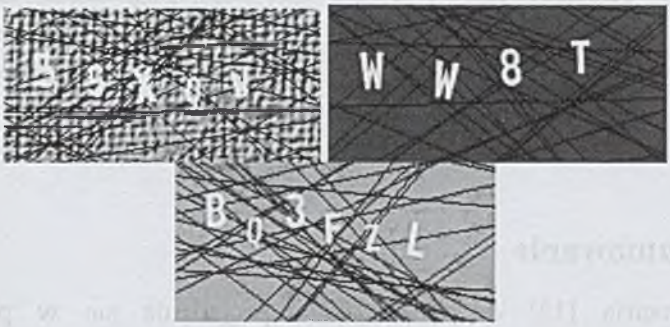
4. Generowanie CAPTCHA w technologii PHP

Wykorzystując technologię PHP można generować kody CAPTCHA przy użyciu obrazów na wiele sposobów. Jednym z nich jest wykorzystanie języka PHP z zainstalowaną biblioteką GD oraz użycie MySQL. Baza danych MySQL służy do przechowywania haseł, które umieszcza się na obrazkach. W takim przypadku wykorzystuje się skrypty pozwalające na utworzenie tabeli w bazie i wygenerowanie haseł do obrazków, skrypty generujące odpowiednie obrazki i skrypty zawierające formularz z obrazkiem zabezpieczającym. Można również generować CAPTCHA bez korzystania z bazy danych. W tym przypadku obrazek tworzony jest z wykorzystaniem odpowiedniej funkcji z pliku (np. JPEG), którego ścieżkę użytkownik podaje jako parametr. W takiej sytuacji można użyć następujących funkcji *imagecolorallocate()* ,*imageline()*,

imagettftext(), imagejpeg().

Poniżej umieszczono fragment skryptu generujący kody CAPTCHA, z wykorzystaniem przygotowanych wcześniej plików graficznych.

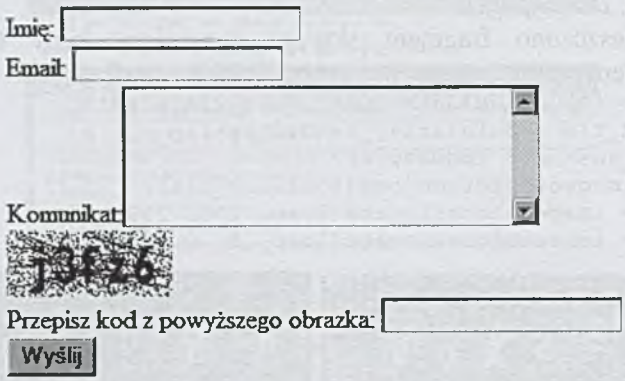
```
$znaki = 'ABCDEFGHIJKLMNPQRSTUWXYZ123456789';
$obrazek_tla = $tla[array_rand($tla)];
$liczba_znakow = rand(4, 6);
$cap = imagecreatefromjpeg($obrazek_tla);
$kolor = imagecolorallocate($cap, 250, 250, 250);
$linie = imagecolorallocate($cap, 0, 0, 0);
```



Rys. 7. Przykład wygenerowanych obrazków - opracowanie własne na podstawie [15]

W przypadku wykorzystania formularza, można utworzyć dwa pliki *.php. Jeden z plików będzie służył do generowania obrazków, drugi będzie zawierał formularz z obrazami zabezpieczającymi. W tej sytuacji grafika generowana jest bezpośrednio w pliku. Poniżej umieszczono fragment skryptu generujący kody CAPTCHA, w przypadku generowania grafiki w pliku php.

```
function
CaptchaSecurityImages($width='120',$height='40',$characters='6
') {
$code = $this->generateCode($characters);
$font_size = $height * 0.75;
$image = @imagecreate($width, $height) or die('Cannot
initialize new GD image stream');
$background_color = imagecolorallocate($image, 255, 255, 255);
$text_color = imagecolorallocate($image, 0, 0, 0);
$noise_color = imagecolorallocate($image, 250, 120, 80);
for( $i=0; $i<($width*$height)/3; $i++ ) {
imagefilledellipse($image, mt_rand(0,$width),
mt_rand(0,$height), 1, 1, $noise_color);}
}
```



Rys. 8. Efekt działania skryptów php - opracowanie własne na podstawie [16]

5. Podsumowanie

Firma Symetria [13] od 1998 roku specjalizuje się w planowaniu i kompleksowej obsłudze projektów związanych z szeroko pojętym e-commerce. Za pomocą metody eye trackingowej przeprowadzono w Symetrii mały test użyteczności kodów weryfikacyjnych CAPTCHA. Na podstawie wyników badań sformułowano cechy dobrego CAPTCHA. Należy o nich pamiętać również w procesie projektowania CAPTCHA w technologii php [11].
Dobre CAPTCHA to takie, które zapewnia bezpieczeństwo serwisu i nie są jednocześnie utrapieniem dla użytkowników, powinny:

- zapewniać dostęp do serwisu użytkownikom niepełnosprawnym,
- zawierać przycisk odśwież,
- informować o ilości koniecznych do wpisania elementów,
- dawać możliwość dodatkowej np. telefonicznej weryfikacji użytkownika
- podlegać czasowym ewaluacjom bezpieczeństwa.

LITERATURA

1. Busz A., Góral A: Witryny na zamówienie. Pięć minut do własnego serwera. Chip Professional nr 4, 2005
2. Kierzkowski A.: PHP 5 Tworzenie stron WWW. Wydawnictwo Helion, Gliwice, 2004
3. Meloni J. C.: PHP Podręcznik tworzenia stron WWW. Mikom,Warszawa 2001
4. Ullman L.: Dynamiczne strony WWW PHP i MySQL. Helion Gliwice 2004
5. http://cups.cs.cmu.edu/soups/2008/proceedings/p44Yan.pdf
6. http://bezpieczenstwo.idg.pl/artykuly/60573/Jak.wyciszyc.blogowy.spam.html
7. http://www.komputerswiat.pl/slownik-komputerowy/s/spam.aspx
8. http://www.pcadvisor.co.uk/news/index.cfm?RSS&NewsID=115122

9. http://bezpieczenstwo.idg.pl/artykuly/56527/Zasypani.przez.spam.html
10. http://www.komputerswiat.pl/blogi/okiem-prawnika/2008/10/spam-w-polskim-prawie.aspx
11. http://www.internetmaker.pl/artykul/5271,1,czy_aby_na_pewno_jest_pan_czlowiekiem_uzytecznosc_kodow_captcha_-_teoria.html
12. http://www.internetmaker.pl/artykul/5255,1,ochrona_przed_spamem_captcha.html
13. http://www.symetria.pl/html/i_44_ogolem.html
14. http://random.irb.hr/signup.php
15. http://wortal.php.pl/wortal/
16. http://www.white-hat-web-design.co.uk/articles/php-captcha.php

Rozdział 7

Brama VPN jako narzędzie zabezpieczenia danych przesyłanych przez Internet

Grzegorz Węgrzyn
Wyższa Szkoła Ekonomii i Informatyki w Krakowie, Zakład Informatyki
Grzegorz.Wegrzyn@wsei.edu.pl

Mariusz Szymczyk
Wyższa Szkoła Ekonomii i Informatyki w Krakowie, Zakład Informatyki
Mariusz.Szymczyk@wsei.edu.pl

Krzysztof Molenda
Wyższa Szkoła Ekonomii i Informatyki w Krakowie, Zakład Informatyki
Krzysztof.Molenda@wsei.edu.pl

Streszczenie

W pracy przedstawiono koncepcję wykorzystania technologii VPN (wirtualne sieci prywatne) do uruchomienia usługi internetowej bramy VPN zapewniającej bezpieczeństwo połączenia internetowego z niezaufanej sieci lokalnej oraz transmisji danych w Internecie. Do realizacji koncepcji wybrano oprogramowanie OpenVPN. Przedstawiono szczegółową konfigurację systemu bramy oraz oprogramowania OpenVPN, przeprowadzono testy wydajnościowe algorytmów szyfrujących, oszacowano przepustowość połączeń i obciążenia serwerów. Omówiono możliwość wdrożenia dodatkowych usług w oparciu proponowane rozwiązanie – systemu zapobiegania włamaniom, zdalnego systemu plików, wyspecjalizowanych systemów proxy zapewniających anonimizację i swobodę dostępu do Internetu.

1. Wprowadzenie

W dobie bardzo szybkiego rozwoju globalnej sieci, coraz większym problemem stają się zagrożenia dotyczące przesyłanych danych internetowych. Zjawiska takie, jak kradzież danych wrażliwych (login/hasło) przez „podsłuchiwanie"

transmisji, zmiana zawartości pakietów sieciowych (np. w celu skierowania przeglądarki www na strony atakujące system użytkownika, podmiana adresów URL podczas odwiedzin zaufanego serwisu www) są coraz powszechniejsze. Ataki tego rodzaju są przeprowadzane głównie przez wysoce wyspecjalizowane złośliwe oprogramowanie i odbywają się w pełni automatycznie (wystarczy jeden „zarażony" system w sieci lokalnej składającej się z setek komputerów). Obecnie używane technologie sieciowe (stos protokołów TCP/IP – warstwy: łącza danych, sieciowa i transportowa), nie zapewniają żadnych mechanizmów bezpieczeństwa i pozwalają na podsłuchiwanie lub zmianę danych sieciowych przesyłanych pomiędzy komputerami w sieci lokalnej. Niebezpieczne staje się przesyłanie jakichkolwiek danych niezabezpieczonych przez warstwy wyższe (np. HTTPS).

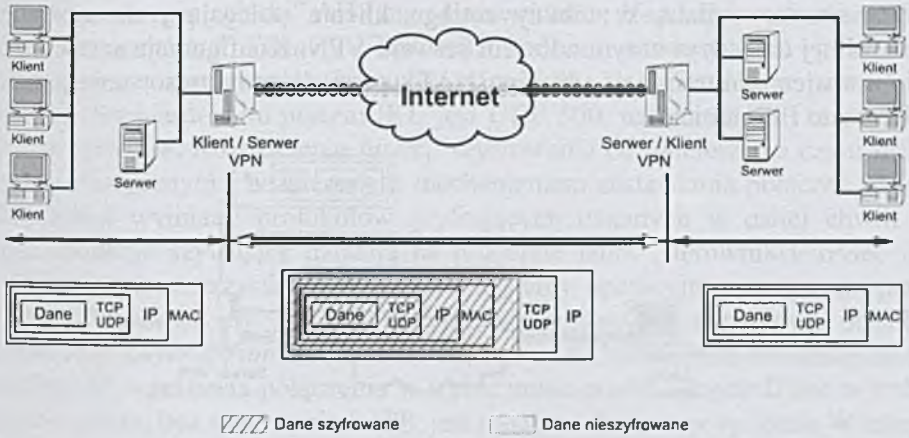
O ile prawie wszyscy użytkownicy systemów komputerowych wiedzą o zabezpieczeniu swojej stacji roboczej przed wirusami (np. trojan, spyware), to większość z nich nie uświadamia sobie, że połączenie jest bezpieczne dopiero po przejściu przez bramę (gateway) naszego dostawcy Internetu. Problem nie występuje, gdy korzystamy z bezpiecznej sieci lokalnej (np. sieci firmowej), do której nie mają dostępu osoby nieupoważnione, pojawia się jednak w chwili włączenia prywatnego komputera do sieci globalnej z wykorzystaniem niepewnej sieci osiedlowej czy ogólnodostępnego hot spotu. W tej sytuacji należy zadbać o zabezpieczenie realizowanego połączenia sieciowego przed ewentualną ingerencją innych użytkowników sieci lokalnej.

2. Sieci VPN

Korzystając ze standardowych połączeń internetowych, których podstawą jest protokół internetowy IP w wersji 4 oraz protokoły warstwy TCP (TCP i UDP), nie ma możliwości zapewnienia poufności przesyłanych danych. Częściowym rozwiązaniem tego problemu jest stosowanie zabezpieczeń w wyższych warstwach stosu protokołów (np. SSL w warstwie aplikacji). W celu zabezpieczenia danych na poziomie sieci można zastosować technologię VPN (wirtualne sieci prywatne). Technologia ta opiera się na zastosowaniu tunelowania i szyfrowania połączeń sieciowych. Enkapsulacja umożliwia zaszyfrowanie całych pakietów przesyłanych jako ładunek w pakietach standardowego połączenia internetowego. Wirtualne łącze zachowuje się wtedy jak połączenie bezpośrednie pomiędzy dwoma klientami końcowymi (tak, jak fizyczne łącze dedykowane). W połączeniu VPN, na dystansie pomiędzy punktami krańcowymi, dane zawarte w pakietach stają się niedostępne dla żadnych pośrednich węzłów zapewniających połączenie fizyczne. Danych nie można podsłuchać, a tym bardziej zmodyfikować.

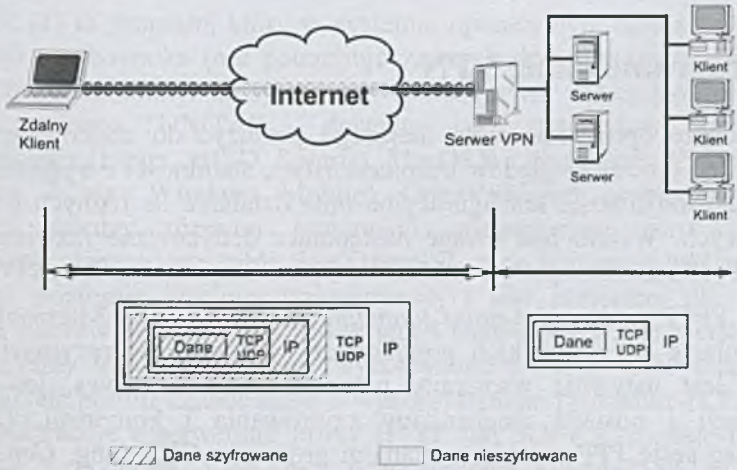
Sieci VPN możemy podzielić na typy w zależności od zastosowanej topologii wirtualnego połączenia [5]:

a) sieć typu *site-site* (rys. 1), w której zestawione jest połączenie między dwoma odległymi sieciami lokalnymi, np. łączące ze sobą 2 rutery brzegowe tych sieci. Tylko pomiędzy ruterami dane są transportowane szyfrowanym tunelem, zaś w obu sieciach lokalnych są propagowane w postaci niezaszyfrowanej. W takiej konfiguracji można uzyskać funkcjonalność jednej spójnej sieci lokalnej wraz z jej wszystkimi podstawowymi cechami (np. komunikacja rozgłoszeniowa).



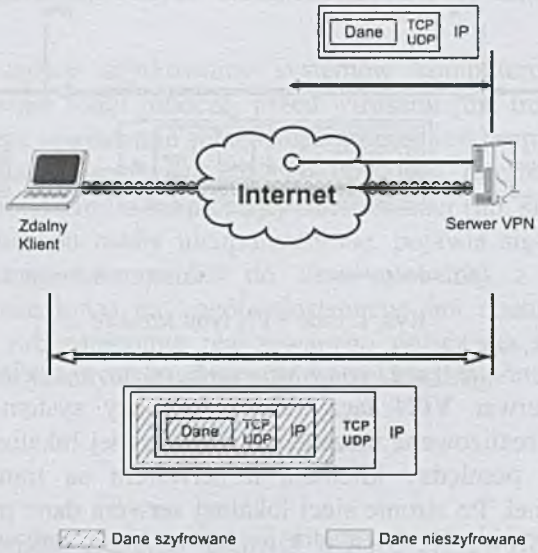
Rys. 1. Sieć VPN typu *site-site*

b) sieć typu *client-site* (ang. *road warrior*), gdzie z siecią lokalną dostępną zdalnie przez serwer VPN łączy się pojedynczy system klienta (rys. 2). Podłączenie jest realizowane niezależnie od fizycznej lokalizacji klienta (klient mobilny). Dane pomiędzy klientem a serwerem są transportowane przez zaszyfrowany tunel. Po stronie sieci lokalnej serwera dane przekazywane są w postaci niezaszyfrowanej. System kliencki uzyskuje bezpośredni dostęp do systemów i zasobów w sieci lokalnej tak, jakby był podłączony bezpośrednio.



Rys. 2. Sieć VPN typu *client-site*

c) sieć typu „brama VPN" (rys. 3), w której klient łączy się z serwerem VPN i traktuje go jako domyślą bramę internetową. W tej konfiguracji serwer VPN kontroluje wszystkie połączenia klienta z internetem. Dane na dystansie klient-brama przekazywane są w zaszyfrowanym tunelu, natomiast na dystansie brama-internet nie są już tunelowane. Funkcjonalnie jest to połączenie typu *client-site* ze zmianą w tablicy rutingu klienta polegającą na zastąpieniu domyślnej trasy prywatnym adresem serwera VPN. Konfiguracja serwera różni się ustawieniami translacji adresów (NAT), pozwalając na trasowanie pakietów klienta do i z internetu.



Rys. 3. Sieć VPN typu „brama VPN"

3. Oprogramowanie VPN

Przy wyborze oprogramowania mającego posłużyć do zbudowania systemu bramy VPN, oprócz względów bezpieczeństwa, stabilności i wydajności, brano pod uwagę możliwości konfiguracyjne oraz działanie na różnych platformach systemowych. Wzięto pod uwagę następujące dedykowane rozwiązania VPN [3]: PPTP, IPSec oraz oparty o bibliotekę OpenSSL program OpenVPN.

Protokół PPTP (ang. *point-point tunneling protocol*) firmy Microsoft, posiada implementacje we wszystkich popularnych systemach operacyjnych (klient i serwer). Jest natywnie wspierany przez systemy Windows, jest łatwy w konfiguracji i posiada mechanizmy szyfrowania i kompresji. Działa on, przesyłając sesje PPP z wykorzystaniem protokołu GRE (ang. *Generic Route Encapsulation*) warstwy IP [2], zaś protokołu TCP (port 1723) używa jako mechanizmu kontroli sesji. Protokół PPTP uważany jest za rozwiązanie mało

bezpieczne. Nie wszystkie systemy firewall i rutery NAT go wspierają. Do poprawnej jego współpracy z mechanizmem NAT wymagane jest śledzenie sesji protokołu GRE. PPTP potrzebuje wsparcia systemu operacyjnego na poziomie jądra (odpowiednie sterowniki ppp).

IPSec jest zbiorem protokołów, w którego skład wchodzą: protokół wymiany kluczy IKE (ang. *Internet Key Exchange*) oraz protokoły zestawiania kanału komunikacyjnego: „płaski" AH (ang. *Authentication Header*) i szyfrowany ESP (ang. *Encapsulated Security Payload*). Podczas nawiązywania połączenia zestawiane są dwa kanały komunikacyjne (dla obu kierunków przesyłania danych). Standardowym portem IKE jest UDP 500, natomiast ESP to protokół 50 z warstwy IP. Rozdzielenie funkcji szyfrowania od tunelowania czyni IPSec bardzo elastycznym i bezpiecznym mechanizmem zestawiania połączeń, dzięki możliwości wymiany protokołów szyfrujących uznanych w danej chwili za słabe. Funkcje szyfrujące działają na poziomie jądra (sterowniki). IPSec jest wspierany przez wszystkie nowoczesne systemy operacyjne – natywnie przez systemy Windows. W systemach Windows stosuje się dodatkowo protokół L2TP (ang. *Layer 2 Tunneling Protocol*), który jest kontenerem dla datagramów warstwy IP – zestawia połączenia w trybie punkt-punkt. Użycie IPSec w trybie bezpośrednim, bez stosowania L2TP, jest możliwe dopiero w systemie Windows 7. Wadą protokołu IPSec jest duży stopień komplikacji, co pociąga za sobą trudną konfigurację i administrację. Pierwotnie IPSec nie był przystosowany do współpracy z systemami NAT (jest mechanizmem przeniesionym wstecznie z IPv6 do IPv4) i do tej pory nie wszystkie rutery NAT pozwalają na jego stosowanie. Wymagane jest śledzenie sesji protokołu ESP przez ruter NAT. Stosowany jest także dodatkowych mechanizm NAT-T (ang. *NAT traversal* – wymaga dodatkowego portu UDP 4500) wspomagający zestawianie połączeń IPSec poprzez ruter NAT.

OpenVPN [1] to program, który w systemie operacyjnym działa całkowicie w przestrzeni użytkownika (nie potrzebuje żadnych dedykowanych sterowników jądra). Do komunikacji z systemem operacyjnym wykorzystuje wirtualny interfejs sieciowy TUN/TAP – dostępny jako standardowy sterownik w systemach unix (Linux, xBSD, Solaris), MacOSX i systemach Windows (2000, XP, Vista, 7 oraz Windows Mobile). OpenVPN bez problemów realizuje połączenia między różnymi systemami operacyjnymi. Do szyfrowania wykorzystuje standardową bibliotekę OpenSSL, a do kompresji bibliotekę LZO. Działa na poziomie warstwy transportowej i jest przyjazny dla systemów firewall i ruterów NAT (wykorzystuje jeden standardowy port warstwy TCP). Jako domyślny transport sieciowy wykorzystuje protokół UDP (standardowo port 1194), ale potrafi działać także z wykorzystaniem protokołu TCP. W trybie TCP współpracuje z serwerami proxy HTTP lub SOCKS, co daje możliwość połączenia z serwerem VPN bez bezpośredniej komunikacji w warstwie sieciowej (konfiguracja sieci bez domyślnej bramy – połączenia realizowane

przez proxy). Potrafi wykorzystywać interfejs TUN/TAP w dwóch trybach – w trybie połączeń punkt-punkt i w trybie rozgłoszeniowym. Wadą OpenVPN jest poleganie na zewnętrznych bibliotekach. Błędy występujące w oprogramowaniu zewnętrznym mogą wpłynąć negatywnie na bezpieczeństwo całego rozwiązania. W samym programie OpenVPN nie wykryto żadnych błędów od roku 2006.

4. Certyfikaty SSL w OpenVPN

Oprogramowanie OpenVPN potrafi tworzyć połączenia szyfrowane SSL w oparciu o dwie metody autoryzacji dostępu [4]:

a) klucza współdzielonego (PSK) – mechanizmu, w którym strony połączenia używają wspólnego stałego klucza służącego do uwierzytelniania i szyfrowania komunikacji,

b) certyfikatów SSL (infrastruktura klucza publicznego, PKI) – mechanizmu autoryzacji pozwalającego na wykorzystanie wszystkich zalet PKI: przyznawania certyfikatów na określony czas, odwoływania skompromitowanych certyfikatów przed upływem daty ich ważności oraz użycia certyfikatu w celach konfiguracyjnych (każdy klient otrzymuje własny certyfikat z unikalnym polem CN – *Common Name*).

Oprogramowanie OpenVPN posiada wygodne środowisko do generowania i zarządzania klienckimi certyfikatami SSL. Jest to zestaw skryptów, dzięki którym administrator może wygenerować certyfikat autoryzujący/nadrzędny (CA), certyfikat serwera, certyfikaty klientów oraz edytować listę CRL (ang. *Certificate Revocation List*) – listę certyfikatów odwołanych.

Serwer OpenVPN umożliwia używanie certyfikatu w celach konfiguracyjnych, odczytując pole CN certyfikatu i dopasowując do predefiniowanych opcji konfiguracyjnych – „per klient". Czyni to dynamiczną konfigurację podczas podłączenia klienta bardzo wygodną. Każdemu klientowi można w ten sposób indywidualne przydzielić stałe opcje i określić jego konfigurację po stronie serwera.

5. Konfiguracja bramy VPN

W implementacji proponowanego rozwiązania system operacyjny serwera to standardowy Linux Debian z rekompilowanym jądrem 2.6.26, zainstalowany w serwerowni firmy OVH (fizyczna lokalizacja to Roubaix – Francja). Serwer posiada 2 rdzeniowy procesor Intel Core Duo 2.33 GHz, 2 GB pamięci RAM i 500GB przestrzeni dyskowej. Serwer podpięty jest do łącza o przepustowości 100 Mbit/s. Najważniejszymi aspektami są tu przepustowość sieci i moc obliczeniowa procesora.

Przy projektowaniu systemu bramy VPN zdecydowano się na użycie połączeń typu punkt-punkt. Najważniejszym aspektem było bezpieczeństwo połączeń klientów. Połączenia klientów są izolowane między sobą. System klienta ma dostęp tylko do internetu i samego systemu bramy VPN. Bezpośrednia komunikacja pomiędzy klientami jest zablokowana na poziomie konfiguracji OpenVPN, a dodatkowo zabezpieczona odpowiednią konfiguracją firewalla.

Po podłączeniu się klienta do VPN, serwer przesyła mu przeznaczoną dla niego konfigurację sieci (sieć typu punkt-punkt z prywatną adresacją). Przykładowo – klient otrzymuje parę adresów 10.0.0.14 i 10.0.0.13, gdzie pierwszy to adres klienta (tunelu po stronie klienta), a drugi to adres tunelu po stronie serwera. Na drugi adres ustawiana jest trasa domyślna w tablicy rutingu klienta. Statyczny adres serwera w sieci VPN to 10.0.0.1. Pod tym adresem serwer jest dostępny dla każdego z klientów podłączonych od VPN. Umożliwia to zaoferowanie klientom dodatkowych usług uruchomionych na serwerze. Dostęp do internetu dla klientów realizowany jest przez mechanizm NAT, który, oprócz funkcji translacji adresów, pełni rolę izolacji klienta od sieci internet (żadne porty sieciowe wirtualnego interfejsu klienta nie są dostępne bezpośrednio z sieci rozległej). NAT realizowany jest przy pomocy standardowych narzędzi sieciowych linuksa (firewall iptables).

Klient ma do dyspozycji dwa serwery OpenVPN pracujące w odmiennych trybach (UDP i TCP). W opisywanym systemie po stronie serwera wykonywany jest skrypt, który zarządza komunikacją pomiędzy dwoma serwerami OpenVPN (każdy z serwerów wykonuje własną wersję tego skryptu). Jeden z serwerów pracuje w trybie UDP, drugi w trybie TCP, oba korzystają z tego samego kompletu klucz/certyfikat. Skrypt jest uruchamiany podczas łączenia się klienta z serwerem, który przy jego pomocy sprawdza, czy dany klient jest już podłączony do drugiego serwera OpenVPN (skrypt serwera UDP sprawdza serwer TCP i odwrotnie). W razie wykrycia otwartej sesji tego klienta z drugim serwerem – odmawia połączenia, klient może być podłączony tylko do jednego z nich w tym samym czasie. Taka konfiguracja pozwala na uczynienie serwisu VPN wysoce dostępnym, przy zachowaniu jednego certyfikatu klienta. W przypadku braku dostępu do zewnętrznego portu UDP 1194 (UDP jest preferowany ze względu na wydajność), klient może połączyć się z drugim serwerem dostępnym pod tym samym adresem IP, na porcie TCP 443, który jest najczęściej otwarty nawet w bardzo restrykcyjnych sieciach korporacyjnych. Serwery te skonfigurowane są w taki sam sposób, a dla klienta posiadają taką samą funkcjonalność. Jednocześnie klient potrzebuje tylko jednego certyfikatu do łączenia się z serwisem, a oba serwery tworzą jeden spójny system VPN.

Połączenia klientów poddane są procesowi kształtowania ruchu sieciowego (ang. *traffic shaping* – kontrolowane jest pasmo dostępne dla każdego klienta. Pasmo ruchu wchodzącego i wychodzącego traktowane jest jak dla łącza synchronicznego (te same wartości download/upload). Arbitralnie przyznano gwarantowane pasmo o wielkości 2 Mbit/s oraz maksymalne o wielkości

4 Mbit/s dla każdego klienta. Kształtowanie ruchu realizowane jest przez standardowe narzędzia systemu linuks – program „tc" z pakietu „iproute2" (potrzebne są również odpowiednie sterowniki jądra linuksa z podsystemu *network scheduler*).

W celu „obrony internetu" przed niektórymi klientami serwisu (np. systemy – ofiary dołączone do sieci typu botnet), dostęp portu smtp (TCP 25) w internecie powinien być traktowany specjalnie (filtrowanie poczty wysyłanej przez klientów – obrona przed spamem ze strony klientów). W opisywanym serwisie stosowane jest dedykowane, przeźroczyste proxy smtp z filtrem antywirusowym i antyspamowym (program smtp-gated). Pakiety kierowane są do proxy za pomocą konfiguracji reguł firewalla linuksowego realizującego funkcje translacji adresów NAT. Klient wysyłający spam jest automatycznie blokowany.

6. Szczegóły konfiguracji OpenVPN oraz przeprowadzone testy

Dostępne algorytmy kryptograficzne zależą od biblioteki SSL i użycie konkretnego z nich definiowane jest na poziomie konfiguracji – domyślnie blowfish ze 128 bitowym kluczem w trybie CBC (ang. *Cipher Block Chaining*). Aby wylistować dostępne metody szyfrowania openssl można użyć polecenia: `openvpn --show-ciphers`. Ze względu na kompromis między bezpieczeństwem a wydajnością wybrany został algorytm szyfrowania AES-128-CBC (o wyborze zadecydowała również kompatybilność z klientem OpenVPN systemu Android).

Testy wydajności algorytmów openssl przedstawiono poniżej:

```
# openssl speed blowfish
type             16 bytes     64 bytes    256 bytes    1024 bytes   8192
bytes
blowfish cbc   84938.66k   89425.86k   91160.58k    91483.14k
91627.52k

# openssl speed aes-128-cbc
type             16 bytes     64 bytes    256 bytes    1024 bytes   8192
bytes
aes-128 cbc    64820.42k 102049.32k 118312.79k   123722.49k
127008.77k

# openssl speed aes-192-cbc
type             16 bytes     64 bytes    256 bytes    1024 bytes   8192
bytes
aes-192 cbc    59339.77k   88614.08k 101157.89k   104770.27k
106838.47k

# openssl speed aes-256-cbc
type             16 bytes     64 bytes    256 bytes    1024 bytes   8192
bytes
```

```
aes-256 cbc   54958.05k  78288.30k  87998.89k  90651.31k
92345.69k
```

Zmierzone wartości oznaczają ilość przetworzonych bajtów w czasie 1000s przy danej wielkości bloku danych. Widoczna jest gorsza wydajność (o ok. 25%) algorytmu aes-128-cbc w porównaniu do blowfish-cbc dla małych (16 bajtów) bloków danych. W przypadku użycia go przez program OpenVPN nie ma to dużego, znaczenia ponieważ pakiety o wielkości 16 bajtów zdarzają się rzadko (np. pakiety uszkodzone). Biorąc pod uwagę cały pakiet – wraz z nagłówkami IP i TCP, gdzie całkowita długość samych nagłówków wynosi 28 (UDP) i 40 (TCP) – przy wielkościach bloku danych 265 i 1024 bajtów widoczny jest wzrost wydajności o odpowiednio 30% i 35%.

W warunkach produkcyjnych, przy maksymalnie wykorzystanym paśmie sieciowym przydzielonym klientowi (4 Mbit/s), wykorzystanie procesora serwera waha się od 2 do 4% (średnio 3%), a wykorzystanie zasobów pamięci jest znikome. Daje to możliwość równoczesnego podłączenia około 20 klientów maksymalnie wykorzystujących pasmo sieciowe (4 Mbit/s) lub 40 klientów, którzy otrzymują pasmo gwarantowane (2 Mbit/s). Powoduje to sumarycznie ok. 60% wykorzystania CPU i obciążenie interfejsu sieciowego ruchem ok. 80 Mbit/s (nie zbadano czy jest możliwe uzyskania tak dużej przepustowości korzystając z infrastruktury firmy OVH). Zależność między ilością przetwarzanych pakietów klienta a wykorzystaniem procesora jest w przybliżeniu liniowa.

Cechą konfiguracji serwera OpenVPN jest możliwość „popychania" (dyrektywa „push") parametrów sieciowych do klientów. Dzięki takiemu zachowaniu się programu, system serwera może całkowicie zrezygnować z wykorzystywania serwera DHCP do konfiguracji sieci klientów. Przydzielanie indywidualnych adresów sieci VPN klientom może odbywać się dynamicznie ze ściśle określonej puli lub w sposób statyczny – przy wykorzystaniu pól CN certyfikatów. Serwer może używać wtedy indywidualnych plików konfiguracyjnych klientów dołączanych dynamicznie do konfiguracji głównej – przetwarzanych na etapie podłączenia klienta.

Zarówno serwer jak i klient mogą wykonywać skrypty podczas łączenia lub rozłączania, co można wykorzystać np. w uniksowych systemach klienckich do konfiguracji ustawień DNS (modyfikacja pliku `resolv.conf`) lub przy dodawaniu dodatkowych tras do tablicy rutingu. Po stronie serwera można stosować skrypty służące np. do dodatkowej kontroli klientów.

Pomimo stosowania kompresji obserwuje się utratę przepustowości – biorąc pod uwagę przydzielone pasmo oraz szerokość pasma fizycznego łącza internetowego klienta. Strata pasma wynosi od 5 do 10% – średnio 7%. Drugim zauważalnym efektem dla połączeń z internetem jest wzrost czasu obiegu pakietów (testy narzędziem ping). Należy tu doliczyć czas wędrówki pakietów do bramy VPN, a następnie do systemu docelowego. W opisywanym systemie

połączenie do bramy VPN wykazuje średni czas obiegu na poziomie ~45 ms. Serwer fizycznie znajduje się poza granicami kraju (we Francji), więc przy testach wykonanych w Polsce dla polskich adresów docelowych, pakiety muszą przebyć trasę Polska–Francja dwukrotnie. Stąd biorą się tak zwiększone czasy obiegu pakietów przy testach prowadzonych na polskich adresach docelowych. W przedstawionych testach, pierwszy odczyt to pomiar na fizycznym łączu, drugi – to pakiety przechodzące przez bramę VPN:

```
gazeta.pl (PL - Wa-wa) 80.252.0.145
rtt min/avg/max/mdev = 15.374/16.435/18.062/1.172 ms
rtt min/avg/max/mdev = 76.235/79.062/82.523/2.263 ms

allegro.pl (PL - Poznan) 193.23.48.134
rtt min/avg/max/mdev = 18.136/19.532/22.238/1.532 ms
rtt min/avg/max/mdev = 80.061/84.482/87.831/3.051 ms

onet.pl (PL - Kraków) 213.180.146.27
rtt min/avg/max/mdev = 14.247/16.087/17.295/1.126 ms
rtt min/avg/max/mdev = 90.076/94.373/98.214/3.140 ms

ripe.net (EU - Holandia) 193.0.19.25
rtt min/avg/max/mdev = 38.590/40.454/43.252/2.021 ms
rtt min/avg/max/mdev = 52.643/53.609/54.204/0.659 ms

slashdot.com (USA - Cal) 216.34.181.45
rtt min/avg/max/mdev = 139.100/139.617/139.915/0.322 ms
rtt min/avg/max/mdev = 151.566/154.721/156.626/2.007 ms

google.pl (multicast? bezpośredni link providera?)
216.239.59.104
rtt min/avg/max/mdev = 62.984/65.050/68.526/2.489 ms
rtt min/avg/max/mdev = 64.582/66.669/68.741/1.478 ms

tunapi.pl (Francja - brama VPN) 91.121.174.23
rtt min/avg/max/mdev = 42.914/44.837/49.611/2.771 ms
rtt min/avg/max/mdev = 44.332/48.177/50.772/2.492 ms
```

Wszystkie testy wykonano z łącza DSL firmy Netia (213.238.117.9) znajdującego się w Krakowie. Największe różnice w czasach obiegu występują oczywiście w przypadku testów adresów docelowych znajdujących się w Polsce (praktycznie podwojony czas obiegu do bramy VPN).

7. Dodatkowe usługi bramy VPN

Dodatkową funkcjonalnością zaprojektowanej sieci – systemu bramy VPN – jest zainstalowany system zapobiegania włamaniom IPS (ang. *intrusion prevention system*), który kontroluje przepływające pakiety klientów przy pomocy techniki DPI (ang. *deep packet inspection*). Funkcję systemu IPS pełni program snort skompilowany w trybie *inline*. Analizuje on wszystkie pakiety klientów (w obu

kierunkach), aż do poziomu warstwy aplikacji (firewall sprawdza tylko nagłówki IP i TCP) oraz reaguje na wzorce ataków na podstawie sygnatur. IPS blokuje pakiety, które wykazują zgodność z sygnaturą ataku. Pakiety kierowane są do programu snort poprzez system firewall (linux iptables) i może to być realizowane dla każdego klienta indywidualnie.

Oprócz funkcji „bezpiecznego serwera sieci internet", opisywany system może pełnić rolę zdalnego serwera plików dostępnego przez VPN. Program SAMBA (odpowiednik serwera plików Windows) został skonfigurowany tak, by udostępniać daną pojemność dyskową przy pomocy protokołu SMB (ang. *Server Message Block*) dla każdego klienta połączonego przez VPN. Zasoby plikowe są indywidualne i izolowane pomiędzy klientami. Usługa jest dostępna jako standardowe „Otoczenie Sieciowe/Moje Miejsca Sieciowe" dla klientów Windows oraz innych systemów operacyjnych wyposażonych w obsługę protokołu SMB.

Istnieje także możliwość zastosowania specjalizowanych systemów proxy w celu przekierowania połączeń lub modyfikacji zawartości transmitowanych danych klientów. Dobrym przykładem jest tu system anonimizacji TOR (ang. *The Onion Routing*) funkcjonujący jako proxy z interfejsem SOCKS. Ta funkcjonalność nie jest obecnie zaimplementowana w opisywanym systemie.

8. Podsumowanie

Przedstawione w pracy rozwiązanie internetowej bramy VPN zwiększa bezpieczeństwo połączenia internetowego i transmisji danych z niezaufanej sieci lokalnej. Zapewnia odporność na podsłuch poprzez analizę pakietów sieciowych, odporność na manipulacje pakietów sieciowych, uniemożliwia śledzenie działalności w Internecie, dostarcza zaufanego serwera DNS. Rozwiązanie to zapewnia również wolność dostępu do usług internetowych (odporność na blokowanie portów usług sieciowych TCP/UDP, adresów IP czy omijanie przeźroczystych serwerów proxy). Dodatkową jego zaletą jest możliwość przydzielenia zawsze tego samego zewnętrznego publicznego adresu IP.

Usługa Bramy VPN w chwili obecnej świadczona jest stosunkowo rzadko i przez niewielką liczbę podmiotów.

LITERATURA

1. Feilner M.: OpenVPN: Building and Integrating Virtual Private Networks. Pact Publishing Ltd., 2006. ISBN 190481185X
2. Microsoft: VPN Tunnels - GRE Protocol 47 Packet Description and Use. Microsoft Knowledge Base, 2007, KB241251.
3. Ryłko K.: VPN za darmo. PC World Komputer, 2005, nr 7.

4. Serafin M.: Sieci VPN. Zdalna praca i bezpieczeństwo danych. Helion, 2008. ISBN: 978-83-246-1521-6
5. Stawowski M.: Projektowanie i praktyczne implementacje sieci VPN. ArsKom, 2004. ISBN: 83-900587-8-2

Rozdział 8

Inżynieria wsteczna w analizie logów komunikatorów internetowych na przykładzie zmieniających się wersji GG

Marek Piotr Stolarski
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Rafał Orlik
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Instytut Fizyki, Politechnika Wrocławska

Borys Łącki
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Streszczenie

Analiza historii rozmów użytkownika sieci Gadu-Gadu (GG) daje nowe narzędzia w walce z przestępczością w sieciach internetowych. Z uwagi na złudne poczucie anonimowości towarzyszące rozmowom prowadzonym za pomocą komunikatorów (np. popularne w Polsce GG) często może dochodzić do łamania prawa. Z powodu zamkniętego protokołu transmisji analiza materiałów archiwalnych jest utrudniona. Wykorzystanie technik inzynierii wstecznej pozwaliło na skuteczne badania materiałów mogących być dowodem w sprawach karnych. Przedstawione narzędzia pozwalają na znaczną automatyzację typowych czynności śledczych.

1. Wstęp

Rozwój dostępu do infrastruktury internetowej spowodował, że coraz większa część komunikacji korzysta z tego kanału. Coraz bardziej popularne

stają się różneho typu komunikatory internetowe, takie jak Gadu-Gadu, Tlen, WP Kontant i inne. Jednak w przeciwieństwie do telefonii, stacjonarnej oraz komórkowej, ten rodzaj komunikacji uchodzi za anonimowy. Przeświadczenie o tym, że w Internecie jest się trudnym, czy też niemożliwym, do zidentyfikowania, skłania część użytkowników do łamania prawa. W tej sytuacji sprawna analiza zabezpieczonych materiałów dowodowych jest kluczową sprawą w postępowaniu prokuratorskim.

W tym rozdziale skupimy się na przedstawieniu analizy logów komunikatorów internetowych na przykładzie progrmau Gadu-Gadu. Wybór ten jest podyktowany faktem, że ten właśnie komunikator jest najpopularniejszy w Polsce. W dalszej cześci skupimy się na analizie archiwów programu Gadu-Gadu w dwóch wersjach, mianowicie, pliku archive.db oraz archives.dat. Formaty te odpowiadają odpowiednio wersji 8.0 (*.db) oraz wcześniejszych (*.dat) komunikatora.

2. Plik archives.dat

Dostępny w starszych wersjach format zapisu wiadomości odebranych oraz wysłanych służył plik archives.dat znajdujący się w katalogu przechowującym profil użytkownika. Jego format, podzielony jest na następujące części: nagłówek, indeks, bloki oraz wiadomości. Ich dokładna struktura zostanie przedstawiona w dalszej części tego rozdziału.

Podstawową jednostką logiczną, która występuje w pliku archives.dat jest nagłówek, zawierający podstawowe informacje na temat profilu właściciela, na przykład numer identyfikacyjny UIN, sumę kontrolną CRC-32, adres kolejnych bloków logicznych. Format nagłówka, ważny z punktu widzenia osoby chcącej czytać archiwum Gadu-Gadu, ma następującą postać:

- 0x00 (4 bajty) – 'RC03' czyli łańcuch identyfikujący archiwum Gadu-Gadu,
- 0x08 (4 bajty) – przesunięcie (liczone od początku pliku), pod tym adresem znajduje się indeks,
- 0x0C (4 bajty) – rozmiar indeksu w bajtach,
- 0x14 (4 bajty) – przesunięcie (liczone od początku pliku) do obszaru zawierające dane,
- 0x24 (4 bajty) – numer identyfikacyjny UIN właściciela profilu, a dokładniej mówiąc UIN^0xFFFFFD66,
- 0x28 (4 bajty) – suma kontrolna, liczona na podstawie algorytmu CRC-32, dla pierwszych N bajtów pliku archives.dat, gdzie N jest wielkością spod adresu 0x14.

Jeżeli suma kontrolna (uwzględniając, że pole przechowujące sumę kontrolną CRC-32 zawiera wartość zero), jest prawidłowa, to można założyć, że plik nie jest uszkodzony. W takim razie analizie można poddać koleje struktury danych

zapisane w pliku archives.dat, mianowicie indeks. Sam indeks zawiera położenie ważnych danych w archiwum, czyli sekcji. Sam indeks ma następującą postać:

- 0x00 (4 bajty) – numer sekcji,
- 0x04 (4 bajty) – liczba bloków, które przynależą do danej sekcji,
- 0x08 (4 bajty) – przesunięcie do pierwszego bloku (liczone od początku obszaru danych, pole 0x14 w nagłówku),
- 0x0C (4 bajty) – przesunięcie do ostatniego bloku.

Jak widać w powyższego opisu, indeks to nic innego jak struktura danych wskazująca na kolejne jednostki logiczne zapisane w pliku archives.dat. Tymi strukturami są bloki. Ich postać jest następująca:

- 0x00 (4 bajty) – suma kontrolna (CRC-32),
- 0x04 (4 bajty) – numer sekcji, do której należy blok.
- 0x08 (4 bajty) – długość bloku,
- 0x0C (4 bajty) – przesunięcie do kolejnego bloku,
- 0x10 (4 bajty) – liczba danych opisywanych przez blok, dane te są zapisane w pliku bezpośrednio po bloku i opisują przechowywane w nim wiadomości.

W tym momencie, podczas analizy archiwum Gadu-Gadu, jesteśmy przy najbardziej interesujących nas danych, mianowicie wiadomościach. Same wiadomości opisane są w dwojaki sposób. Pierwsza część znajduje się w bloku i opisuje przesunięcie do samej wiadomości, druga zaś – przechowuje wiadomość i pola ją opisujące. Z punktu widzenia analizy archiwum najważniejsze dane to:

- 0x00 (4 bajty) – flagi opisujące stan wiadomości, z punktu widzenia analizy archiwum to pole przechowuje informację skasowana (zero) i nieskasowana (jeden),
- 0x04 (4 bajty) – przesunięcie do wiadomości jako takiej,
- 0x08 (4 bajty) – wielkośc wiadomości,
- 0x0C (4 bajty) – przesunięcie do bloku, do którego należy dana wiadomość.

Ostatni etap, potrzeby do odczytania wiadomości, to analiza danych wskazywanych przez kolejne dane zapisane w strukturze blok. Sam opis wiadomości różni się ze względu na jej typ – przychodząca albo wychodząca. Dla wiadomości przychodzącej struktura ją opisująca ma postać:

- 0x00 (4 bajty) – czas wysłania wiadomości (dokładniej liczba sekund od 1. stycznia 1970r),
- 0x04 (4 bajty) – numer nadawcy,
- 0x08 (4 bajty) – wartość zero (0),
- 0x0C (4 bajty) – czas odebrania wiadomości,
- 0x10 (4 bajty) – długość wiadomości w bajtach,
- 0x14 (0x10 bajtów) – zakodowana wiadomość.

Wiadomość, zarówno przychodząca jak i wychodząca, zakodowana jest przy użyciu bardzo prostego algorytmu opartego na operacji XOR: i-ty bajt o_i wiadomości odkodowanej jest wynikiem operacji XOR na i-ty oraz (i-1)-tym bajcie wiadomości zakodowanej. Mianowicie $o_i=e_i \wedge e_{i-1}$, przy dodatkowym założeniu, że e_0=0xFF.

Lekko zmodyfikowaną postać ma struktura opisująca wiadomość wychodzącą:

- 0x00 (4 bajty) – czas wysłania wiadomości,
- 0x04 (4 bajty) – numer nadawcy,
- 0x08 (4 bajty) – liczba odbiorców (dla wiadomości przychodzącej to pole ma wartośc 0),
- 0x0C ([0x08]*4 bajty) – numery UIN odbiorców,
- 0x0C + [0x08]*4 (4 bajty) – czas dostarczenia wiadomości,
- 0x10 + [0x08]*4 (4 bajty) – długość wiadomości,
- 0x14 + [0x08]*4 – zakodowana wiadomość.

W przypadku, gdy wiadomość wychodząca odpowiada de facto wysłaniu krótkiej wiadomości tekstowej (SMS) postać struktury danych ją opisującej jest następująca:

- 0x00 (4 bajty) – czas wysłania wiadomości,
- 0x04 (4 bajty) – nazwa odbiorcy zakończona znakiem 0x00 (w sumie N znaków),
- 0x04 + N (4 bajty) – długość wiadomości (K bajtów),
- 0x08 + N (K bajtów) – zakodowana wiadomość.

Jak widać, format archiwum zapisanego w pliku archives.dat, nie jest bardzo przyjazny dla ewentualnego użytkowniaka. Występuje w nim znaczna liczba wzajemnie zależnych od siebie struktur logicznych, od nagłówka, poprzez indeks, bloki, aż po wiadomości. Wydaje się, że taki a nie inny wybór formatu archiwum nie jest do końca przemyślany, jeżeli chodzi o łatwość dostępu do poszczególnych danych (wiadomości). Także z punktu widzenia ochrony wiadomości wybrany sposób szyfrowania danych nie zapewnia praktycznie żadnego poziomu bezpieczeństwa. Na szczęście dla użytkownoków, wymienione powyżej mankamenty zostały usunięte w kolejnej wersji archiwum, pliku archive.db.

3. Plik archive.db

W komunikatorze Gadu-Gadu, począwszy od wersji 8.0, zmieniony został format pliku przechowującego archiwum wiadomości przychodzących/wychodzących. Mianowicie, twórcy Gadu-Gadu zrezygnowali z własnego formatu pliku na rzecz relacyjnej bazy danych opartej o bibliotekę SQLlite. Takie podejście spowodowało, że

przecjowywnanie logów, a w szczególności dostęp do wybranych wiadomości, został znacznie uproszczony.

Użycie bazy danych SQLite wymusiło opis danych w oparciu o język SQL. Sam plik archive.db to w istocie prosta baza danych zdefiniowana jako:

```
CREATE TABLE chats (
chat_id                    INTEGER PRIMARY KEY NOT NULL,
interlocutor_id     NUMERIC NOT NULL,
is_initialized_by_user NUMERIC NOT NULL,
start_date                 TEXT NOT NULL,
first_communication_item_id  INTEGER DEFAULT 0
);
```

Tabela 'chats' to z punktu widzenia analizy logów komunikatora nic innego jak podział wszystkich rozmów właściciela profilu z uwzględnieniem (i) rozmówcy oraz (ii) czasu wysłania/odebrania wiadomości. Najważniejsze pola to 'chat_id' oraz 'interlocutor_id' określające odpowiednio (i) numer porządkowy rozmowy, oraz (ii) numer porządkowy rozmówcy (klucz w tabeli 'interlocutors'). Pozstałe pola: 'is_initialized_by_user' to określenie typu wiadomości (0 – przychodząca, 1 – wychodząca), 'start_date' – data i godzina rozpoczęcia rozmowy (w formacie yyyy-MM-ddThh:mm:ss), 'first_communication_item_id' – klucz w tabeli 'communication_items' przechowującej poszczególne wiadomości.

Jak zostało wspomniane powyżej, sama treść wiadomośc jest przechowywana w tabeli 'communication_items', której definicja jest przedstawiona poniżej:

```
CREATE TABLE communication_items (
communication_item_id INTEGER PRIMARY KEY AUTOINCREMENT
NOTNULL,
chat_id                    NUMERIC NOT NULL,
is_sent_by_user     NUMERIC NOT NULL,
start_date                 TEXT NOT NULL,
content_type               NUMERIC NOT NULL,
content                    TEXT,
plain_text_content         TEXT
);
```

Pole 'chat_id' tej tabeli wiąże dany komunikat z poszczególnym rozmówcą określonym poprzez pole 'interlocutor_id' tabeli 'chats'. Nowe w stosunku do tabeli 'chats' pola 'content_type', 'content' oraz 'plain_text_content' określają odpowiednio: typ wiadomości (0 – rozmowa, 1 – SMS, 2 – transfer pliku), treść komunikatu (łańcuch HTML dla rozmowy, plik XML dla pozostałych), łańcuch wykorzystywany przy przeszukiwaniu archiwym (może być to łańcuch pusty).

Ostatnia, trzecia, tabela zdefiniowana w pliku archive.db to 'interlocutors' postaci:

```
CREATE TABLE interlocutors(
interlocutor_id INTEGER PRIMARY KEY NOT NULL,
identification_type NUMERIC NOT NULL,
identification TEXT NOT NULL
);
```

Zawiera ona, w zależności od potrzeby (wiadomość, SMS) albo numer UIN rozmówcy albo inny jego identyfikator (np. imię i nazwisko w przypadku SMS'ów).

Jak widać, w porównaniu do archiwum przechowywanego w pliku archives.dat, zastosowany w nowej wersji komunikatora Gadu-Gadu format logów jest znacznie bardziej ujednolicony. W szczególności, wszelkiego typu zapytania o (i) rozmowy z danym rozmówcą, (ii) rozmowy z zadanego przedziału czasowego, i inne sprowadzają się do napisania odpowiedniego zapytania typu 'SELECT * FROM ...' w języku SQL. Cały mechanizm znalezienia odpowiednich rekordów spoczywa na silnuku bazy danych, zaś osoba analizująca logi komunikatora ma dużo większą swobodę w przeszukiwaniu archiwum.

Także sprawa zapewnienia prywatności została znacznie lepiej potraktowana. Mianowicie, sam plik archive.db może być zakodowany przy użyciu znacznie silniejszego algorytmu. W tym wypadku jest to AES256, w przeciwieństwie do zastosowanej w pliku archives.dat metody opartej na jednokrotnej operacji XOR. Oczywiście, w przypadku zaszyfrowanego pliku archive.db jego analiza jest uzalezniona od (i) znajomości hasła, (ii) możliwości jego szybkiego złamania (metoda brute-force, słownikowa i inne).

4. Podsumowanie

Przedstawione w tym rozdziale rozwiązania pozwalają na odczyt, a tym samym dalszą analizę, logów komunikatora Gadu-Gadu. Zastosowanie wyspecjalizownych programów pozwala na znaczne ułatwienie typowej pracy informatyka śledczego, polegającej na wyszukiwaniu np. zadanych fraz w rozmowach prowadzonych przez właściciela profilu Gadu-Gadu. Z drugiej jednak strony, istnienie dwóch całkowiecie różnych formatów przechowyania logów (achives.dat, archive.db) sprawia, że sam program do analizy musi być starannie zaplanowany. Wydaje się celowym konwertowanie obu plików do wybranego formatu pośredniego, który jest obsługiwany przez program analizujący. W takim wypadku wprowadznie w przyszłości nowego typu archiwum wymusza jedynie napisanie wyspecjalizowanej aplikacji konwertującej.

LITERATURA

Rozdział powstał wyłącznie w oparciu o własne opracowania Autorów.

Rozdział 9

Efektywne wyszukiwanie podobieństw w tekstach źródłowych

Marek Piotr Stolarski
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Rafał Orlik
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Instytut Fizyki, Politechnika Wrocławska

Mateusz Kocielski
iConsulting Marek Piotr Stolarski
iconsulting@iconsulting.pl

Streszczenie

*Znajdowanie podobieństw w tekstach źródłowych jako narzędzie kontroli własności intelektualnej. Metodyka tworzenia efektywnego, zarówno pod względem prędkości jak i dokładności, systemy wykrywania podobieństw. Typowe metody oszukiwania systemu i sposoby ich unikania. Normalizacja tekstu źródłowego taka jak konwersja z *.pdf, *.html, *.odt, *.doc do *.txt; wykorzystanie słowników synonimów itp. Różne języki w jednym dokumencie. Sposoby radzenia sobie z nowymi słowami, które nie występują w słowniku.*

1. Wprowadzenie

W dzisiejszym świecie informacja jest janważniejszą rzeczą. Kto kontroluje informację ten kontroluje innych. To sprawia, że jest ona bardzo podatna, na przykład może być ukradziona z czyjejś strony WWW. W tym rozdziale przedstawimy nasz system wykrywania podobieństw w tekstach źródłowych. System ten, który początkowo był tworzony pod kątem wykorzystania go w

jednostkach akademickich, po małych zmianach może być użyty na przykład przez agencje prasowe, portale WWW itp. W następnym rozdziale przedstawimy metodologię tworzenia takiego systemu oraz typowe problemy, które muszą być rozwiązane aby był on odporny na atak. Cały system używa programów Open Source takich jak Python (aplikacje po stronie serwera), PostgreSQL (baza danych do przechowywania tekstów), Apache (strona użytkownika).

2. Metodologia tworzenia efektywnego algorytmu znajdowania podobieństw

Najprostszą drogą znajdowania podobieństw jest użycie akgorytmów wyszukiwana łańcuchów, na przykład algorytm naiwny, Rabina-Karpa, Knutha-Morrisa-Pratta. Takie podejście pracuje dobrze dla identycznych tekstów. Metoda ta, aczkolwiem poprawna, będzie bardzo wolna i łatwa do obejścia. Można ją przyspieszyć poprzez podzielenie tesktu źródłowego na mniejsze fragmenty, np. długości 10 wyrazów, i wyznaczenie dla każdego z nich funkcji skrótu, np. MD5. Wtedy, poprzez proste odpytywania bazy danych o te teksty, które zawierają identyczną sygnaturę, można otrzymać te dokumenty, które składają się z takich samych fragmentów tekstów. Z powodu bardzo małego prawdopodobieństwa tego, że różne fragmenty tesktu dadzą w wyniku tą samą sygnaturę MD5, można przyjąć, że taki algorytm wykrywa podobieństwa pomiędzy tekstem źródłowym a innymi tekstami.

Podejście wymienione wcześniej można mardzo łatwo oszukać, gdy osoba kradnąca fragment cudzego testu źródłowego, zmieniająca go nieznacznie, np. poprzez zmianę konstrukcji zdania. Ostatecznie więc, taki algorytm systemu wykrywania podobieństw nie będzie działał dla dostatecznie mądrego złodzieja.

Tak więc, musimy stworzyć inną metodę, która będzie porównywalnie szybka jak metoda bazująca na sygnaturach MD5, ale w przeciwieństwie do niej nie będzie tak łatwa do oszukania.

3. Jak znormalizować tekst źródłowy?

Z powodu wielu możliwych do wykorzystania typów dokumentów tekstowych, takich jak *.pdf, *.ps, *.odt, *.doc, *.txt oraz inne, doskonały system wykrywania podobieństw musi dokonać procesu normalizacji. Najprostszą drogą jest konwersja dokumentu do formatu tekstowego. Biorąc pod uwagę istnienie alfabetów, w których występują specyficzne znaki narodowe (znaki diakrytyczne), trzeba dokonać wyboru pomiędzy np. kodowaniem ISO-8859-1 lub UTF-8/UTF-16. Użycie kodowanie Unicode wydaje się być najlepszym rozwiązaniem, ponieważ każdy znak diakrytyczny ma w nim swoją reprezentację. Niestety, takie podejście nie jest najlepsze. W rzeczywistości

użycie Unicode sprawia, że analiza staje się trudniejsza, ponieważ niektóre znaki mogą być zakodowane w kilka równoważnych sposobów. To sprawia, że użycie kodowania Unicode staje się złym wyborem z praktycznego punktu widzenia.

Należy także wziąć pod uwagę inne zagadnienie. Mianowicie, przypuśćmy, że dokument źródłowy został poddany procesowi OCR (Optical Character Recognition). Podczas niego część znaków mogła być błędnie rozpoznana, np. polska litera „a" mogła być zamieniona na „a". Dla człowieka taki błąd jest oczywiście łatwy do wykrycia i usnięcia, ale w przypadku komputerów staje się niebanalnym problemem. Tak więc, musimy sibie poradzić z niezwykle istotnym zagadnieniem: jak uniknąć takich sytuacji? W naszym systemie zdecydowaliśmy się traktować wszystkie znaki diakrytyczne na równi z ich łacińskimi odpowiednikami. Mianowicie, wszystkie wymienione znaki „z", „ż", „ź" są konwertowana do „z". Co więcej, podczas procesu normalizacji wszystkie znaki są zamieniana na ich odpowiedniki pisane małą literą. W tym momencie widać wyraźnie, że wybór kodowania padł na ISO-8859-1.

Oczywiście, w dokumencie źródłowym występują także znaki istotne w tekście pisanym, ale, gdy mówiony, są trudne do rozróżnienia. Także więc znaki takie jak „," (przecinek), „." (kropka), „;" (średnik) należy traktować w sposób odmienny. My zdecydowaliśmy się na ich usunięcie z tekstu źródłowego. Ostatecznie więc tekst "Litwo! Ojczyzno moja! ty jesteś jak zdrowie. Ile cię trzeba cenić, ten tylko się dowie, Kto cię stracił. Dziś piękność twą w całej ozdobie Widzę i opisuję, bo tęsknię po tobie" będzie znormalizowany do postaci "litwo ojczyzno moja ty jestes jak zdrowie ile cie trzeba cenic ten tylko sie dowie kto cie stracil dzis pieknosc twa w calej ozdobie widze i opoisuje bo tesknie po tobie".

Taka metoda będzie wystarczająca dla prostego systemu wykrywania podobieństw. Jednak jak radzić sobie z innymi metodami jego unikania? To zagadnienie będzie przedstawione w kolejnym rozdziale.

4. Typowe metody unikania wykrycia podobieństw i sposoby radzenia sobie z nimi

W poprzednim rodziale przedstawiliśmy proces normalizacji dokumentu źródłowego. Co się jednak stanie, gdy ktoś zmieni nieznacznie test oryginalny? Mianowicie, w dokumencie źródłowym zamieni zdanie „Tomek poszedł to szkoły" na „On poszedł do szkoły". Tak długo jak „Tomek" i „On" znaczą dokładnie to samo, oba zdania mają to samo znaczenie. Inny przykład takiego celowego działania: zamiast przepisać „Tomek poszedł do szkoły. Pogoda była mglista" ktoś napisał „Pogoda była mglista, gdy Tomek szedł do szkoły". Trzeci możliwy przypadek to użycie synonimów: „Mieszkanie było puste" zamiast „Mieszkanie było do wynajęcia". Oba zdania znaczą dokładnie to samo, jednak

do wyrażenia tej myśli zostały użyte inne wyrazy. Poraz kolejny prosty system wykrywania podobieństw przestanie działać poprawnie.

Wszystkie wymienione powyżej przykłady pokazują, że projektowanie mądrego systemu wykrywania podobieństw nie jest zagadnieniem trywialnym. Aby uniknąć wspomnianych problemów można użyć następujących technik: zamienić wszystkie słowa na ich rdzenie („robię" na „robić"), sprawdzać czy w tekście źródłowym występują synonimy, a jeżeli tak, to zamienić je na jednego wybranego reprezentanta z każdej z grup synonimów.

Te wszystkie metody użyte razem sprawiają, że jeżeli ktoś chcący oszukać taki system wykrywania podobieństw będzie zmuszony do tak znaczących zmian w tekście źródłowym, że trudno już będzie mówić o kradzieży.

5. Słownik i słowa, które w nim nie występują

W tym momencie nasz system wykrywania podobieństw wydaje się działać niemalże bezbłędnie. Jako dane wejściowe otrzymuje dokument źródłowy, który następnie jest konwertowany do postaci tekstowej. Przy uzyciu technik sprawiających, że oszukanie systemu staje się trudniejsze, słowa tekstu źródłowego zostają zamienione na odpowiadające im rdzenie. Ponieważ komputery pracują korzystając wyłącznie z liczb, rdzenie są zamieniane na np. liczby całkowite.

Pozostaje jednak otwarte pytanie: co zrobić ze słowami, których nie znamy, ponieważ nie występują w wykorzystywanym przez nas słowniku rdzeń → możliwe odmiany? Jednym z możliwych rozwiązań jest obliczenie nowego numeru idetyfikacyjnego dla słowa. Należy przy tym pamiętać, że słowo to może być jedną z odmian rdzenia. Zazwyczaj rdzeń ten może być w łatwy sposób wyznaczony, gdyż w ogólnym przypadku każde słowo ma formę prefix-core-suffix, gdzie prefix to (najczęściej) „nie-", „naj-", itp. Tak więc poprzez usunięcie typowych przedrostków otrzymamy formę core-suffix naszego nowego słowa. Co więcej, tylko kilka pierwszych liter należeć będzie to rdzenia. W tym momencie mamy praktycznie rozwiązany nasz problem: wyznacznie identyfikatora dla nowego słowa. Mianowicie, przypiszmy każdej literze alfabetu pewną liczbę. Wtedy, poprzez sumowanie ich z ustalonymi wagami (zależnymi od odległości od początku słowa: duże wartości wag dla początkowych liter; coraz mniejsze, dążące do zera dla pozostałych) wyznaczymy poszukiwaną wartość.

Z algorytmem przedstawionym powyżej nadal są związane pewne otwarte pytania, takie jak np. postać funkcji wagowej, wartości liczbowe odpowiadające każdej z liter. Z drugiej jednak strony jest to dobry punkt wyjścia. Co więcej, dla każdego ze słów, które mają ten sam rdzeń, metoda ta powinna dać tą samą (albo bliskie siebie) wartości. Ostatecznie więc, wszystkie słowa, występujące w

dokumencie źródłowym, są zamienione na ich reprezentację liczbową i w takiej formie są ostatecznie zapisywane w bazie danych.

6. Dwa i więcej języków w tekście źródłowym

System wykrywania podobieństw, który został opisany w poprzednim rozdziale, działa bezproblemowo w przypadku dokumentu napisanego w jednym języku. Co stanie się jednak, gdy drugi i więcej języków pojawi się w rozpatrywanym tekście? Zazwyczaj język obcy oznacza cytowanie tekstu zagranicznego. Wtedy więc zasadniczy język dokumentu może być stosunkowo łatwo określony poprzez zliczenie jaka część tekstu należy do danego języka. Aby dodatkowo usprawnić ten proces moża wymagać od użytkownika systemu wstępnego określenia używanych w dokumencie języków. Najprostsza metoda jest postaci: jak długo dane słow może być zidentyfikowane jako należące do konkretnego języka, tak długo do jego reprezentacji liczbowej można dodać stałą wartość (różną dla każdego z języków). Jeżeli tak otrzymane liczby dla różnych języków nie przeplatają się, to system wykrywania podobieństw będzie działał bez istotniejszych zmian.

7. System wykrywania podobieństw

Proces normalizacji został zakończony. Tekst źródłowy został przetransformowany ze swojej oryginalnej postaci, np. pliku *.pdf, do reprezentacji liczbowej. Podczas tych czynności pewne dodatkowe techniki były użyte, takie jak wykrywanie synonimów, zastępowanie słów poprzez odpowiadające im rdzenie. Można więc rozpocząć sprawdzanie czy występują podobieństwa pomiędzy tekstem źródłowym a tekstami zachowanymi w bazie danych. Powstaje jednak pytanie: jak wykrywać owe podobieństwa? W rozpatrywanym systemie użyliśmy bardzo prostego, ale skutecznego, algorytmu. Mianowicie, niech tekst znormalizowany zostanie podzielony na ramki, każda o długości N słów. Następnie, dla każdej ramki S odpytujemy bazę danych czy istnieją inne ramki S_i (należące do innych dokumentów), takie że moc koniunkcji zbiorów ($S \cap S_i$) jest większa niż $M<N$. Można przypuszczać, że dwie ramki (S z dokumentu źródłowego oraz S_i z innego dokumentu) mają wspólne M/N słów. Stąd, współczynnik podobieństwa r pomiędzy nimi może być zdefiniowany jako $r=M/N$. Oczywiście, nasz system może jedynie stwierdzić czy istnieje niezerowe prawdopodobieństwa takiego zdarzenia: dla całkowicie różnych ramek $r=0$, dla dokładnie takich samych $r=1$. Informacja ta jest zachowywana następnie w bazie danych. Przedstawiony algorytm wyszukiwana podobieństw jest najbardziej wrażliwą częścią systemu. Jego administrator musi niezwykle starannie dobrać wartości współczynników N oraz M. Liczby te nie mogą być zbyt małe (algorytm „widzi" jedynie część zdania), ani zbyt duże (zbyt dużo następujących po sobie zdań brane jest pod uwagę).

Dość zgrubne oszasowanie wartość parametru N jest takie, że powinno ono odzwieciedlać średnią długość jednego-dwóch zdań. W przypadku parametru M można przyjąć, że M wynosi ok. 75-80% wartości N. Etap odpytywania bazy danych jest najwolniejszą częścią systemu. Dane doświadczalne pokazuję, że o ile proces normalizacji tekstu źródłowego zajmuje czas rzędu minut, to już wyszukiwanie w bazie danych może wymagać czasu rzędu godzin i więcej, w zależności od rozmiaru bazy danych i uwarunkowań sprzętowych komputera.

8. Podsumowanie

Przedstawiliśmy krok po kroku metodę projektowania systemu do wykrywania podobieństw w tekstach źródłowych: począwszy od momentu otrzymania dokumenu, poprzez proces jego normalizacji, skończywszy na przechowywaniu go w bazie danych. Zaprezentowaliśmy typowe sposoby oszukiwania systemu wykrywania podobieństw (zamiana kolejności zdań, wykorzystywanie synonimów) oraz metody im przeciwdziałania. Pokazaliśmy także jak radzić sobie z problemem słów niewystępujących w słowniku.

LITERATURA

Rozdział powstał wyłącznie w oparciu o własne opracowania Autorów.

Część 2.
Efektywność systemów informatycznych

Rozdział 10

Ryzyko w przedsięwzięciach informatycznych

Dariusz Dymek
Uniwersytet Ekonomiczny w Krakowie
dariusz.dymek@uek.krakow.pl

Streszczenie

Budowa rozwiązań wykorzystujących technologię informatyczną jest obecnie codziennością większości organizacji, a ich umiejętne wykorzy-stanie może decydować o przewadze konkurencyjnej. Jak każdemu dzia-łaniu, również przedsięwzięciom informatycznym towarzyszy ryzyko, a skuteczne zarządzanie ryzykiem może decydować o powodzeniu przed-sięwzięcia. Specyfika przedsięwzięć informatycznych sprawia, że do tego zagadnienia należy podchodzić ze szczególną uwagą, a zarządzanie ryzy-kiem w przedsięwzięciach informatycznych musi mieć charakter komple-ksowy i stały, nie ograniczony jedynie do projektu informatycznego, lecz obejmujący również eksploatację systemu. W szczególności musi być ono podporządkowane celom strategicznym organizacji i zgodne jej polityką w zakresie zarządzania ryzykiem.

1. Projekt, a przedsięwzięcie

W języku polskim pojęcie projektu jest wykorzystywane w dwóch znaczeniach. W węższym znaczeniu, projekt jest rozumiany jako praca analityczno-badawcza, koncepcja budowy i funkcjonowania jakiegoś systemu, dokumentacja techniczna czy plan działania. W szerszym znaczeniu projekt jest działaniem kompleksowym w skład, którego wchodzą opracowania studialne, modele (projekty rozwiązań) oraz efekty rzeczowe [12]. Wyróżnikiem projektów jako formy działalności są ich tymczasowość (określony początek i koniec), ukierunkowanie na cel, złożoność, wyjątkowość (unikalność) oraz określone zasoby i koszty. W dalszej części posłu-giwać będziemy się definicją Wysockiego i McGary'go, którzy definiują projekt jako *„sekwencję niepowtarzalnych, złożonych i powiązanych ze sobą działań mających wspólny*

cel, przeznaczonych do wykonania w określonym terminie, bez przekraczania ustalonego budżetu, zgodnie z założonymi wymaganiami" [13].

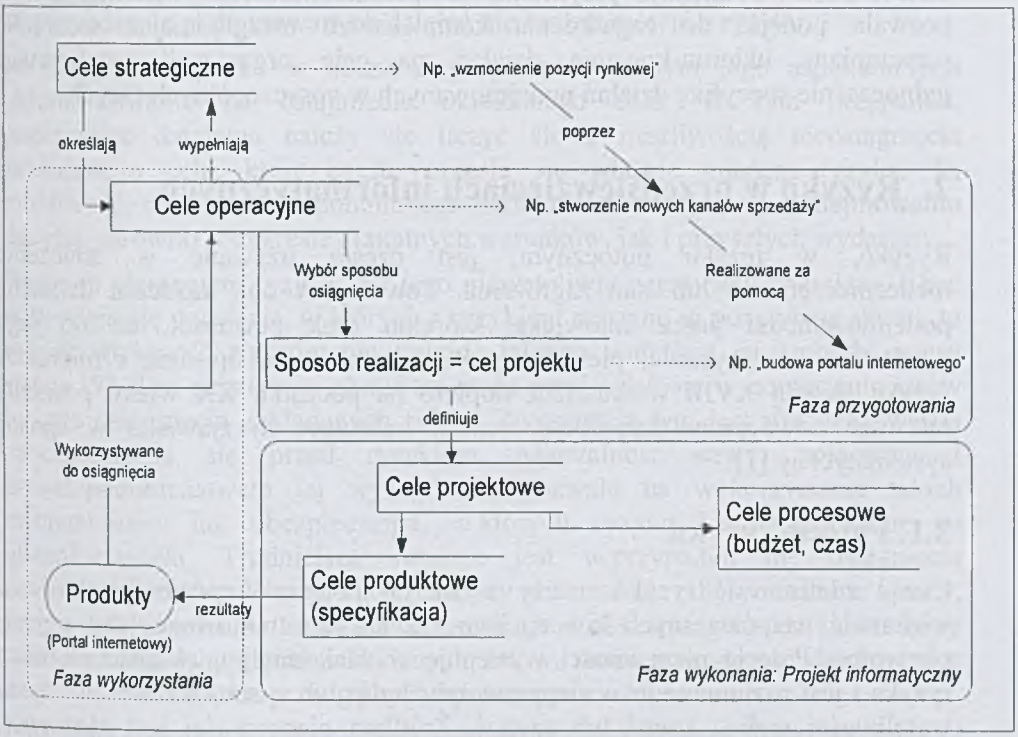
Istotnym aspektem jest instytucjonalny charakter projektu. Projekt stanowi wydrę-bnioną jednostkę organizacyjną powołaną na czas jego trwania, o zdefiniowanych zadaniach i dysponującą własnym zespołem i budżetem. Na czele zespołu projekto-wego stoi Kierownik Projektu, który kieruje całością projektu i odpowiada za jego realizację. Ciałem nadzorującym projekt jest Komitet Sterujący, na czele którego stoi Przewodniczący Komitetu Sterującego (Sponsor). Do zadań Komitetu Sterującego należy kontrolowanie przebiegu projektu oraz podejmowanie decyzji wykraczających poza kompetencje kierownika projektu, w szczególności w zakresie terminów, zasobów, budżetu i wymagań.

W przypadku projektów informatycznych kolejnym i wyróżnikiem jest forma realiza-cji projektów. Ze względu na wysoki poziom złożoności oraz wymaganą wiedzę z zakresu technologii informatycznej, znakomita większość projektów informaty-cznych jest realizowana przez wyspecjalizowane firmy, zewnętrzne w stosunku do organizacji klienta. Oznacza to, że projekt informatyczny jest realizowany na styku dwóch niezależnych organizacji, w ścisłej współpracy, która bardzo często trwa długo po zakończeniu samego projektu. Taka sytuacja powoduje powstanie, rzadko spotykanych w przypadku innego rodzaju projektów, specyficznych uwarunkowań.

Aby w pełni opisać te specyficzne uwarunkowania musimy wyjść od działań podejmowanych przed rozpoczęciem projektu i celu realizacji projektu informaty-cznego. Motorem podejmowania działań przez daną organizację jest przyjęta przez nią strategia. W ramach niej, organizacja określa cele strategiczne oraz zakładany horyzont czasowy ich osiągnięcia. Następnym krokiem jest wybór sposobów osiąg-nięcia celów strategicznych, czyli określenie celów operacyjnych, w ramach których wskazywane są konkretne rozwiązania i specyfikowane potrzebne narzędzia. Ogół tych działań zaliczymy do fazy przygotowania bezpośrednio poprzedzającej fazę wykonania, która w przypadku, gdy potrzebne rozwiązania bazują na technologii informatycznej, przyjmuje formę projektu informatycznego. Ostatnią fazą jest faza eksploatacji, w której rezultaty projektu są wykorzystywane do osiągnięcia pierwotnie zakładanych celów operacyjnych, a w konsekwencji celów strategicznych. Wszystkie te trzy fazy składają się na przedsięwzięcie informatyczne[1]. Ogólny schemat przedsięwzięcia informatycznego przedstawia rysunek 1, a szczegółowy opis zależ-ności pomiędzy celami poszczególnych faz jest opisany w [7].

---

[1] Podobny podział przedsięwzięcia na fazy występuje np. dla przedsięwzięć inwestycyjnych. Najczęściej wyróżnianie są fazy: przediwestycyjna (planowania), budowy (realizacji), operacyjna (eksploatacji) oraz faza likwidacji [11].



Rys. 1. Ogólny schemat przedsięwzięcia informatycznego [7]

Jak widać na powyższym schemacie projekt informatyczny stanowi jedynie jedną z faz przedsięwzięcia. Takie wyodrębnienie faz w przypadku przedsięwzięć informa-tykcznych jest podyktowane kilkoma istotnymi czynnikami. Do najważniejszych należą:

- ukierunkowanie przedsięwzięcia na cele organizacji,
- specyfika i różnorodność działań podejmowanych w poszczególnych fazach,
- skład osobowy zespołów pracujących w różnych fazach i ich formalny status wewnątrz organizacji.

Warto podkreślić, że nie jest możliwe włączenie tych faz w projekt informatyczny, bez naruszenia podstawowych elementów definicji projektu, takich jak ograniczenia czasowe czy określoność budżetu. Dodatkowo, wspomniana wcześniej praktyka wykorzystywania zewnętrznych wykonawców sprawia, że mogą oni brać udział w przedsięwzięciu najwcześniej pod koniec fazy przygotowawczej, na etapie definio-wania wymagań. Z drugiej strony, wybór przyszłego wykonawcy, bez znajomości tych wymagań jest znacząco utrudniony.

Rozróżnienie pomiędzy projektem, a przedsięwzięciem informatycznym pozwala podejść do zagadnienia kompleksowo uwzględniając wcześniej wspomniane ukierun-kowanie działań na cele organizacji, zachowując jednocześnie specyfikę działań podejmowanych w poszczególnych fazach

2. Ryzyko w przedsięwzięciach informatycznych

Ryzyko, w języku potocznym, jest często używane w znaczeniu niebezpieczeństwo lub stan zagrożenia. Towarzyszy ono każdemu działaniu podejmowanemu przez człowieka, któremu brak pewności, co do jego rezultatów. Historycznie, pierwsze próby zrozumienia, opisania i mierzenia ryzyka sięgają XVIII wieku, lecz dopiero na początku XX wieku powstały pierwsze nowoczesne koncepcje ryzyka, traktujące to zjawisko w sposób systematyczny [1].

2.1. Pojęcie ryzyka

Chcąc zdefiniować ryzyko należy wyjść od pojęcia niepewności, które w większości współczesnych koncepcjach ryzyka jest traktowane jako pojęcie pierwotne. Pojęcie niepewności występuje w znakomitej większości definicji ryzyka i jest rozumiane jako nieprzewidywalność lub synonim zawodności czy wątpliwości wobec kogoś lub czegoś. Źródłem niepewności jest złożoność, nieokreśloność i nieciągłość zjawisk społecznych i ekonomicznych.

Rozróżnienie pomiędzy ryzykiem, a niepewnością jest oparte przede wszystkim na kwestii mierzalności. W takim podejściu ryzyko jest mierzalną formą niepewności (F.H.Knight). Bazując na tej samej koncepcji, ale wykorzystując rachunek prawdo-podobieństwa J.M. Keynes definiuje ryzyko następująco: o ryzyku mówimy wtedy, gdy jesteśmy w stanie określić zakres przyszłych zdarzeń oraz przypisać im prawdopodobieństwo wystąpienia. W przeciwnym wypadku, gdy zakres przyszłych zdarzeń jest nieokreślony i tym samym nie możemy im przypisać prawdopo-dobieństwa wystąpienia, mamy do czynienia z niepewnością. Takie podejście, bazu-jące na rachunku prawdopodobieństwa, odcisnęło bardzo silny ślad na dzisiejszym podejściu do ryzyka, które przez wielu postrzegane jest właśnie przez pryzmat prawdopodobieństwa wystąpienia zdarzeń o charakterze negatywnym. Należy jednak pamiętać, że nie jest to jedyna koncepcja, choć niewątpliwie ze względu na jej matematyczne podstawy jest ona wygodna w użyciu.

Warto zwrócić uwagę na inny element związany z ryzykiem. O ile źródłem niepew-ności jest szeroko rozumiane otoczenie, o tyle źródłem ryzyka jest sam człowiek, podejmujący działania w warunkach braku pełnej informacji o jego skutkach (w warunkach niepewności). Ten brak informacji, określany również mianem deficytu informacji, jest naturalną konsekwencją nieprzewidywalności przyszłości i towarzyszy każdemu działaniu człowieka. Należy podkreślić, że

deficyt informacji jest istotnym elementem definicji ryzyka: przy pewności niepowodzenia podejmowanych działań nie można mówić o ryzyku.

W przypadku ryzyka w obszarze gospodarki ważnym jego aspektem jest ukierun-kowanie na osiągnięcie określonego celu. W tym przypadku, podejmując działania należy się liczyć się z możliwością nieosiągnięcia zakładanego celu, które często określa się właśnie mianem ryzyka. Ta możliwość, wynika ze wspomnianego deficytu informacji przy podejmowaniu decyzji, zarówno w zakresie aktualnych warunków, jak i przyszłych wydarzeń.

Ważnym elementem ryzyka jest jego niewątpliwie negatywny charakter. Choć poja-wiają się podejścia, w których z ryzykiem wiązane są pozytywne skutki, to jednak większość autorów nie traktuje takiego podejścia jako ryzyka *sensu stricte* [9]. Ten negatywny charakter może przejawiać się w poniesieniu straty lub nie osiągnięcia zakładanych celów. Związana z tym jest również kwestia zabezpieczenia się przed ryzykiem. Mierzalność straty, połączona z prawdopodobieństwem jej wystąpienia, pozwala na wykorzystanie takich mechanizmów jak ubezpieczenia, w których ryzyko jest przenoszone na ubezpieczyciela. Trudniejsza sytuacja jest w przypadku nie osiągnięcia zakładanych celów. Wtedy, choć jesteśmy w stanie określić poniesione koszty, to nie są one tożsame ze stratą. Szczególnie w obszarze gospodarki, utracony czas oraz nie osiągnięcie zakładanych celów, może powodować daleko idące nega-tywne skutki, które mogą być nierekompensowalne.

Powiązanie ryzyka z człowiekiem i podejmowanymi przez niego działaniami sprawia, że w zależności od dziedziny tej działalności możemy wyróżnić różne rodzaje ryzyka. I tak, Kaczmarek [9] wyróżnia ponad 20 rodzajów ryzyka, w tym ryzyko ubezpie-czeniowe, ekonomiczne, kredytowe, finansowe, produkcyjne, organizacyjne, nowych technologii, medyczne, farmaceutyczne, ekologiczne, chemiczne, prawne, polityczne, socjologiczne, kulturowe i wiele innych. Warto zwrócić uwagę, że podział na rodzaje ryzyka nie jest rozłączny i poszczególne rodzaje ryzyka mogą zawierać w sobie elementy ryzyka innego rodzaju. Wynika to właśnie z faktu ukierunkowania na dziedzinę działalności, a ta podlega wpływom zdarzeń pochodzących z różnych obszarów otoczenia.

2.2. Rodzaje ryzyka w przedsięwzięciach informatycznych

Przedsięwzięcia informatyczne, szczególnie te podejmowane w obszarze gospodarki, możemy zaliczyć do działań o charakterze zarówno organizacyjnym jak i ekonomicz-nym. Charakter ekonomiczny wynika z ukierunkowania na cele organizacji, natomiast charakter organizacyjny przejawia się w instytucjonalnej formie przedsięwzięcia oraz jego wpływie na funkcjonowanie organizacji. Z tego powodu, mówiąc o ryzyku przedsięwzięć informatycznych konieczne jest uwzględnienie wszystkich tych aspektów i związanych z nią ryzyk.

Wyróżnione wcześniej fazy przedsięwzięcia informatycznego grupują działania nie tylko ze względu na ich następstwo czasowe, ale również ze względu na

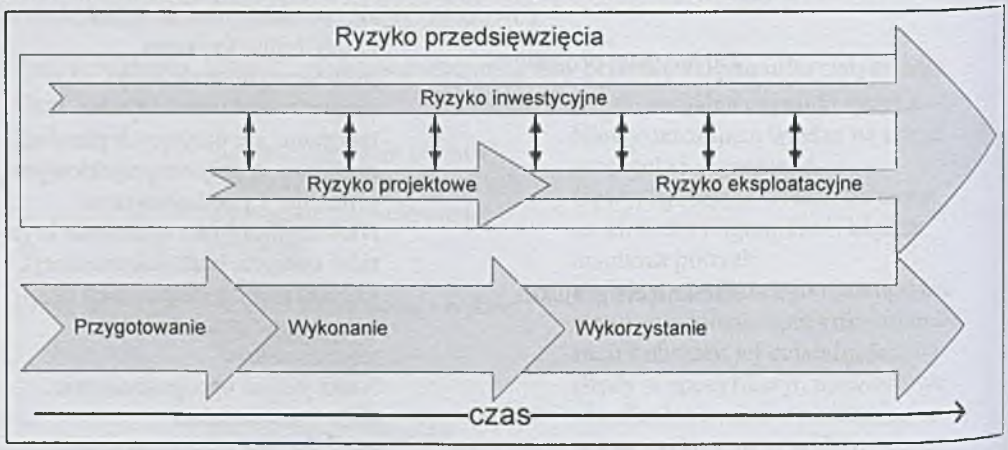
charakter podejmowanych działań. I tak, faza przygotowania odpowiada za określenie sposobu osiągnięcia celów operacyjnych poprzez wskazanie potrzebnych narzędzi i sposobów działania, faza wykonania odpowiada za wykonanie tych narzędzi (budowę systemów informatycznych) i ich wdrożenie (przekazanie do eksploatacji), natomiast faza wykorzystania odpowiada za właściwe użycie stworzonych narzędzi do osiągnięcia wcześniej wyznaczonych celów. Każda z tych faz posiada własne, dobrze określone cele. W konsekwencji, w każdej z nich pojawia się specyficzne ryzyko związane z możliwością ich nieosiągnięcia. Zestawienie wybranych czynników ryzyka w rozbiciu na fazy zawiera tabela 1. Nie jest to zestawienie kompletne. Mnogość potencjalnych czynników ryzyka w każdej z faz oraz ich różnorodność związana ze specyfiką podejmowanych działań powoduje, że sporządzenie kompletnej listy w przypadku ogólnym jest prawie niemożliwe. Już dla samej fazy wykonania, w ramach ryzyka projektowego, w zależności od przyjętych kryteriów można wyróżnić kilkadziesiąt czynników ryzyka [10]. Niemniej jednak, można zauważyć interdyscyplinarność czynników ryzyka, które dotyczą zarówno zagadnień technicznych (wybór technologii), zarządzania (wybór metodyki projektowej), jak i psychologicznych (dobór członków zespołu projektowego ze względu na cechy osobowe).

Tab. 1. Wybrane czynniki ryzyka w rozbiciu na fazy przedsięwzięcia informatycznego

Faza	Działania	Potencjalne czynniki ryzyka
Przygotowanie	Operacjonalizacja celów i definiowanie alternatywnych rozwiązań, zdefiniowanie problemu	Niewystarczająca wiedza na temat otoczenia i organizacji, niewystarczająca wiedza na temat możliwości i ograniczeń, błędna diagnoza potrzeb
	Analiza możliwości i wybór rozwiązania	Wybór niewłaściwego rozwiązania, nieuwzględnienie specyfiki organizacji i obszaru jej działalności
	Definiowanie projektu	Błędy w specyfikacji, niewłaściwe określenie terminów i budżetu, wybór wykonawcy, wybór technologii
Wykonanie (projekt informatyczny)	Organizowanie projektu	Błędy w zakresie identyfikacji celów, użytkowników, wymagań; wybór niewłaściwej metodyki zarządzania projektem, błędy w zakresie definiowania infrastruktury projektu
	Zarządzanie jakością	Nieokreślenie kryteriów jakościowych, niewystarczające nakłady na działania projakościowe,
	Tworzenie zespołu projektowego	Niewłaściwy dobór członków zespołu projektowego(ze względu na kwali-fikacje, cechy osobowe,

		doświa-dczenie), niewłaściwe określenie zakresu kompetencji (odpowiedzia-lności i decyzyjności), brak czytelnej struktury organizacyjnej, brak szkoleń, brak systemu motywowania i oceniania
	Planowanie projektu	Błędy w planowaniu (podział na zadania, pracochłonność zadań, harmonogram projektu, zależności między zadaniami, dostępność zaso-bów i środków), błędna identyfikacja zagrożeń dla projektu, dopasowanie „na siłę" do oczekiwań organizacji, nieokreślenie zasad komunikacji w zespole,
	Zarządzanie projektem	Brak nadzoru nad postępem prac, duża zmienność wymagań, omijanie procedur, brak zarządzania zmianami, brak zarządzania ryzykiem projektowym, brak osób odpowiedzialnych, sztywne trzymanie się przyjętych planów, problemy w zespole projektowym, problemy z poddostawcami
	Zamknięcie projektu	Nieweryfikowanie spełnienia kryte-riów odbioru, brak dokumentacji, przekazanie do eksploatacji bez nale-żytego przeszkolenia użytkowników
	Eksploatacja systemu	Niska jakość oprogramowania, słaba infrastruktura teleinformatyczna (dostępność, niezawodność, bezpie-czeństwo danych, itd.), złe wyszko-lenie użytkowników, brak efektyw-ności wykorzystania (działania wspierające)
Wykorzystanie	Utrzymanie (ang. maintenance) systemu (doskonalenie, naprawianie, dostosowanie, zapobieganie)	Niska jakość oprogramowania, duża szybkość i skala zmian w otoczeniu, duża szybkość i skala zmian wyma-gań, słabe administrowanie systemem (zespół ludzki, procedury, itd.)

Kolejnym istotnym aspektem związanym z przedstawionymi czynnikami ryzyka jest kwestia horyzontu czasowego ich oddziaływania. Działania w fazie przygotowania dotyczą fundamentalnych elementów całego przedsięwzięcia, tym samym zasięg ich oddziaływania obejmuje cały czas trwania przedsięwzięcia. Oznacza to, że występujące w tej fazie ryzyko, które będziemy dalej nazywać ryzykiem inwesty-cyjnym, a dotyczące głównie błędnego zdefiniowania potrzeb i możliwości ich zaspo-kojenia, nie kończy się wraz z zakończeniem fazy przygotowania, lecz rozciąga się na cały czas trwania przedsięwzięcia, łącznie z fazą eksploatacji – gdyż może się tak zdarzyć, że dopiero w fazie eksploatacji wystąpią negatywne skutki decyzji podjętych w fazie przygotowania. Z kolei, większość czynników ryzyka projektowego, związa-nego z fazą wykonania, w sposób naturalny przestaje oddziaływać bezpośrednio na przedsięwzięcie w momencie zakończenia projektu. Niemniej jednak, ryzyko projek-towe oddziałuje pośrednio na ryzyko w fazie wykorzystania (ryzyko eksploatacyjne). Dzieje się tak, gdyż ryzyko eksploatacyjne jest w znacznym stopniu związane z jakością produktów końcowych projektu (oprogramowania) oraz jakością wdrożenia tych produktów. Ogólny schemat tych zależności przedstawia rysunek 2.



Rys. 2. Rodzaje ryzyka w poszczególnych fazach przedsięwzięcia informatycznego

Aby zilustrować te zależności rozpatrzmy przypadek przedstawiony na rysunku 1. Jako sposób na realizację celu „*wzmocnienie pozycji rynkowej*" wybrano „*budowę nowych kanałów dystrybucji*" poprzez „*budowę portalu internetowego ze zintegro-wanym sklepem internetowym*". Działanie w ramach każdej z faz mogą być przyczyną nieosiągnięcia zakładanych celów. I tak, może się okazać, że grupa docelowa odbior-ców danego produktu nie dokonuje zakupów przez Internet (brak wiedzy i błędy analizy w fazie przygotowania), a w konsekwencji nawet perfekcyjna realizacja pozo-stałych faz nie pozwoli osiągnąć zakładanych celów. Może też się okazać, że stwo-rzone oprogramowanie zawiera wiele błędów i usterek, co w konsekwencji zniechęca potencjalnych klientów do

korzystania z niego (niska jakość oprogramowania) lub, że w fazie wykorzystania nie podjęto działań mających wypromować świadomość istnienie takiej możliwości zakupów u potencjalnych klientów (brak działań wspierających – marketingu i reklamy). Jak widać, w każdej z faz możliwe jest popełnienie błędów mogących przekreślić możliwość osiągnięcia zakładanych celów.

Sukces przedsięwzięcia jest wypadkową działań podejmowanych we wszystkich jego fazach. Tym samym, ryzyko występujące w każdej z faz składa się na ryzyko całego przedsięwzięcia. Oznacza to, że ryzyko przedsięwzięć informatycznych powinniśmy rozpatrywać kompleksowo z uwzględnieniem zależności pomiędzy poszczególnymi rodzajami ryzyka występującymi w różnych fazach tych przedsięwzięć.

3. Działanie w warunkach niepewności - zarządzanie ryzykiem

Ryzyko jest rezultatem działań podejmowanych przez człowieka w warunkach braku pełnej wiedzy na temat ich skutków. Złożoność otoczenia oraz zachodzących w nim procesów sprawia, że osiągnięcie pełnej wiedzy o tych skutkach jest niemożliwe, a tym samym człowiek musi działać w warunkach występującego ryzyka, niezależnie od tego czy jest tego świadomy czy nie. Uświadomienie sobie faktu istnienia ryzyka, jego identyfikacja oraz określenie konsekwencji stanowi punkt startowy do świado-mego wyboru sposobu działania z uwzględnieniem poziomu ryzyka i jego akceptacją, czyli liczeniem się z jego potencjalnymi konsekwencjami. Właśnie to założenie stanowi podstawę zarządzania ryzykiem opisanego poniżej.

3.1. Zarządzanie ryzykiem – koncepcja i procesy

Niewątpliwie celem zarządzania ryzykiem jest świadome kształtowanie zarówno poziomu podejmowanego ryzyka jak i ograniczanie (minimalizacja) jego potencja-lnych skutków. Na potrzeby definicji zarządzania ryzykiem, T.T.Kaczmarek przyj-muje w odniesieniu do gospodarki, że „*ryzyko to zagrożenia dla osób, rzeczy i interesów przedsiębiorstwa w następstwie prowadzonej działalności, istniejących zależności oraz zdarzeń*". Na tej podstawie definiuje on zarządzanie ryzykiem jako „*logicznie uporządkowany zbiór reguł i zasad, w sposób jednolity i stały stosowanych w odniesieniu do ryzyka prowadzonej działalności*" [9], kładąc tym samym nacisk na aspekt procesowy zarządzania ryzykiem. Z kolei A.Stabryła definiując zarządzanie ryzykiem w odniesieniu do projektów, cytuje Y.Y.Chonga i E.M. Browna [4], którzy opisują je w następujący sposób: „*Analizę ryzyka można porównać do kreślenia mapy potencjalnych zagrożeń oraz szacowania szkód mogących być ich wynikiem. Zarządzanie ryzykiem to wykorzystanie tej mapy i podejmowanie*

decyzji, jak uniknąć wspomnianych zagrożeń" [12]. Obie te definicje wskazują wyraźnie, że pojęcie zarzą-dzania ryzykiem jest pewnym skrótem myślowym, który należy rozumieć jako podejmowanie działań z uwzględnieniem istniejących i potencjalnych zagrożeń, tak, aby tych zagrożeń uniknąć lub zminimalizować ich skutki.

Powyższe definicje mają charakter opisowy, wskazujący na istotę zarządzania ryzykiem. W praktyce, przekładają się one na różne metodyki zarządzania ryzykiem, oparte na wyodrębnieniu etapów działania oraz wskazaniu metod i technik wykorzystywanych na każdym z etapów. Liczba etapów oraz ich zakres zależy od konkretnej metodyki. Dla przykładu, PMI wyróżnia sześć etapów działania. Są to: planowanie zarządzania ryzykiem, identyfikacja, analiza jakościowa, analiza ilościowa, planowanie środków przeciwdziałania oraz monitoring i kontrola. Z kolei P.Buła wskazuje sześć obszarów działania, pogrupowanych w cztery etapy: wybór obszaru (etap 1), identyfikacja, ustalenie indykatorów i budowa katalogu (etap 2), ocena ryzyka (etap 3) i manipulowanie ryzykiem (etap 4) [2]. Również T.T.Kaczmarek wyróżnia cztery etapy: prognozowanie zagrożeń, identyfikacja ryzyka, opracowanie strategii zarządzania oraz controlling ryzyka [9]. Natomiast A.Stabryła w swojej metodyce proponuje sześć faz: identyfikacja, analiza, sformułowanie wariantów, ocena, sterowanie oraz kontrola, monitoring i ocena podjętych działań [12].

Wspólnym elementem tych metodyk jest koncentrowanie się wokół czterech podstawowych rodzajów aktywności: identyfikacji, oceny, planowania i kontroli. Identyfikacja obejmuje określenie źródeł potencjalnego zagrożenia oraz wskazanie istniejących zależności pomiędzy nimi, a elementami przedsięwzięcia. Z kolei, ocena zawiera określenie prawdopodobieństwa wystąpienia danego czynnika ryzyka oraz związanych z nim negatywnych skutków. W ramach oceny powstaje również tzw. system wczesnego ostrzegania [4], który w przyszłości poprzez stałe mierzenie poziomu ryzyka ma za zadanie uprzedzać o wystąpienie zdarzeń wpływających na ryzyko. Uzyskane wyniki są następnie wykorzystywane do zaplanowania konkretnych działań w zakresie obniżenia poziomu ryzyka. Do tych działań zaliczymy takie działania jak unikanie ryzyka np. poprzez wybór rozwiązań alternatywnych czy kompensację, czyli redukcję lub eliminację skutków. Istotnym elementem jest stworzenie planów awaryjnych dla najbardziej prawdopodobnych lub niosących ze sobą największe zagrożenie czynników. Ostatnim elementem jest kontrola, która obejmuje stały monitoring poziomu ryzyka, uruchamianie planów awaryjnych oraz stałe śledzenie i ocena skutków podejmowanych działań.

Ryzyko i zarządzanie ryzykiem jest związane bezpośrednio z podejmowaniem decyzji. W konsekwencji musi pozostać w rękach osób podejmujących te decyzje. Podkreśla to T.T.Kaczmarek, który określa to następująco „*Nie może ona (strategia podejścia do ryzyka) być wydzielonym zadaniem dla managera do spraw ryzyka, ale powinna być wpisana w zakresy obowiązków managerów operatywnych działów. Stanowi ona integralną część procesu planowania,*

kierowania i raportowania" ([9], str. 96). Z drugiej strony ten sam autor podkreśla, że „*...zarządzanie ryzykiem wymaga skoordynowanej współpracy specjalistów z wielu dziedzin, ... Ponadto potrzebny jest manager, który w sposób całościowy będzie umiał zarządzać ryzykiem przedsiębiorstwa.*" ([9], str.109). Te dwa pozornie sprzeczne zdania pokazują trudność zarządzania ryzykiem. Z jednej strony musi się ono odbywać na każdym poziomie, na którym podejmowane są decyzje, a z drugiej strony musi być skoordynowane na poziomie całości działań. Takie podejście jest stosunkowo łatwe do wdrożenia w przypadku niektórych rodzajów ryzyka jak ryzyko kredytowe czy ubezpieczeniowe, gdzie istnieją formalne metody i techniki zarządzania ryzykiem, które mają ugruntowaną pozycję i stosunkowo łatwo podlegają standaryzacji. Pozwala to na stworzenie standardów organizacyjnych w zakresie zarządzania ryzykiem, w ramach których zawarte są wytyczne i procedury dla menagerów na różnym poziomie szczebla decyzyjnego. W przypadku przedsięwzięć informatycznych sprawa jest bardziej skomplikowana, szczególnie ze względu na ich niepowtarzalność oraz zmiany w składzie osób decyzyjnych w poszczególnych fazach przedsięwzięcia.

3.2. Zarządzanie ryzykiem w przedsięwzięciach informatycznych – aspekt osobowy

Odnosząc koncepcję zarządzania ryzykiem do przedsięwzięć informatycznych należy wyjść od rodzajów ryzyka, jakie mogą wystąpić. W ramach przedsięwzięć informatycznych wyróżniono wcześnie trzy rodzaje ryzyka powiązane z fazą ich powstawania, a tym samym z działaniami (i decyzjami) podejmowanymi w ramach tych faz. Dla każdego z tych rodzajów ryzyka istnieją sprawdzone metody zarządzania ryzykiem[1], stale rozwijane w odpowiedzi na zmieniające się wymagania. Tym samym dysponujemy metodami, które mogą być wykorzystane w zarządzaniu ryzykiem w przedsięwzięciach informatycznych. W tym miejscu powstaje pytanie, czy te stworzone często niezależnie metody tworzą jednolity i spójny system zarządzania ryzykiem?

Pytanie to, jest tym bardziej ważne, że jak pokazano na rysunku 2, poszczególne rodzaje ryzyka występują równocześnie oraz są od siebie zależne. Aby odpowiedzieć na to pytanie, trzeba wrócić do poruszonej wcześniej kwestii kompetencji w zakresie zarządzania ryzykiem.

Wymaganie, aby zarządzanie ryzykiem pozostawało w gestii osób podejmujących decyzje, sprawia, że punktem wyjścia do określenia kompetencji są osoby bezpoś-rednio zaangażowane w przedsięwzięcie lub wpływające na jego przebieg. W tabeli 2 przedstawiono skład zespołów zaangażowanych w

[1] Przedstawienie tych metod wykracza poza ramy niniejszych rozważań. Wykaz wybranych opracowań obejmujących zarządzanie ryzykiem w poszczególnych fazach przedsięwzięcia jest zawarty w literaturze uzupełniającej.

realizację poszczególnych faz w odniesieniu do hierarchii celów przedsięwzięcia (patrz również rysunek 1).

Tab. 2. Zaangażowanie osobowe w poszczególnych fazach [6]

Fazy	Cele	Zaangażowane zasoby osobowe
Przygotowania	Cele strategiczne Cele operacyjne Cel projektu	Kierownictwo wysokiego i średniego szczebla, analitycy biznesowi z ewentualnym wsparcie zewnętrznych firm konsultingowych
Wykonania	(rozwiązanie) Cele projektowe:	Sponsor, Komitet Sterujący, Kierownik Projektu, Kierownik Jakości, Zespół projektowy
	Procesowe	Kierownik Projektu, Zespół projektowy
	Produktowe	Kierownik Projektu, Kierownik Jakości, Zespół projektowy
Wykorzystania	Cele operacyjne	Użytkownicy, Dział Informatyki, Kierownictwo średniego szczebla

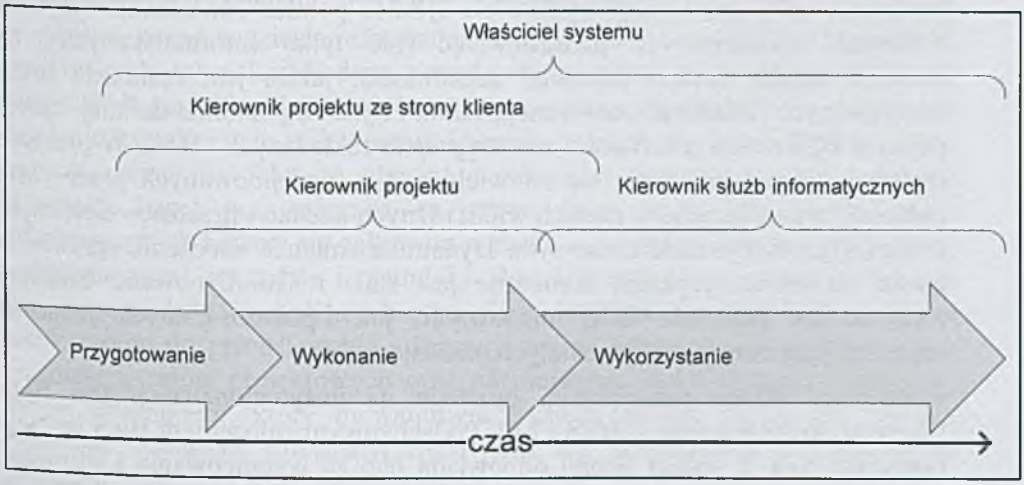
Wspomniana wcześniej zasada „*lokalizacji zarządzania ryzykiem na poziomie menadżerów operacyjnych*" oznacza, że każda z osób biorących udział w przedsię-wzięciu, która posiada możliwości decyzyjne powinna realizować na swoim poziomie zadania związane z zarządzaniem ryzykiem. Rodzi to kolejne pytanie. Kto nadzoruje i koordynuje te działania, czyli kto jest managerem ryzyka?

Stosunkowo łatwo wskazać managera ryzyka dla fazy wykonania. Zgodnie z więk-szością metodyk zarządzania projektami rola ta przypada Kierownikowi projektu, który odpowiada za całość prac prowadzonych w ramach projektu informatycznego. Znacznie bardziej skomplikowana sytuacja ma miejsce w przypadku pozostałych faz.

W przypadku fazy przygotowania, analizując skład osób zaangażowanych w jej przeprowadzenie trudno wskazać pojedynczą osobę odpowiadającą za zarządzanie ryzykiem. W części metodyk projektowych, faza przygotowania jest włączana do projektu informatycznego, tym samym wskazując osobę Kierownika projektu, jako odpowiedzialnego za ryzyko. Ale w praktyce, biorąc pod uwagę, że projekt jest realizowany przez wykonawcę zewnętrznego, Kierownik projektu jest dostępny dopiero po wybraniu wykonawcy, co jest jednym z końcowych elementów tej fazy. Pewnym rozwiązaniem tego problemu jest, spotykane w praktyce, powołanie Kierownika projektu ze strony klienta. Osoba ta odpowiada za przeprowadzenie końcowych działań fazy przygotowania i pełni rolę partnera oraz współpracownika Kierownika projektu w fazie wykonania. Stanowi ona również naturalnego kandydata na menadżera ryzyka w fazie eksploatacji. Problemem jest kwestia kompetencji takiej osoby. Specyfika projektów informatycznych wymusza, aby z jednej strony była to osoba o znaczących kompetencjach w zakresie technologii informatycznych, a z drugiej strony posiadała wystarczającą wiedzę z obszaru merytorycznego, w którym dane przedsięwzięcie jest realizowane połączone ze znajomością

i zrozu-mieniem celów całego przedsięwzięcia. Kolejnym problemem jest jej umiejscowienie w strukturze decyzyjnej organizacji klienta oraz zasad współpracy w ramach projektu informatycznego z Kierownikiem projektu.

Osobne pytanie dotyczy roli Kierownika projektu ze strony klienta w fazie wykorzys-tania. Potencjalnie nieograniczony czas trwania tej fazy oraz natura podejmowanych w tej fazie działań sprawia, że większy nacisk jest kładziony na działania poza-informatyczne. Dlatego w niektórych koncepcjach pojawia się rola właściciela systemu lub głównego użytkownika, osoby reprezentującej użytkowników systemu informatycznego (rezultatu projektu), do obowiązków której należy dbanie o właściwe wykorzystanie tego systemu i kordy-nacja wszystkich działań związanych z syste-mem, w tym również prac związanych z jego utrzymaniem. Najczęściej osobą tą jest kierownik jednego z działów operacyjnych (merytorycznych). W fazie wykorzystania musimy dodatkowo brać pod uwagę kwestie zapewnienia funkcjonowania systemu informatycznego, która to kwestia leży najczęściej w gestii służb informatycznych klienta. Oznacz to, że w tej fazie jest dwóch naturalnych kandydatów na menadżerów ryzyka, są to właściciel systemu oraz kierownik służb informatycznych. Obydwoje odpowiadają za zarządzanie ryzykiem w swoim, dobrze określonym obszarze, aczkolwiek efekt końcowy zależy od ich sprawnego współdziałania. Zasięg czasowy możliwej aktywności każdej z wymienionych osób przedstawia rysunek 3.



Rys. 3. Ramy czasowe aktywności poszczególnych osób w odniesieniu do faz przedsięwzięcia informatycznego

Jak łatwo zauważyć, że wśród wymienionych osób, będący naturalnymi kandydatami na menadżerów ryzyka w różnych fazach przedsięwzięcia, brakuje osoby, której czas zaangażowania w przedsięwzięcie obejmowałby początkowe etapy fazy przygotowa-nia. Jest to związane ze specyfiką działań w tym okresie, obejmujących głównie kwestie definicji strategii organizacji i jej

operacjonalizacji. W zakresie decyzyjnym jest to domeną całego wyższego kierownictwa organizacji, stąd trudność we wskazaniu pojedynczej osoby mogącej odpowiadać wycinkowo za zarządzanie ryzykiem w kontekście przyszłego przedsięwzięcia. Na szczęście w zakresie definio-wania strategii, kwestie zarządzania ryzykiem są stawiane na jednym z najważniej-szych miejsc, co pozwala założyć, że są one realizowane w sposób należyty. Późniejsze działania w obszarze zarządzania ryzykiem powinny stanowić ich naturalną kontynuację.

3.3. Instytucjonalizacja zarządzania ryzykiem

Nierozstrzygniętą pozostaje kwestia koordynacji i współdziałania wymienionych wcześniej osób. Ich liczba, przy braku koordynacji i współpracy, sama w sobie staje się zagrożeniem dla realizacji przedsięwzięcia [6]. Kluczem do rozwiązania tego problemu jest istnienie jednolitej i spójnej polityki zarządzania ryzykiem na poziomie całej organizacji, obejmującej swoim zakresem wszystkie podejmowane działania, w tym realizowane przedsięwzięcia informatyczne. W takiej sytuacji, zarządzanie ryzykiem na poziomie projektu informatycznego stanowiłoby wypadkową polityki zarządzania ryzykiem klienta oraz wykonawcy.

Potrzeba wypracowania jednolitej polityki w zakresie zarządzania ryzykiem na pozio-mie całej organizacji wynika nie tylko z potrzeb zarządzania ryzykiem w ramach realizowanych przedsięwzięć (nie tylko informatycznych), ale również dotyka bardzo istotnego zagadnienia, jakim jest realizacja celów strategicznych. Jest to konsek-wencją faktu, że już na poziomie definicji ryzyka pojawia się kwestia możliwości nieosiągnięcia zakładanych celów. W praktyce, realizacja celów przekłada się na wiele zadań, podejmowanych przez różne jednostki organizacyjne w ramach wielu różnych działań i przedsięwzięć, często w rozległym horyzoncie czasowym. Dynamika zmian w otoczeniu sprawia, że chcąc zarządzać ryzykiem konieczne jest stałe i skoordynowane działanie, zarówno na poziomie całej organizacji, jak i poszcze-gólnych jednostek organizacyjnych czy podejmowanych przedsięwzięć.

Koncepcją takiego rozwiązania, opartego na instytucjonalizacji jest Biuro Nadzoru Strategicznego – BNS [7,8]. Podstawowym założeniem BNS jest jego podwójna rola. Z jednej strony odpowiada ono za wypracowanie i wdrożenie polityki w zakresie zarządzania ryzykiem, z drugiej strony pełni rolę pomocniczą dla menadżerów operacyjnych dostarczając im narzędzi i metod wspierających zarządzanie ryzykiem. Ważne jest, że BNS nie przejmuje kompetencji istniejących jednostek organizacyjnych czy instytucji projektowych, takich jak Komitet Sterujący, a jedynie koordynuje ich działania. Innym ważnym zadaniem BNS jest gromadzenie informacji na temat występujących ryzyk, ich wpływu na podejmowane działania oraz podejmowanych działań oraz ich skuteczności. Dzięki temu powstaje organizacyjne baza wiedzy w obszarze zarządzania ryzykiem.

Warto podkreślić, że podobne rozwiązanie, aczkolwiek związane z innymi rodzajami ryzyk niż ryzyko przedsięwzięć, są już powszechnie wykorzystywane. Dotyczy to przykładowo ryzyka kredytowego czy ubezpieczeniowego. Instytucje finansowe czy ubezpieczeniowe posiadają wyspecjalizowane jednostki organizacyjne odpowie-dzialne za wypracowanie i wdrożenie polityki w zakresie zarządzania ryzykiem danego rodzaju. Wydaje się, że zmiany, jakie zachodzą w koncepcji zarządzania organizacjami, a związane z przechodzeniem z funkcjonowania procesowego na projektowe, wymuszą powstanie podobnych rozwiązań dla innych rodzajów ryzyka.

4. Podsumowanie

Specyfika przedsięwzięć informatycznych tkwi w ich ścisłym zintegrowaniu z funkcjonowaniem organizacji. Technologia informatyczna dostarczając narzędzi wspomagających organizację w jej działaniu, jednocześnie wpływa na funkcjo-nowanie tej organizacji. W dobie silnej rynkowej konkurencji, samo posiadanie syste-mów informatycznych już nie wystarcza, gdyż posiadają je już praktycznie wszyscy [3]. O przewadze rynkowej stanowi obecnie efektywność wykorzystania tych syste-mów oraz umiejętność ich użycia do osiągnięcia celów organizacji.

Mimo wielu lat rozwoju inżynierii oprogramowania i metod zarządzania projektami, przedsięwzięcia informatyczne należą dalej do jednych z najbardziej ryzykownych. Jak pokazują dane, ponad połowa projektów informatycznych nie osiąga zakładanych celów w zaplanowanym czasie i budżecie[1]. Na podobnym poziomie kształtuje się zadowolenie klientów z uzyskanych rezultatów. Warto zwrócić uwagę na fakt, że chociaż poprawa w tym zakresie w ostatnich kilkunastu latach jest znacząca, to jednak dalej około 20% przedsięwzięć informatycznych kończy się całkowitą porażką, i to praktycznie bez względu na wykorzystywane narzędzie i metodyki. Rodzi to pytanie o przyczynę takiego stanu rzeczy. Jedną z możliwych odpowiedzi jest wskazanie na brak podejścia całościowego do przedsięwzięć informatycznych, integrującego kwestie celów organizacji, celów projektowych oraz późniejszego wykorzystania rezultatów. Innym czynnikiem, który niewątpliwie wpływa na ten obraz jest wysoka złożoność systemów informatycznych oraz ich koncepcyjny i abstrakcyjny charakter.

Warto zwrócić uwagę na interdyscyplinarny charakter przedsięwzięć informatycznych. Ukierunkowane na cel organizacji w sposób naturalny odnoszą się do obszaru jej działalności, integrując w sobie dodatkowo takie aspekty jak organizacja i zarządzanie, współpracę różnych podmiotów czy technologię

[1] Dane na podstawie raportów The Standish Group – organizacji specjalizującej się w gromadzeniu i analizie danych na temat efektywności projektów informatycznych (http://standishgroup.com/).

informatyczną, i to nie tylko w zakresie projektu informatycznego (faza wykonania), ale również pozostałych faz przedsięwzięcia. Oznacza to, że odnosząc podstawowe aspekty zarządzania ryzykiem do przedsięwzięć informatycznych musimy podchodzić do zagadnienia kompleksowo, pamiętając o specyfice poszczególnych faz.

LITERATURA

1. Bernstein P.,L.: Przeciw bogom – niezwykłe dzieje ryzyka. WIGG-Press, Warszawa 1997.
2. Buła P.: Zarządzanie ryzykiem w jednostkach gospodarczych. Aspekt uniwersalistyczny. Akademia Ekonomiczna, Kraków 2003.
3. Carr N.,G.: IT doesn't matter. Harvard Business Review. May, 2003.
4. Cebula P.: Systemy wczesnego ostrzegania w przedsiębiorstwie. Wydawnictwo Uniwersytetu Ekonomicznego, Kraków 2008.
5. Chong Y.,Y., Brown E.,M.: Zarządzanie ryzykiem projektu. Oficyna Ekonomiczna, Kraków 2001.
6. Dymek D.: Nieciągłość nadzoru w przedsięwzięciach informatycznych jako czynnik ryzyka. Zeszyty Naukowe Wyższej Szkoły Ekonomiczno-Społecznej w Ostrołece, nr 6, 2008, str. 55-60.
7. Dymek D.: Ciągłość realizacji celów w przedsięwzięciach opartych na technologii informatycznej. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie. Prace z zakresu informatyki. Przyjęte do druku.
8. Dymek D.: The Institutional Aspect of the Responsibility Distribution in the IT Enterprise. Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Application, pp. 100-104, Acta Press, Innsbruck, 2008.
9. Kaczmarek T.,T.: Ryzyko i zarządzanie ryzykiem – ujęcie interdyscyplinarne. Wyd.3, Difin, Warszawa 2008.
10. Kuraś M., Zając A.: Czynniki powodzenia i ryzyka projektów informatycznych. Zeszyty Naukowe nr 522, Akademia Ekonomiczna, Kraków, 1999.
11. Rogowski W., Michalczewski A.: Zarządzanie ryzykiem w przedsięwzięciach inwestycyjnych. Oficyna Ekonomiczna, Kraków 2005.
12. Stabryła A.: Zarządzanie projektami ekonomicznymi i organizacyjnymi. Wydawnictwo Naukowe PWN, Warszawa 2006.
13. Wysocki R.,K., McGary R.: Efektywne zarządzanie projektami, Helion, Gliwice 2005.

Literatura uzupełniająca

Alberts C.,J.: Common element of risk. Technical Note CMU/SEI-2006-TN-014, April, 2006.

Dymek D.: Risk Management in IT Based Enterprises. W Information Systems in Management III pod redakcją Karnowski W. i Orłowski A., Warsaw University of Life Science – SGGW Department of Informatics, pp.16-22, Warszawa 2009.

Jedynak P.,Szydło S.: Zarządzanie ryzykiem. Wydawnictwo Ossolineum, Wrocław 1997.

Jones R.: Zarządzanie projektami – sztuka przetrwania. MT Biznes, Warszawa 2009.

Kaczmarek T.,T.: Zarządzanie zdywersyfikowanym ryzykiem w świetle badań interdyscyplinarnych. Wydawnictwo WSZiM, Warszawa 2003.

Kendall R.: Zarządzanie ryzykiem dla manedżerów. Wydawnictwo K.E.Liber, Warszawa 2000.

Kerzner H.: Advanced project management – edycja polska. Helion, Gliwice 2005.

Lendzion J.,P., Stankiewicz-Mróz A.: Wprowadzenie do organizacji i zarządzania. Oficyna Ekonomiczna, Kraków 2005.

Pańkowska M.: Zarządzanie zasobami informatycznymi. Difin, Warszawa 2001.

Pawlak M.: Zarządzanie projektami. Wydawnictwo Naukowe PWN, Warszawa 2006.

Phillips J.: Zarządzanie projektami IT. Helion, Gliwice 2005.

Pritchard C.,L.: Zarządzanie ryzykiem w projektach. WIG-Press, Warszawa 2002.

Szyjewski Z.: Zarządzanie projektami informatycznymi. Agencja Wydawnicza Placet, Warszawa 2001.

Tyszka T., Zaleśkiewicz T.: Racjonalizacja decyzji. Polskie Wydawnictwo Ekonomiczne, Warszawa 2001.

Wiliams. R., Ambrowe K., Centrem L.: A roadmap of risk diagnostic methods: developing an integrated view of risk identification and analysis techniques. Technical Note CMU/SEI-2004-TN-002, September 2004.

Wiliams R.C., Ambrowe K., Bentrem L., Merendino T.: Risk based diagnostics. Technical Note CMU/SEI-2004-TN-013, September 2004.

Zachorowska A.: Ryzyko działalności inwestycyjnej przedsiębiorstw. Polskie Wydawnictwa Ekonomiczne, Warszawa 2006.

Rozdział 11

Informatyzacja przedsiębiorstw – koszty oraz ocena efektów

Tomasz Lis
Politechnika Częstochowska
tomlis1@wp.pl

Marek Lis
Politechnika Częstochowska
lism@el.pcz.czest.pl

Streszczenie

W rozdziale podjęto problem informatyzacji przedsiębiorstw z perspektywy kosztów oraz efektów.

1. Wstęp

Osiąganie przez przedsiębiorstwa sukcesu rynkowego nie jest wyłącznie związane z prowadzeniem skutecznej polityki w zakresie ich zarządzania. Wdrożenie odpowiedniego oprogramowania pozwala na zwiększenie skuteczności szeregu działań zachodzących w łańcuchu dostaw, w dużym stopniu zwiększając jego przejrzystość. Systemy informatyczne, dzięki zdolności do szybkiego analizowania bardzo dużych ilości informacji, pozwalają na przeprowadzanie symulacji procesów dystrybucji i zaopatrzenia dając w ten sposób możliwość optymalizacji utrzymywanych zapasów i polepszenie poziomu obsługi klienta [1].

- Skuteczny system informatyczny jest doskonałym narzędziem wspomagającym działanie każdego przedsiębiorstwa, na wszystkich płaszczyznach jego zarządzania. Decyzje o jego wdrożeniu zaliczane są według J. Kisielnickiego oraz H. Sroki do najtrudniejszych zagadnień związanych z chęcią unowocześniania organizacji. Jak stwierdzają autorzy: „funkcjonowanie systemu informacyjnego dotyczy zagadnień przetwarzania informacji, natomiast wyznaczanie wartości informacji do dziś stanowi nierozwiązany problem. Wartość

informacji jest związana z podejmowaniem decyzji, a więc z prawdopodobieństwem zaistnienia różnych sytuacji decyzyjnych (...),

- inwestowanie w informatykę ma na celu unowocześnienie systemu zarządzania. Sprawność funkcjonowania takiego systemu zależy od wielu czynników, a szczególnie od kwalifikacji kadry kierowniczej, a więc jej przygotowania do posługiwania się informatyką,
- efektywność korzystania z informatyki zależy nie tylko od organizacji, ale też od sprawności otoczenia. Organizacja powiązana jest z innymi organizacjami wieloma kanałami. Sprawność powiązań, czyli tzw. interface wpływa na sprawność całego układu,
- użytkownicy systemu informacyjnego mają różne preferencje, które nie zawsze są zgodne, a często bywają konfliktowe. Problemem jest określenie wzajemnych relacji pomiędzy użytkownikami danego systemu [2].

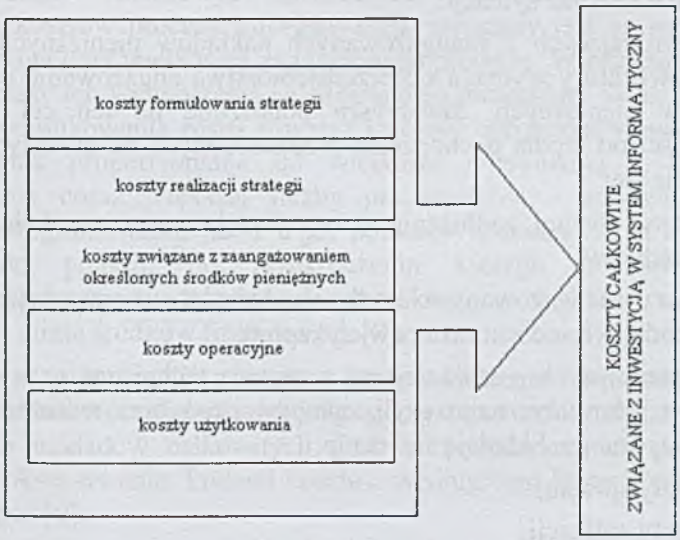
W celu zgodnego z zamierzeniami wykorzystania narzędzi informatycznych, w procesach zarządzania przedsiębiorstwem, niezbędne jest wybranie, najlepszego spośród oferowanych na rynku rozwiązania aplikacyjnego. Przystępując do zakupu systemu informatycznego należy zachować szczególną ostrożność. Wynika to z faktu, że inwestycja w systemy informatyczne ma dla organizacji znaczenie strategiczne. W związku z tym, przed przystąpieniem do jej realizacji należy przeprowadzić szereg badań oraz analiz mających określić szansę powodzenia przedsięwzięcia.

2. Koszty wdrażania narzędzi informatycznych w przedsiębiorstwach

Proces informatyzacji firmy może być związany z zakupem nowego sprzętu komputerowego (zakładamy, iż przedsiębiorstwo posiada już odpowiedni system informatyczny lecz dla poprawy wydajności jego działania potrzebne są nowe jednostki centralne – czasami wystarcza zakup nowego szybkiego serwera), zakupem systemu informatycznego (zakładamy, że w przedsiębiorstwie istnieje odpowiednia infrastruktura sprzętowa, a dezaktualizacji uległ funkcjonujący dotychczas system – może być to spowodowane: rozbudową przedsiębiorstwa, poszerzeniem zakresu działalności, itp.), zakupem zarówno sprzętu jak i oprogramowania (sytuacja spotykana najczęściej w przypadku zakładania firmy, znacznej jej modyfikacji, bądź zupełnego zdezaktualizowania zarówno możliwości obliczeniowych komputerów, jak i możliwości funkcjonalnych systemu informatycznego). Oczywiste jest, że każdy z procesów inwestycyjnych w obszarze informatyzacji związany jest z różnymi nakładami finansowymi. Skala tych nakładów zależna jest od wielkości firmy,

zakresu prowadzonej działalności, a także jakości sprzętu i usług w jakie inwestujemy.

Czas związany z wyborem systemu informatycznego jaki ma zostać wdrożony w organizacji nie może być zbyt długi gdyż funkcjonowanie według starych zasad wiąże się z występowaniem określonych strat. Ponieważ na rynku dostępnych jest wiele różnych systemów informatycznych, to dla wyboru rozwiązania najbardziej optymalnego, należy wcześniej określić szereg wymagań funkcjonalnych, stanowiących kryterium przydatności. Następnie przyrównując do nich właściwości poszczególnych systemów dokonuje się selekcji ofert. Czynności na tym etapie polegają na „porównywaniu wymagań użytkowników i wypracowanej w fazie reorganizacji nowej koncepcji działania biznesu z pożądaną funkcjonalnością systemu" [3]. Kolejnym krokiem jest porównanie kosztów zakupu, wdrożenia i użytkowania oraz formy i warunków wdrożenia. Po wybraniu rozwiązania najlepszego pod względem wymagań technicznych powinno się sprawdzić jeszcze jak dany system funkcjonuje w rzeczywistości. Do największych błędów kadry kierowniczej przedsiębiorstw decydujących się na inwestycje w systemy informatyczne jest pogląd, że całkowite koszty tej operacji ograniczają się do zakupu. W rzeczywistości są one znacznie większe i co najważniejsze towarzyszą całemu okresowi użytkowania [4].



Rys. 1. Koszty inwestycji w system informatyczny

Koszty procesu informatyzacji przedsiębiorstw są sumą [5,6,7]:

kosztów formułowania strategii informatyzacji – powstają one w związku z potencjałem ludzkim, który musi być wydelegowany do zaplanowania oraz kontroli przebiegu prac informatyzacyjnych. Najczęściej wywołuje je:

- konieczność uczestniczenia w procesie kadry kierowniczej, konsultantów oraz ekspertów,
- konieczność powołania odpowiedniego zespołu, którego zadaniem jest określenie pozycji funkcjonalnej przedsiębiorstwa, a także przeprowadzenie analizy sytuacyjnej systemu,
- przygotowywanie niezbędnych spotkań oraz szkoleń związanych z planowanym procesem oraz co z tym związane konieczność wydelegowania pracowników do ich realizacji.

kosztów realizacji przyjętej strategii – podobnie jak poprzednio, związane są z czynnikiem ludzkim. W tym przypadku jest to przede wszystkim konieczność angażowania do projektu odpowiednich ekspertów oraz konsultantów, najczęściej zatrudnionych w firmach zewnętrznych – działania outsourcingowe. Wynikiem mających miejsce na tym etapie czynności jest pojawienie się zmian na płaszczyźnie informacyjnej, technicznej oraz proceduralnej:

- opracowanie projektu modernizacyjno-rozwojowego,
- opracowanie działań modernizacyjno-rozwojowych,
- wdrożenie zaplanowanych i przygotowanych do realizacji działań,
- koszty związane z pojawianiem się niemożliwych do przewidzenia zdarzeń oraz sytuacji.

kosztów wynikających z zaangażowanych nakładów pieniężnych – każdemu procesowi, który wymaga od przedsiębiorstwa angażowania odpowiednich środków pieniężnych, towarzyszą ponoszone na ten cel wydatki. W zależności od źródła pochodzenia przeznaczonych na inwestycje pieniędzy wyróżnia się:

- koszty obsługi zadłużenia – organizacja zmuszona była do wzięcia kredytów,
- koszty zaangażowanych środków własnych – organizacja poświęcając środki własne zamraża pewien kapitał.

kosztów operacyjnych – związane są z oprawą techniczną oraz wykonawczą projektu informatycznego czyli zakupywanego oraz wdrażanego systemu informatycznego. Składają się na nie:

- koszty sprzętu,
- koszty aplikacji,
- koszty osobowe,
- koszty utrzymywania ciągłości pracy oraz koszty konserwacji,
- koszty lokalizacji.

kosztów użytkowania – w czasie użytkowania każdego sprzętu technicznego w tym także komputerów mogą pojawić się zakłócenia ich pracy. Wynikają one zazwyczaj z awarii, których źródłami są ludzie bądź też sam sprzęt.

Takim zdarzeniom podlegają również systemy informatyczne. Z tym, że w ich przypadku to czynnik ludzki jest najbardziej aktywnym źródłem powstających awarii. Z racji, często trudnych do przewidzenia efektów przestojów w pracy aplikacji komputerowych przewidywanie kosztów awarii jest procesem trudnym. Najczęściej dokonuje się szacunkowego wyznaczenia współczynnika awaryjności, a następnie porównania otrzymanej wartości do wyników tego samego współczynnika wyznaczonego dla porównywalnych pod względem jakości i przeznaczenia systemów informatycznych.

Do podstawowych kosztów awaryjności zalicza się:

- koszty naprawianego sprzętu,
- koszty oprogramowania,
- koszty prawne,
- koszty wydatkowane na osoby zajmujące się naprawą błędów,
- koszty związane z utratą potencjalnych zysków,
- koszty wynikające z odpływu klientów,

koszty związane z utratą i koniecznością odzyskiwania danych.

Poziom nakładów finansowych związanych z poszczególnymi elementami całkowitych kosztów procesu informatyzacji, związany jest ze stopniem jego skomplikowania oraz wielkością zamierzonego projektu. W przypadku prostych aplikacji, koszty ich wytworzenia są stosunkowo niewielkie. Wraz ze wzrostem poziomu skomplikowania rosną również i koszty oprogramowania. Zmiany te nie są jednak proporcjonalne do wielkości i wynikają z konieczności zaangażowania coraz większej liczby pracowników – zarówno w firmie tworzącej oprogramowanie jak i u jej odbiorcy. Poziom kosztów wzrastając dochodzi do punktu, po przekroczeniu którego lawinowo wzrasta prawdopodobieństwo niedokończenia projektu, co z kolei wiąże się z bezpowrotną utratą środków finansowych [8].

Kierownictwo organizacji planując przeprowadzenie procesu wdrożenia nowego, bądź unowocześnienia istniejącego systemu informatycznego, oprócz planowanych nakładów całkowitych, jest zainteresowane ich podziałem ze względu na okres trwania. Podział kosztów według tego kryterium przedstawia się następująco [9]:

1) jednorazowe:
   - zakup środków technicznych,
   - zatrudnienie dodatkowych pracowników – przede wszystkim: analityków, administratorów,
   - wykonanie (planowanie, realizacja, testowanie i użytkowanie) systemu informatycznego,

- przygotowanie pomieszczeń do instalacji w nich sprzętu komputerowego,
- instalacja oraz wdrożenie systemu informatycznego,
- przeszkolenie kadry – użytkowników systemu.

bieżące – związane z tokiem użytkowania oprogramowania, ich występowania nie da się uniknąć:

- nadzór nad wdrożonym oprogramowaniem oraz związana z tym konserwacja,
- amortyzacja infrastruktury komputerowej,
- dzierżawa łącz telekomunikacyjnych,
- transmisja danych,
- koszty wynikające z użytkowania i związane z zakupem materiałów eksploatacyjnych i części zamiennych.

Koszty użytkowania systemów informatycznych ujmuje się często w ramy dwóch głównych kategorii:

1) „koszty bezpośrednie (księgowane), na które z kolei składają się koszty:
   - sprzętu komputerowego i oprogramowania (wydatki kapitałowe związane z zakupem oraz dzierżawą nowych systemów informatycznych, sprzętu komputerowego, jego rozbudową i uaktualnianiem oprogramowania do nowszych wersji),
   - zarządzania (koszty związane z administracją sieci komputerowych, systemów informatycznych, archiwizacją danych, czy też usług outsourcingowych),
   - pomocy technicznej (opłaty związane z podnoszeniem umiejętności, szkolenia, wyjazdy służbowe, świadczenia za wsparcie techniczne, itp.),
   - projektowania (koszty związane z projektowaniem i rozwojem aplikacji, testowaniem oprogramowania, tworzeniem dokumentacji, przystosowaniem i utrzymaniem istniejących systemów informatycznych),
   - opłaty telekomunikacyjne (wydatki związane z dzierżawą łącz telekomunikacyjnych i dostępem do serwerów),
2) koszty pośrednie (nie księgowane), które dzielą się na:
   - koszty generowane przez użytkowników systemów informatycznych: wzajemne wsparcie, okazyjne douczanie, zarządzanie plikami i danymi, rozwój aplikacji, czynniki rozpraszające, satysfakcja użytkowników,

- koszty wynikające z czas przestojów systemu – w wyniku tego składnika kosztów pośrednich, obniżona zostaje produktywność organizacji, podstawowym ich wykładnikiem są utracone przez przedsiębiorstwo korzyści" [10].

Oprócz poniesienia przez organizację pewnych nakładów finansowych, procesowi informatyzacji towarzyszą efekty niewymierne, których bezpośrednie przedstawienie w formie pieniężnej jest w zasadzie niemożliwe. Mówi się wówczas o efektach niemierzalnych. Zalicza się do nich [11]:

1) powstanie chaosu organizacyjnego związanego ze zmianami w organizacji – jest to szczególnie widoczne w pierwszym okresie użytkowania systemu informatycznego,
2) zagrożenia związane z podłączeniem firmowej sieci komputerowej, do sieci ogólnodostępnej, np. Internetu,
3) obserwowane zmniejszenie się panujących w organizacji więzi międzyludzkich.
4) Każdy proces inwestycyjny ma doprowadzać do otrzymania zamierzonych efektów. Ze względów finansowych wyróżnia się efekty wymierne i niewymierne. Taki podział obowiązuje również w dziedzinie inwestycji w szeroko rozumiane technologie informacyjne, w tym, w systemy informatyczne. Efekty niewymierne mogą być pozytywne, bądź negatywne - mówi się wówczas o stratach, ofiarach czy poświęceniu, wymierne z kolei, rozważa się zazwyczaj ze względu na ich zyskowność i rentowność.

3. Ocena opłacalności oraz efektywności procesu informatyzacji

W celu wyeliminowania projektów niepraktycznych i nierentownych opracowane zostały różne metody oceny inwestycji. Ich wyniki mają dawać w miarę jednoznaczną odpowiedź na pytanie: czy określony projekt inwestycyjny powinien zostać zrealizowany czy też nie. Należy jednak zaznaczyć, że nie istnieje metoda relatywnie najlepsza i uniwersalna, a kryterium wyboru jednej z nich sprowadza się do przeanalizowania szeregu czynników pośrednich związanych
z projektem. W pracy pod redakcją J. K. Grabary oraz J. S. Nowaka wymienione zostały cztery podstawowe grupy metod oceny inwestycji w technologie
informacyjne [12]:

1) analizy finansowej - w metodach analizy finansowej bierze się pod uwagę tylko i wyłącznie efekty finansowe planowanej inwestycji. Ich celem jest w miarę jednoznaczne określenie opłacalności projektu, a

cechą charakterystyczną zwracanie szczególnej uwagi na przepływy pieniężne. Wiele projektów inwestycyjnych nie może być ocenionych przy pomocy tych metod, ciężko jest na przykład stwierdzić za ich pomocą czy uzyskane zostaną pożądane efekty strategiczne - a tego właśnie typu oczekiwania związane są z inwestycjami informatycznymi. Z tego też powodu często spotyka się rozwiązania poszerzane o pewne dodatkowe kryteria. Mówi się wówczas o rozwiązaniach hybrydowych. Do metod analizy finansowej zaliczyć możemy metodę: wewnętrznej stopy zwrotu, bieżącej wartości netto, wskaźnik rentowności inwestycji, zrównoważonej karty wyników - metoda hybrydowa.

2) podejście wielokryterialne - w przeciwieństwie do poprzedniej klasy metod
w podejściu kryterialnym do oceny projektów inwestycyjnych stosuje się szereg różnych wskaźników. Oprócz finansowych, uwzględnia się tu kryteria biznesowe, a także technologiczne. Biznesowe - kompatybilność
z prowadzoną strategią firmy, niepewność organizacyjna, technologiczne – zgodność proponowanych w projekcie rozwiązań z istniejącą już infrastrukturą informatyczną, sposobu i jakości określanych wymagań technicznych, itp.. Przed przystąpieniem do oceny przedsięwzięcia inwestycyjnego należy określić oraz wybrać wskaźniki, których obliczenie w jak największym stopniu przedstawi szanse badanego projektu,

3) analizy wskaźnikowej – u podstaw metody analizy finansowej leży koncepcja wskaźnika opłacalności zarządzania. Już od dawna zdawano sobie sprawę, że sukces powinien być mierzony przy użyciu wskaźnika wyrażającego stopień opłacalności -rentowności procesów. Przy badaniu opłacalności, korzystano z miary przedstawiającej liczbę zatrudnionych, skalę zaangażowanego kapitału. Rentowność była również badana na podstawie porównywania danych z innymi podobnymi przedsięwzięciami. Porównywanie to jest popularne również obecnie. Istnieje tworzona od wielu lat baza danych, zawierająca szereg informacji na temat wcześniejszych projektów i procesów inwestycyjnych. Przy realizacji projektu dzięki skonfrontowaniu planowanych wyników i panujących realiów, z informacjami zawartymi w tej bazie, możemy stwierdzić czy ma on szanse na odniesienie sukcesu i co należy zrobić aby tak właśnie się stało. Bada się w ten sposób skutki podejmowania różnych decyzji – baza udostępniana jest komercyjnie.

4) analizy portfela inwestycji – metody analizy portfela inwestycji zostały stworzone w celu umożliwienia otrzymania odpowiedzi na podstawowe

pytania jakie zadawane są przed przystąpieniem do projektowania inwestycji. Do pytań tych zaliczyć możemy:

- czy w danym momencie i sytuacji, organizacja powinna inwestować w technologie informacyjne,
- na jaka skalę przeprowadzić informatyzację, jakie obszary działalności firmy skomputeryzować, a jakie pozostawić niezmienione,
- w jakie rozwiązania techniczne i systemy informatyczne zainwestować.

W analizie portfela przyjmuje się, że przy inwestycji, bierze się pod uwagę tylko najlepsze z istniejących bądź planowanych warunków będących niezbędnymi dla osiągania zamierzonych celów przedsiębiorstwa.

Inny podział metod badania opłacalności inwestycji znaleźć można w pracy E. Nowaka, E. Pielichaty oraz M. Poszwy, którzy wyróżniają dwie podstawowe grupy [13]:

1) proste metody badania opłacalności – charakteryzują się one statycznym podejściem do badanego problemu, nie uwzględniając zmieniającej się w czasie wartości pieniądza. Przy sprawdzaniu rachunku opłacalności wykorzystywane są dane z jednego roku, bądź też przeciętne wartości nominalne z lat poprzednich – nie dyskontuje się ich, przez co uzyskiwane wyniki cechują się niska jakością. Metody te stosowane są w przypadku badania małych inwestycji, bądź też jako etap wstępny dużych projektów. Do prostych metod badania opłacalności autorzy zaliczają:
   - okres zwrotu,
   - stopa zwrotu,
   - próg rentowności.
2) dyskontowe metody badania opłacalności – pozwalają na przeprowadzenie analizy całego okresu trwania inwestycji. W przeciwieństwie do opisywanych wcześniej metod prostych uwzględnia się tu zależną od czasu i zmienna wartość pieniądza. „Dyskontowanie jest zabiegiem sprowadzającym nadwyżkę pieniężną z różnych lat do wartości bieżącej w roku bazowym, przez co uzyskuje się porównywalność w czasie" [13]. Wśród najczęściej używanych metod dyskontowych wymienia się:
   - wartość zaktualizowaną netto,
   - wewnętrzną stopę zwrotu,
   - wskaźnik rentowności inwestycji,
   - zdyskontowany okres zwrotu.

Rozważając zagadnienie kosztów oraz przyczyn i efektów inwestowania w systemy informatyczne dla przedsiębiorstw, nie można pominąć tematyki zakłóceń występujących przy procesie informatyzacji. Wyróżnia się dwie

główne formy zakupu systemu informatycznego. Pierwsza z nich opiera się na wyborze rozwiązania ogólnie dostępnego i uniwersalnego – znajdującego się w ofercie odpowiedniej firmy informatycznej. Druga forma wiąże się z zaangażowaniem firmy consultingowej, która przeprowadzając odpowiednie badania w przedsiębiorstwie opracowuje dokumentację zawierającą możliwości oraz przewidywaną specyfikację techniczną dla indywidualnego procesu informatyzacji. Na jej podstawie wyspecjalizowana firma zajmująca się tworzeniem oprogramowania opracowuje i wdraża system informatyczny. W pierwszym przypadku, jeśli zakupiony, gotowy system, nie posiada odpowiednich i zgodnych z wymaganiami możliwości, to efektem jego wdrożenia nie będzie usprawnienie funkcjonowania przedsiębiorstwa, a wręcz odwrotnie, mogą pojawić się zakłócenia i przestoje w jego pracy. Decydując się na informatyzację kompleksową – tworzenie oprogramowania od podstaw, według indywidualnych potrzeb – przed otrzymaniem zamierzonego efektu może pojawić się szereg przyczyn, z powodu których poniesione nakłady nie zostaną zwrócone, a poświęcony czas zostanie bezpowrotnie stracony. Przyczyny te mogą leżeć zarówno po stronie organizacji zamawiającej jak i dostarczającej system informatyczny. E. Yourdon wśród powodów niepowodzenia takiej formy informatyzacji wymienia [14]:

1) termin wykonania projektu został skrócony o ponad połowę w porównaniu
z wcześniejszymi ustaleniami – związane jest to zazwyczaj z chęcią wyprzedzenia konkurencji,
2) znacznie zmniejszona została liczba osób biorących udział w tworzeniu systemu – może być to skutkiem oszczędności, bądź niezrozumienia wymagań projektowych,
3) budżet projektu został zmniejszony do poziomu uniemożliwiającego otrzymanie zamierzonego efektu, bądź też końcowy twór nie posiada wielu niezbędnych komponentów, innym efektem takiej sytuacji jest zatrudnienie firmy tańszej, nie posiadającej jednak wymaganych kompetencji oraz umiejętności,
4) organizacja zamawiająca stawia zbyt wygórowane wymagania techniczne,
5) w trakcie tworzenia oprogramowania pojawiły się w firmie zamawiającej zmiany organizacyjne i nowe kierownictwo ma odmienne zdanie na temat informatyzacji,
6) każdemu nieudanemu procesowi informatyzacji przedsiębiorstwa towarzyszą wymierne oraz niewymierne straty. Zalicza się do nich: zmniejszenie ilości klientów, niedotrzymywanie terminów biznesowych, pozostawanie w tyle za konkurencją. Krańcowym efektem może być nawet upadek przedsiębiorstwa.

Wraz z rozwojem technologii informatycznej oraz coraz szerszym jej stosowania w podmiotach gospodarczych, powstawało wiele metod pozwalających na ocenę kosztów całkowitych towarzyszących procesowi informatyzacji. B. Pilawski wymienia wśród nich [15]:

1) metodyka TCO Chart of Accounts – jedna z najstarszych metod dostępnych na rynku, jej autorem była firma Gartner. Opiera się ona na podziale kosztów całkowitych na koszty bezpośrednie i pośrednie, a głównym wykorzystywanym w niej narzędziem są ankiety i wywiady. Ponieważ na otrzymywane wyniki mają bezpośrednio wpływ ankietowani, którzy celowo bądź nie mogą wpływać na manipulowanie otrzymywanymi wynikami. Z tego też powodu narzędzia te zalicza się do głównych słabości metody. Jej mankamentem jest również założenie że komputery w firmie są włączone na okrągło – co w rzeczywistości spotykane jest tylko sporadycznie,
2) metodyka TEI – w przeciwieństwie do poprzedniej (która opierała się na parametrze kosztów całkowitych TCO) metoda ta wykorzystuje pojęcie ROI – rachunek zwrotu z nakładów i zalecana jest jako uzupełniająca w stosunku do wcześniej zastosowanych w przedsiębiorstwie metod oceny kosztów – jest to metoda szacunkowa. Jej podstawowym celem jest wyznaczanie pełnego efektu gospodarczego przedsięwzięcia. Wykorzystuje się tu trzy kategorie: elastyczność, korzyści i koszty związane z ocenianym przedsięwzięciem, które obarczone są określonym elementem ryzyka. Uwzględnienie ryzyka ma zwiększyć poziom realności otrzymywanych wyników,
3) metodyka firmy Storactive Inc. – jest jedną z metod, w których koszty nakładów i użytkowania odnosi się do wskaźnika zwrotu z nakładów. Metoda ta, opiera się na pomiarze strat wynikających z utraty zgromadzonych informacji. Ponieważ koszty z tym związane są różne w zależności od organizacji, to w metodyce przyjęto pewne odgórne założenia:
   - „czas braku dostępu do komputera na stanowisku należy liczyć podwójnie, gdyż czas stracony należy później, gdy dostęp do komputera zostanie przywrócony odrobić,
   - komputery przenośne są 40% droższe w eksploatacji od stacjonarnych,
   - okres stosowania komputerów przenośnych jest o połowę krótszy, niż stacjonarnych" [15].
4) metodyka Applied Information Economics firmy Hubbard Decision Research – jej podstawą jest wyznaczanie nakładów inwestycyjnych na informatykę oraz pogląd, że nawet dla wielkości niewymiernych można określić odpowiednie formuły ekonomiczne, które pozwolą na obliczenie ich wartości.

Z kolei J. Nowak metody i techniki oceny kosztów oraz efektów procesów informatyzacyjnych dzieli na [16]:

1) Return-On-Investment (ROI) – zwrot z inwestycji – metody te opierają się na szacowaniu formalnych inwestycji. Przykładem prostej metody jest payback – zapłata jej celem jest prognozowanie czasu w jakim nakłady na inwestycje zostaną odzyskane. Najczęściej stosowaną metodą z tej grupy jest IRR – wewnętrzny stosunek zwrotu. Cały projekt określa się w niej przy użyciu jednego parametru jakim jest wewnętrzna stopa zwrotu. „Metody ROI są używane przeważnie przez organizacje z surową dyscypliną finansową. Techniki są oficjalne, a kalkulacje zwykle wykonuje personel księgowy na podstawie informacji uzyskanych od osób pracujących nad danym projektem. Mimo, że IRR boryka się z problemem ryzyka sposobów ustalenia właściwej poprzeczki dla stopy zwrotu" [16]. Metod tych używa się do szacowania inwestycji prostych, które cechuje mała liczba elementów decydujących o niepewności,
2) Cost – benefit analysis (CBA) – analiza kosztu i korzyści – podejście to cechuje szacowanie finansowe każdego z elementów, który ma wpływ na procesy informatyzacji. Składowe, dla których jasne określenie ich wartości jest trudne do wykonania, oznaczane są pewną wartością, w oparciu o przyjęte pojęcie wartości. Takie określanie elementów jest główną słabością metody, gdyż kierownictwo często nie zgadza się z przypisywanymi miarami. Metoda jest stosowana w przypadkach, w których właściwe jest stosowanie metody ROI lecz koszty i korzyści są trudne do bezpośredniego wyznaczenia,
3) Multi-objective, Multi criteria methods (MOMC) – metody wielu celów I wielu kryteriów „Kierownictwo ma możliwość oszacowania względnych wartości różnych pożądanych wyników w kategoriach swoich preferencji: ma możliwość uszeregowania celów przez określenie wagi preferencji dla każdego celu. (...) Metody te najlepiej sprawdzają się tam, gdzie istnieje wiele potencjalnych celów, które mają służyć różnym mechanizmom lub ludziom w organizacji. Są one przydatne przede wszystkim na etapie, gdy decyduje się strategia. Są także użyteczne kiedy istnieje wiele alternatywnych projektów i trudno jest wybrać, bo każdy przynosi inny rezultat" [16],
4) Boundary values – wartości graniczne – istotą metod z tej grupy, jest określenie zależności nakładów całkowitych przeznaczanych na informatykę do innych połączonych wartości,
5) Return on management (ROM) – „metoda ta polega na wyrażeniu rezultatu wprowadzenia nowego systemu jako zmiany wartości dodanej przez management będąca rezultatem wprowadzenia nowego systemu. ROM jest definiowana jako wartość pozostała po potrąceniu z całkowitego dochodu kosztu i wartości dodanej przez każdy zasób, włączając w to kapitał, lecz wyłączając management i jego koszt" [16],

6) Information economics – informacja ekonomiczna – jest to jedna z metod z grupy technik koszt-korzyść, przy czym rozszerza podstawę jej funkcjonowania o trzy procesy:
   - o „połączenie wartości, które wyglądają na wynikające z działania najważniejszej zmiany obejmującej różne obszary działania,
   - o przyspieszenie wartości, która usiłuje zdefiniować wartość przyszłych systemów uzależnionych od wyników testu wprowadzonego systemu,
   - o wzbogacenie pracy, dostarczające oceny dodatkowej wartości podniesionych umiejętności i zrozumienia dla organizacji, które personel może uzyskać dzięki używaniu IT" [16].

4. Zakończenie

Wprowadzenie do firm techniki informatycznej w znacznym stopniu przyczyniło się do usprawnienia prac ewidencyjno biurowych - automatyzacji i związanego z nią przyspieszenia wykonywanych czynności. Jednak największa zaleta tego procesu nie jest związana bezpośrednio z samym sprzętem, a ze sposobem podejścia do metod zarządzania przedsiębiorstwem.

Koszty procesów informatyzacyjnych, na które składa się zakup sprzętu komputerowego, zakup i wdrożenie specjalnego oprogramowania - późniejsze dbanie o jego sprawność, a także przeszkolenia pracowników powodują, że szczególnie w małych i średnich przedsiębiorstwach instaluje się proste systemy, których możliwości nie odpowiadają w pełni istniejącym wymaganiom. Jedynie organizacje duże stać na wprowadzenie do ich struktur systemów informatycznych dobrej klasy, posiadających wszystkie niezbędne funkcje. Jednym z podstawowych błędów przy podejmowaniu decyzji o konieczności zakupu odpowiedniego oprogramowania jest niedoszacowanie projektu, czego efektem może być jego niedokończenie.

LITERATURA

1. Witkowski K., Fidali A.: Miejsce operatora logistycznego w procesie integracji łańcucha dostaw, W: Logistyka przedsiębiorstw w warunkach przemian, pod red. Jarosława Witkowskiego, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2002.
2. Kisielnicki J., Sroka H.: Systemy informacyjne biznesu, informatyka dla zarządzania, Wydawnictwo PLACET, Warszawa 1999.
3. Kubiak B. F., Korowicki A.: Modelowanie procesów biznesowych i dylematy wyboru zintegrowanego systemu informatycznego, W.: Informatyczne wspomaganie procesów logistycznych, pod redakcją Janusza K. Grabary, WNT, Warszawa 2004.

4. Logistyka dystrybucji. Praca zbiorowa pod red. Krzysztofa Rutkowskiego, Wydawnictwo Difin, Warszawa 2002.
5. Wyraz W.: Zastosowanie prostych metod oceny do pomiaru efektywności przedsięwzięć informatycznych, W.: Problemy systemów informacyjnych w zarządzaniu przedsiębiorstwem, Wydawnictwo Wydziału Zarządzania Politechniki Częstochowskiej, Częstochowa 2000.
6. Nowicki A.: Strategia doskonalenia systemu informacyjnego w zarządzaniu przedsiębiorstwem, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 1999.
7. Yourdon E.: Współczesna analiza strukturalna, WNT, Warszawa 1996.
8. Szyjewski Z.: Metodyki zarządzania projektami informatycznymi, Wydawnictwo PLACET, Warszawa 2004.
9. Nowicki A.: Informatyka dla ekonomistów, PWN, Wrocław 1998.
10. Piwowarski M.: Koncepcja minimalizacji całkowitych kosztów informatyzacji przedsiębiorstw (TCO), W: Efektywność zastosowań systemów informatycznych, praca zbiorowa pod redakcją Janusza K. Grabary i Jerzego S. Nowaka, WNT, Warszawa-Szczyrk 2001.
11. Wyraz W.: Zastosowanie prostych metod oceny do pomiaru efektywności przedsięwzięć informatycznych, W.: Problemy systemów informacyjnych w zarządzaniu przedsiębiorstwem, Wydawnictwo Wydziału Zarządzania Politechniki Częstochowskiej, Częstochowa 2000.
12. Efektywność zastosowań systemów informatycznych, pod redakcją Janusza K. Grabary i Jerzego S. Nowaka, Tom 1, WNT, Warszawa-Szczyrk 2002.
13. Nowak E., Pielichaty E., Poszwa M.: Rachunek opłacalności inwestowania, PWE, Warszawa 1999.
14. Yourdon E.: Marsz ku klęsce, poradnik dla projektanta systemów, WNT, Warszawa 2000.
15. Pilawski B.: Ile kosztuje informatyka?, W: Informatyka w gospodarce globalnej – problemy i metody, praca zbiorowa pod redakcją Jerzego Kisielnickiego, Janusza K. Grabary i Jerzego S. Nowaka, WNT, Warszawa-Szczyrk 2003.
16. Efektywność zastosowań systemów informatycznych, praca zbiorowa pod redakcją Janusza K. Grabary i Jerzego S. Nowaka, Tom II, WNT, Warszawa-Szczyrk 2001,
17. Nowicki A.: Informatyka dla ekonomistów, PWN, Wrocław 1998.
18. Informatyczne wspomaganie procesów logistycznych, pod redakcją Janusza K. Grabary, WNT, Warszawa 2004.
19. Informatyka w gospodarce globalnej – problemy i metody, praca zbiorowa pod redakcją Jerzego Kisielnickiego, Janusza K. Grabary i Jerzego S. Nowaka, WNT, Warszawa-Szczyrk 2003.

Rozdział 12

Analiza i ocena metod drugiej generacji wymiarowania funkcjonalnego systemów oprogramowania

Beata Czarnacka-Chrobot
Szkoła Główna Handlowa, Katedra Informatyki Gospodarczej
bczarn@sgh.waw.pl

Streszczenie

Inżynieria oprogramowania to dyscyplina, która nie może się pochwalić takim stopniem dojrzałości pod względem wymiarowania tego, co jest jej przedmiotem, jak inne dyscypliny inżynierskie, w których wymiarowanie ma wszak znaczenie podstawowe. Zasadniczym powodem tego stanu rzeczy jest brak jednoznacznej miary rozmiaru produktu programowego, co stanowi z kolei kluczową przyczynę problemów w obiektywnym i wiarygodnym szacowaniu tak podstawowych atrybutów przedsięwzięć zmierzających do realizacji systemów oprogramowania, jak pracochłonność, koszty i czas, czy atrybutów wydajnościowych. Brak jednoznacznej miary rozmiaru produktu programowego można zatem uznać za jeden z zasadniczych problemów tej dyscypliny wiedzy i życia, z którego wynikają kilkudziesięcioletnie poszukiwania skutecznych rozwiązań w tym obszarze. W ostatnim czasie zaowocowały one akceptacją przez międzynarodowe organizacje standaryzacyjne ISO i IEC koncepcji wymiarowania rozmiaru oprogramowania bazującej na jego funkcjonalności wraz z pięcioma opartymi na niej metodami. Dwie z tych metod są uznawane za metody drugiej generacji wymiarowania rozmiaru funkcjonalnego produktu programowego. Celem niniejszego rozdziału jest analiza i ocena zasadniczych cech takich metod oraz ich porównanie z metodami wymiarowania rozmiaru funkcjonalnego pierwszej generacji na przykładzie aktualnej wersji metody stworzonej przez organizację COSMIC (Common Software Measurement International Consortium).

1. Wprowadzenie

W definicji inżynierii oprogramowania przyjętej przez Institute of Electrical and Electronics Engineers (IEEE) czytamy, że jest ona *„zastosowaniem systematycznego, zdyscyplinowanego, kwantyfikowalnego podejścia do rozwoju, obsługi i utrzymania oprogramowania"* [9]. Podejście kwantyfikowalne oznacza wszak, że wymiarowanie powinno stanowić immanentną cechę tej dyscypliny wiedzy i życia. Tymczasem inżynieria oprogramowania nie może się pochwalić wysokim stopniem dojrzałości pod względem wymiarowania tego, co jest jej przedmiotem – atrybutów oprogramowania, w tym przede wszystkim jego rozmiaru. Stanowi to zasadniczą przyczynę problemów w obiektywnym i wiarygodnym szacowaniu tak podstawowych atrybutów przedsięwzięć zmierzających do jego realizacji, jak pracochłonność, koszty i czas, a także atrybutów wydajnościowych (np. produktywność, stopa błędów). Brak jednoznacznej miary rozmiaru produktu programowego można zatem uznać za jeden z zasadniczych problemów tej dyscypliny.

Z tych między innymi powodów wynikają kilkudziesięcioletnie poszukiwania skutecznych rozwiązań w obszarze wymiarowania rozmiaru produktów programowych. Właśnie w ostatnim czasie zaowocowały one akceptacją przez międzynarodowe organizacje standaryzacyjne ISO (International Organization for Standardization) i IEC (International Electrotechnical Commission) jednej z koncepcji jego wymiarowania wraz z kilkoma opartymi na niej metodami. Po wielu latach weryfikacji różnorodnych podejść okazało się, że na ich uznanie w pełni zasługuje jak dotąd jedynie koncepcja szacowania i pomiaru rozmiaru oprogramowania bazująca na jego funkcjonalności – atrybucie o priorytetowym znaczeniu dla użytkownika. Wymiarowanie w oparciu o tę koncepcję rozmiaru produktów programowych, a na tej podstawie kluczowych atrybutów przedsięwzięć zmierzających do ich rozwoju można więc określić mianem *wymiarowania funkcjonalnego*.

2. Metody wymiarowania rozmiaru funkcjonalnego oprogramowania

Zasady wymiarowania rozmiaru funkcjonalnego (ang. *Functional Size Measurement* - FSM) oprogramowania zostały znormalizowanie w sześcioczęściowym standardzie ISO/IEC 14143 [13]. Wszystkie jego części, podobnie jak niżej wymienione normy ISO/IEC dotyczące poszczególnych metod FSM, są zgodne z normą ISO/IEC 15939 [17], wyznaczającą ogólne procedury dla procesu wymiarowania oprogramowania w zgodzie z normą ISO/IEC 15288 [16], określającą z kolei procesy cyklu życia systemu. Na potrzeby takiego wymiarowania w standardzie ISO/IEC 14143 zaproponowano definicję rozmiaru funkcjonalnego (ang. *functional size*) jako *„rozmiaru oprogramowania otrzymanego przez ilościowe określenie wymagań*

funkcjonalnych użytkownika" [14, s. 2]. Przy czym przez wymagania funkcjonalne użytkownika (ang. *Functional User Requirements* - FUR), rozumie się z kolei *„podzbiór wymagań użytkownika opisujący to, co oprogramowanie ma robić, w kategoriach zadań i usług"* [14, s. 2]. Użytkownik został zaś zdefiniowany jako osoba lub rzecz (np. inne aplikacje, czujniki, sprzęt komputerowy), która komunikuje się lub wchodzi w interakcje z wymiarowanym oprogramowaniem [14, s. 3]. Elementarną jednostkę FUR zdefiniowaną i wykorzystywaną przez metodę FSM do celów pomiarowych określa się natomiast mianem podstawowego komponentu funkcjonalnego (ang. *Base Functional Component* – BFC) [14, s. 1].

Twórcą koncepcji wymiarowania funkcjonalnego oprogramowania był A. Albrecht, który ok. 30 lat temu zaproponował metodę *Function Point Analysis* (FPA), opartą na pomiarze i estymacji funkcjonalności wymaganej przez zleceniodawcę oraz jemu dostarczonej [1]. Stąd grupę metod, która powstała na skutek rozwoju tego podejścia określa się mianem metod wymiarowania rozmiaru funkcjonalnego (ang. *Functional Size Measurement Methods* – FSMM). Od tamtej pory opracowano wiele ich wariantów (ocenia się, że co najmniej 20), wśród których do metod uznanych przez ISO i IEC za zgodne z zasadami FSM opisanymi w normie ISO/IEC 14143 znajdują się obecnie:

- Metody, które bywają uznawane za metody FSM pierwszej generacji [6, s. 8], w tym:
  - metoda punktów funkcyjnych w wersji IFPUG (International Function Point Users Group) [19] - najpopularniejsza z technik wymiarowania rozmiaru funkcjonalnego, uznana przez ISO/IEC w części dotyczącej takiego rozmiaru;
  - metoda punktów funkcyjnych MkII (Mark II) rozwijana przez UKSMA (United Kingdom Software Metrics Association) [20] - technika popularna w Wielkiej Brytanii, oferująca wyższy poziom szczegółowości wymiarowania w zestawieniu z metodą IFPUG, również w części dotyczącej rozmiaru funkcjonalnego zaakceptowana przez ISO/IEC;
  - metoda punktów funkcyjnych NESMA (Netherlands Software Metrics Association) [21] – holenderska i uproszczona wersja metody punktów funkcyjnych IFPUG, także posiadająca częściową aprobatę ISO/IEC.
- Metody uznawane za metody FSM drugiej generacji [6, s. 8]:
  - metoda (pełnych) punktów funkcyjnych w wersji COSMIC (Common Software Measurement International Consortium) [18] - technika wymiarowania rozmiaru funkcjonalnego w całości zaakceptowana przez ISO/IEC jako międzynarodowy standard takiego wymiarowania;
  - metoda FSM w wersji zaproponowanej przez FiSMA (Finnish Software Measurement Association) [22], także w całości uznana przez ISO/IEC.

Dotychczasowe doświadczenia pokazują, że w celu uznania określonej metody FSM za standard ISO/IEC musi ona przejść weryfikację trwająca od 2 do 4 lat.

Pierwsze trzy z wymienionych powyżej metod zostały znormalizowane nie w całości proponowanej przez rozwijające je organizacje, a w części - jednak tej najistotniejszej, bo dotyczącej wymiarowania rozmiaru funkcjonalnego oprogramowania. W podejściach proponowanych przez IFPUG, UKSMA i NESMA metody te obejmują także wyznaczenie tzw. czynnika korygującego, który ma za zadanie dostosowanie rozmiaru funkcjonalnego do środowiska konkretnego przedsięwzięcia poprzez uwzględnienie wpływu wymagań technicznych i jakościowych na proces projektowania i implementacji aplikacji. Jednakże ta część tych metod nie posiada aprobaty ISO i IEC - przyjęte przez te organizacje założenia wykluczają bowiem zależność FSM od tego typu wymagań. Metody COSMIC i FiSMA zostały natomiast uznane za międzynarodowe standardy w całości, dlatego bywają one uważane za metody FSM drugiej generacji.

Metody FSM posiadające akceptację ISO/IEC różnią się możliwościami wymiarowania oprogramowania w odniesieniu do różnych domen funkcjonalnych (kategorii oprogramowania). Dlatego przed wyborem określonej metody należy w pierwszym rzędzie ocenić jej adekwatność wobec rodzaju produktu, którego rozmiar funkcjonalny ma podlegać wymiarowaniu. I tak w normach opisujących powyżej wymienione metody czytamy, iż [15, s. 5, 18-20]:

- Nie istnieją ograniczenia co do domen funkcjonalnych dla zaakceptowanej części metody IFPUG i NESMA, podobnie jak dla metody FiSMA.
- Metoda UKSMA jest przeznaczona dla dowolnego typu oprogramowania, ale pod warunkiem, że można w nim zidentyfikować tzw. logiczne transakcje (rodzaj BFC w tej metodzie). Reguły metody nie wspierają kompleksowych algorytmów charakterystycznych dla oprogramowania naukowego i inżynieryjnego ani systemów czasu rzeczywistego, jako że opracowano je z myślą o systemach biznesowych.
- Metoda COSMIC sprawdza się dla systemów wspomagających zarządzanie (sterowanych danymi), systemów czasu rzeczywistego (sterowanych czasem) oraz dla rozwiązań hybrydowych łączących oba ww. typy oprogramowania. Natomiast jej zastosowanie nie daje poprawnych wyników w odniesieniu do systemów o złożonych algorytmach matematycznych lub innych wyspecjalizowanych i kompleksowych zasadach funkcjonowania oraz systemów przetwarzających zmienne ciągłe. Jednakże dla ww. domen możliwe jest zmodyfikowanie metody w celu jej lokalnego wykorzystania (tzw. lokalna kastomizacja).

Wymagania wobec właściwej metody FSM są różne w zależności od charakteru organizacji. Przykładowo, instytucje finansowe zwykle wybierają metodę, która w sposób prawidłowy wymiaruje aplikacje biznesowe, natomiast przedsiębiorstwo chemiczne w związku z jego działalnością podstawową

wymaga raczej takiej metody wymiarowania, która jest odpowiednia dla domen funkcjonalnych określanych jako systemy czasu rzeczywistego. Dlatego wybór właściwej metody należy rozpocząć od podziału oprogramowania organizacyjnego na domeny funkcjonalne. Wybór adekwatnej do potrzeb metody zależy także od planowanego sposobu wykorzystania jej rezultatu. Jeżeli przedsiębiorstwo czy instytucja zamierza wykorzystać rezultaty wymiarowania również w celu porównania swojej produktywności z danymi branżowymi, zaleca się wybór metody stosunkowo popularnej w danej branży, dla której istnieją takie dane. W przypadku, gdy potrzebuje jedynie pobieżnego, orientacyjnego oszacowania rozmiaru funkcjonalnego, wymagania wobec odpowiedniej metody jego wymiarowania będą zredukowane.

3. Rozwój i struktura metody COSMIC

W rezultacie kontrowersji dotyczących możliwości wykorzystania różnych powstałych do połowy lat 90. ubiegłego wieku technik wywodzących się z metody zaproponowanej przez A. Albrechta w odniesieniu do wymiarowania innego rodzaju oprogramowania niż aplikacje biznesowe w 1997 r. na University of Quebec w Montrealu opracowano koncepcję metody pełnych punktów funkcyjnych (ang. *Full Function Points* - FFP) [25]. Jej celem było rozszerzenie możliwości zastosowania koncepcji wymiarowania rozmiaru funkcjonalnego oprogramowania na te jego kategorie, których odbiorcą funkcjonalności (użytkownikiem) *nie jest człowiek* lub *nie jest nim tylko człowiek*. Zauważono bowiem, iż ówczesne podejścia stanowią w praktyce skuteczne metody wymiarowania dla produktu programowego powstającego w celu wspomagania operacji informacyjnych, gospodarczych oraz administracyjnych, czyli ogólnie rzecz ujmując – systemów wspomagających procesy zarządzania (sterowanych danymi). "*Problem z podejściem opartym na punktach funkcyjnych jest taki, że ocenia ono ograniczoną grupę rodzajów aplikacji: zwykle duże (...) systemy, realizowane w organizacjach takich, jak banki (....) i instytucje handlu detalicznego, nie radząc sobie z systemami hybrydowymi (...) ze znacznym komponentem komunikacyjnym*" [10]. Odbiorcą funkcjonalności systemów sterowanych danymi jest bowiem zwykle człowiek. Istnieją jednak i inne kategorie oprogramowania, których odbiorcą funkcjonalności nie jest człowiek, a urządzenia i/lub inne oprogramowanie, w tym przede wszystkim systemy czasu rzeczywistego (sterowane czasem) - ich funkcjonalność także powinna podlegać wymiarowaniu. Kategorie te różnią się znacznie od systemów wspomagających zarządzanie, przede wszystkim właśnie odbiorcą funkcjonalności.

W związku z powyższym punktem wyjścia dla stworzenia nowej metody FSM stało się rozszerzenie definicji użytkownika wykorzystywanej wtedy przez IFPUG: "*Użytkownikiem mogą być nie tylko osoby (specyfikujące FUR), ale też oprogramowanie i urządzenia mechaniczne wchodzące w interakcje z*

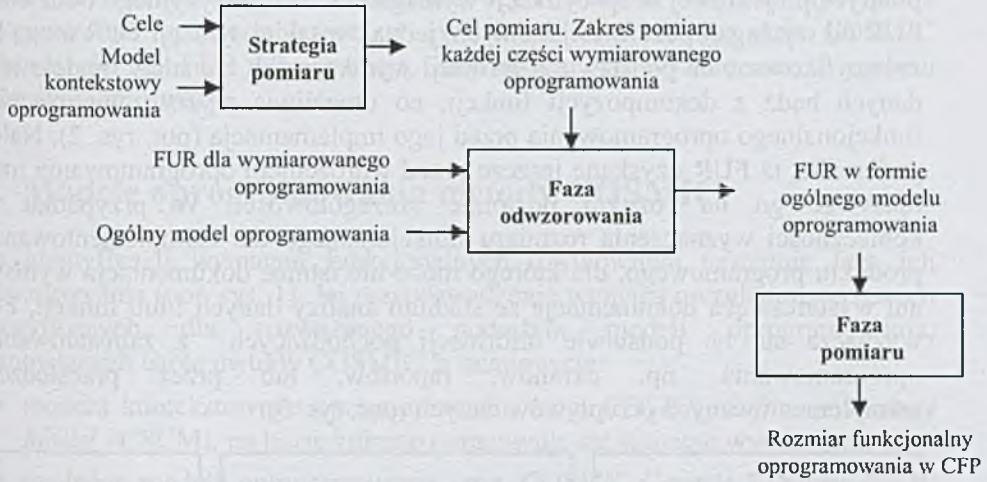
mierzonym oprogramowaniem" [24, s. 35]. Jej sens jest zgodny z późniejszą definicją przyjętą w normie ISO/IEC 14143. Dlatego koncepcja pełnych punktów funkcyjnych z 1997 roku miała dotyczyć wymiarowania rozmiaru funkcjonalnego przede wszystkim systemów czasu rzeczywistego, infrastrukturalnych aplikacjach wspomagających oraz oprogramowania wbudowanego w urządzenia i sterującego pracą maszyn. Przedmiotem jej zainteresowania były również aplikacje biznesowe.

W 1998 na nieformalnym spotkaniu w Londynie członków międzynarodowej grupy ekspertów powstałej w celu opracowania standardu ISO/IEC dotyczącego FSM powołano do życia organizację COSMIC. Od tego czasu metoda oparta na FPP jest rozwijana pod jej patronatem. W wyniku podjętych przez tę organizację prac, przy których korzystano z dorobku różnych wariantów metody punktów funkcyjnych, w tym głównie metody IFPUG, powstała nowa wersja metody FFP o nazwie COSMIC-FFP, która bywa uznawana za pierwszą metodę FSM drugiej generacji. Po licznych testach metody COSMIC-FFP w wersji z 2001 r., przeprowadzanych zgodnie z zaleceniami ISO i IEC w bankowości, lotnictwie, przemyśle samochodowym oraz w korporacjach telekomunikacyjnych, w 2003 roku metoda ta została zaakceptowana przez te organizacje standaryzacyjne w normie ISO/IEC 19761 jako w pełni zgodna ze standardem ISO/IEC 14143. Później, w celu lepszego zsynchronizowania zasad rozważanej metody ze standardem FSM, zostały one zaktualizowane: najpierw w roku 2003 (jeszcze jako metoda COSMIC-FFP), a następnie w 2007 r. w ramach metody COSMIC w wersji 3.0 [5][1], która obowiązuje aktualnie. Nie wyklucza się kolejnych modyfikacji w miarę nabywania doświadczeń praktycznych w użytkowaniu rozważanej metody, jednak jej tzw. ogólny model oprogramowania, stanowiący podstawę FSM w metodzie COSMIC, ciągle pozostaje taki sam. Proces wymiarowania rozmiaru funkcjonalnego produktu programowego przy wykorzystaniu obecnie obowiązującej wersji metody COSMIC przebiega w następujących fazach (por. rys. 1):

1. Opracowanie strategii pomiaru, realizowane przed rozpoczęciem właściwego wymiarowania na bazie analizy celów oraz specyficznego dla opisywanej metody tzw. modelu kontekstowego oprogramowania, w wyniku czego następuje wyznaczenie celu i zakresu wymiarowania.
2. Odwzorowanie wymagań funkcjonalnych użytkownika (FUR) dla mierzonego oprogramowania - przy uwzględnieniu specyficznego dla opisywanej metody tzw. ogólnego modelu oprogramowania oraz celu i zakresu wymiarowania - w FUR wyrażone w formie ogólnego modelu oprogramowania.
3. Właściwe wymiarowanie na bazie FUR wyrażonych w formie ogólnego modelu oprogramowania, w rezultacie czego uzyskuje się rozmiar funkcjonalny produktu programowego wyrażony w punktach funkcyjnych

[1] Dla uproszczenia nazwy począwszy od wersji 3.0 zrezygnowano z „FFP".

COSMIC (ang. *COSMIC Function Points* – CFP), przy czym 1 CFP jest definiowany jako rozmiar pojedynczego tzw. przepływu danych (por. tabela 1).



Rys. 1. Struktura metody punktów funkcyjnych COSMIC

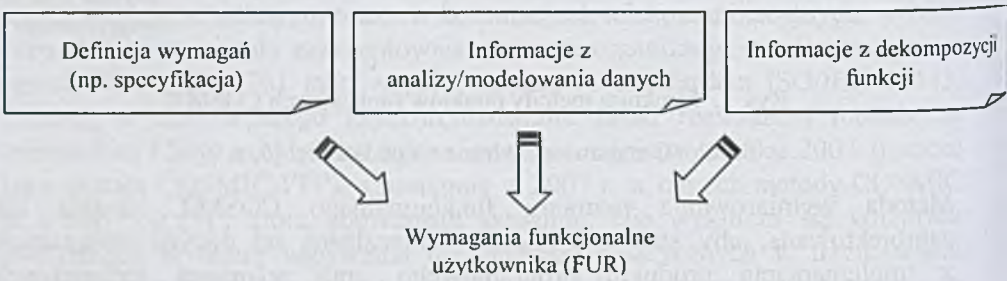
Źródło: Opracowanie własne na podstawie: [6, s. 9]

Metoda wymiarowania rozmiaru funkcjonalnego COSMIC została tak zaprojektowana, aby stanowić podejście niezależne od decyzji związanych z implementacją produktu programowego oraz wymagań technicznych i jakościowych, które odpowiadają na pytanie o to, *jak* ma działać oprogramowanie. Dlatego w jej ramach wymiarowanie odbywa się na bazie wymagań funkcjonalnych użytkownika, zdefiniowanych w zgodzie z pierwszą częścią normy ISO/IEC 14143, dla określonej części oprogramowania. Są one przedmiotem odniesienia dla zbioru modeli, zasad, reguł i procesów, w wyniku zastosowania których otrzymuje się wartość liczbową reprezentującą rozmiar funkcjonalny poszczególnych części oprogramowania w CFP. Przy czym funkcjonalność dotyczy przetwarzania informacji, które musi realizować oprogramowanie dla *wszystkich* swoich użytkowników, dlatego w rozważanej metodzie dla uściślenia wprowadza się termin *użytkowników funkcjonalnych* (ang. *functional users*), oznaczający osoby i rzeczy *„wysyłające i będące zamierzonymi odbiorcami danych do i z wymaganej funkcjonalności"* [6, s. 11][1].

Z punktu widzenia pomiaru rozmiaru funkcjonalnego produktu programowego metodą COSMIC najistotniejszą cechą oprogramowania jest zatem jego funkcjonalność, którą dostarcza ono lub ma dostarczać użytkownikom funkcjonalnym. Jak wynika z rys. 1, zanim nastąpi faza odwzorowania FUR dla

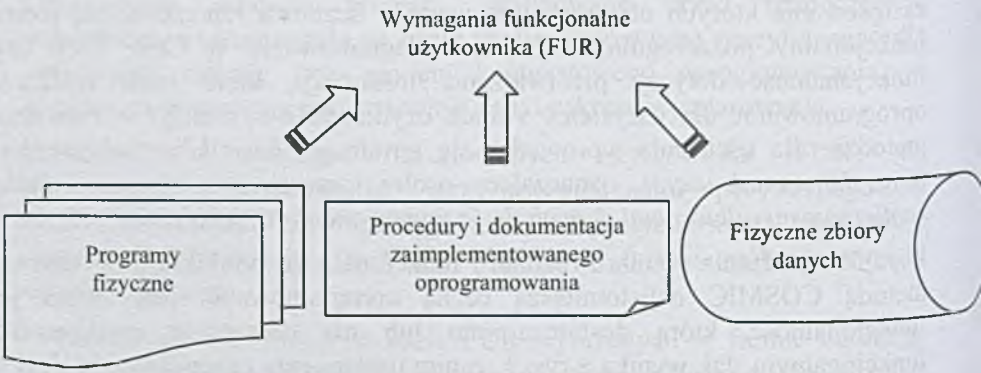
[1] Pozycja ta stanowi zasadnicze źródło informacji zawartych w punktach 2 i 3 niniejszego opracowania.

wymiarowanego oprogramowania w FUR w postaci ogólnego modelu oprogramowania, co musi się odbyć przed zastosowaniem reguł i procedur właściwego pomiaru, należy te wymagania zidentyfikować. Nierzadko w praktyce projektowej w specyfikacji wymagań nie ma precyzyjnego oddzielenia FUR od wymagań pozafunkcjonalnych, jednak w takiej sytuacji FUR mogą być zidentyfikowane na podstawie informacji wynikających z analizy modelowania danych bądź z dekompozycji funkcji, co umożliwia wyznaczenie rozmiaru funkcjonalnego oprogramowania przed jego implementacją (por. rys. 2). Należy podkreślić, iż FUR uzyskane jeszcze przed wdrożeniem oprogramowania mogą opisywać go na różnym poziomie szczegółowości. W przypadku zaś konieczności wyznaczenia rozmiaru funkcjonalnego dla zaimplementowanego produktu programowego, dla którego może nie istnieć dokumentacja wymagań ani wystarczająca dokumentacja ze stadium analizy danych i/lub funkcji, FUR wyznacza się na podstawie informacji pochodzących z zainstalowanego oprogramowania, np. ekranów, raportów, lub przez prześledzenie zaimplementowanych przepływów danych (por. rys. 3).



Rys. 2. Pre-wdrożeniowy model COSMIC wyznaczania wymagań funkcjonalnych użytkownika

Źródło: [5, s. 11]



Rys. 3. Post-wdrożeniowy model COSMIC wyznaczania wymagań funkcjonalnych użytkownika

Źródło: [5, s. 12]

Sposób wyznaczania wymagań funkcjonalnych użytkownika jest zatem różny w zależności od tego, czy oprogramowanie już istnieje, czy też ma być dopiero stworzone. W każdym jednak przypadku FUR stanowią opis funkcjonalności, która ma być lub już jest dostarczana użytkownikom funkcjonalnym przez wymiarowane oprogramowanie, w związku z czym w metodzie COSMIC zakłada się, że FUR albo istnieją, albo mogą być wyznaczone w oparciu o kreślone informacje.

4. Modele oprogramowania metody COSMIC

Po identyfikacji wymagań funkcjonalnych użytkownika następuje faza ich odwzorowania (por. rys. 1). Jej przeprowadzenie wymaga uwzględnienia dwóch specyficznych dla rozważanego podejścia modeli oprogramowania, stanowiących istotę metody COSMIC, a mianowicie:

- modelu kontekstowego oprogramowania (ang. *COSMIC Software Context Model* - CSCM), na bazie którego opracowuje się strategię wymiarowania;
- ogólnego modelu oprogramowania (ang. *COSMIC Generic Software Model* - CGSM), w formie którego wyrażane są wymagania funkcjonalne użytkownika.

Jeszcze przed fazą odwzorowania istnieje konieczność zdefiniowania zakresu wymiarowania każdej części mierzonego produktu programowego, przy czym należy uwzględnić kontekst jej działania, tzn. jakiegokolwiek inne oprogramowanie i/lub sprzęt, z którym się ona komunikuje. Model kontekstowy oprogramowania wprowadza zasady niezbędne do tej definicji. Znaczenie koniecznych do ich zrozumienia terminów znajdzie Czytelnik w tabeli 1. Zasady te są następujące:

- oprogramowanie jest ograniczone przez sprzęt (z jednej strony przez urządzenia konwersacyjne czy sensory, przekaźniki, z drugiej zaś przez różne typy pamięci trwałej);
- typowe oprogramowanie jest zbudowane z warstw (por. rys. 4 i 5);
- warstwa oprogramowania może zawierać jedną lub więcej oddzielnych równorzędnych części oprogramowania, a każda z jego części może składać się z osobnych równorzędnych komponentów - przy czym część oprogramowania w dowolnej warstwie może być dekomponowana na równorzędne komponenty na różnych poziomach szczegółowości (np. na poszczególne moduły funkcjonalne czy klasy obiektów), a przy użyciu omawianej metody mogą być one oddzielnie wymiarowane, niezależnie od tego poziomu;
- każda część wymiarowanego oprogramowania powinna być zdefiniowana co do swojego zakresu pomiaru, który musi być całkowicie zawarty w ramach jednej warstwy – wynika to z faktu, że każda warstwa ma swoje własne funkcje, jak również może być rozwijana w innej technologii;

- zakres pomiaru części wymiarowanego oprogramowania powinien zależeć od celu pomiaru – jeżeli jest nim np. estymacja pracochłonności, a każda część oprogramowania jest rozwijana w innej technologii, to powinno się ów zakres wyznaczyć oddzielnie dla każdej części, jako że pracochłonność będzie zależeć od wykorzystywanej technologii;
- użytkownicy funkcjonalni części oprogramowania powinni być zidentyfikowani na podstawie FUR dla tej części wymiarowanego oprogramowania jako osoby i/lub rzeczy będące nadawcami i/lub zamierzonymi odbiorcami danych (por. rys. 6 i rys. 7) – nie zaś jako sprzęt i/lub dodatkowe oprogramowanie, które umożliwia realizację tych funkcji (np. system operacyjny czy sterowniki urządzeń - por. rys. 4 i rys. 5);
- część oprogramowania komunikuje się ze swoimi użytkownikami funkcjonalnymi poprzez przepływy danych przekraczające jej granice i może ona przenosić dane do i z pamięci trwałej znajdującej się w jej granicach (por. rys. 6 i rys. 7);
- wymagania funkcjonalne użytkownika dla oprogramowania mogą być wyrażone na różnych poziomach uszczegółowienia opisu części oprogramowania;
- poziom uszczegółowienia FUR, na którym powinno dokonywać się wymiarowania, powinien odpowiadać poziomowi uszczegółowienia procesów funkcjonalnych;
- jeżeli nie jest możliwy pomiar na poziomie uszczegółowienia procesów funkcjonalnych (problem spotykany zazwyczaj we wstępnych etapach przedsięwzięć konstrukcyjnych polegających na budowie nowych produktów), to FUR dla oprogramowania powinny być wymiarowane przy wykorzystaniu aproksymacji i przeskalowane do poziomu uszczegółowienia procesów funkcjonalnych.

Tab. 1. Definicje pojęć wykorzystywanych przez modele oprogramowania metody COSMIC

Termin	Definicja
Cel wymiarowania (ang. *purpose of measurement*)	Stwierdzenie definiujące przyczyny, dla których podejmuje się wymiarowanie, oraz przeznaczenie jego rezultatów (np. podstawa do oszacowania pracochłonności).
Granica (ang. *boundary*)	Konceptualny interfejs między wymiarowanym oprogramowaniem a jego użytkownikami funkcjonalnymi.
Warstwa (oprogramowania – ang. *layer*)	Część oprogramowania stanowiąca rezultat funkcjonalnego podziału architektury oprogramowania, która razem ze sprzętem składa się na cały system komputerowy (por. rys. 4 i rys. 5), gdzie: - warstwy są zorganizowane w sposób hierarchiczny, - istnieje tylko jedna warstwa oprogramowania na każdym poziomie hierarchii, - pomiędzy usługami funkcjonalnymi dostarczanymi przez oprogramowanie w dowolnych dwóch warstwach w architekturze oprogramowania wymieniających bezpośrednio dane istnieje zależność hierarchiczna typu „nadrzędne/podrzędne"; - dowolne dwie warstwy w architekturze oprogramowania wymieniające dane tylko część tych danych interpretują identycznie.
Komponent równorzędny (części oprogramowania – ang. *peer component*)	Fragment oprogramowania stanowiący rezultat funkcjonalnego podziału wymagań funkcjonalnych użytkownika dla części oprogramowania w ramach warstwy na ich współpracujący ze sobą zbiór w taki sposób, aby każda z tych części zaspakajała określoną porcję FUR dla tej części oprogramowania (por. rys. 4 i rys. 5). Każdy komponent równorzędny musi spełniać następujące założenia: - w zestawie komponentów równorzędnych części oprogramowania występujących w jednej warstwie nie istnieje zależność hierarchiczna (w przeciwieństwie do warstw); - wszystkie komponenty równorzędne części oprogramowania musza współpracować ze sobą w celu właściwego działania tej części oprogramowania; - komponenty równorzędne dowolnej jednej warstwy przesyłające/dzielące między sobą dane definiują te dane identycznie (rozpoznają te same atrybuty i podgrupy

	danych).
Proces funkcjonalny (ang. *functional process*)	Elementarny komponent zbioru wymagań funkcjonalnych użytkownika obejmujący unikalny, spójny i niezależnie wykonywany zestaw przepływów danych (por. rys. 8). Jest inicjowany (wyzwalany) przez przepływ danych typu wejście (por. niżej) pochodzący od użytkownika funkcjonalnego, który informuje część oprogramowania, że użytkownik zidentyfikował zdarzenie inicjujące. Kończy się w momencie wykonania wszystkich działań, które są wymagane w reakcji na to zdarzenie.
Zdarzenie inicjujące (wyzwalające – ang. *triggering event*)	Zdarzenie powodujące, że użytkownik funkcjonalny części oprogramowania inicjuje (wyzwala) co najmniej jeden proces funkcjonalny. W zestawie FUR każde zdarzenie, które wyzwala proces funkcjonalny jest dla tego zestawu wymagań niepodzielne. Zdarzeniami wyzwalającymi mogą być np. zdarzenia czasowe. Zdarzenie wyzwalające albo zaszło, albo nie (dla danego zestawu FUR).
Podproces (ang. *sub-process*)	Część procesu funkcjonalnego, która albo przemieszcza dane (do oprogramowania od użytkownika funkcjonalnego/na zewnątrz oprogramowania do użytkownika funkcjonalnego lub do/z pamięci trwałej), albo operuje na danych (por. rys. 8).

Termin	Definicja
Operowanie na danych (ang. *data manipulation*)	Działanie na danych inne niż ich przepływ do/na zewnątrz procesu funkcjonalnego bądź pomiędzy procesem funkcjonalnym a pamięcią trwałą (por. rys. 8).
Obiekt zainteresowania (ang. *object of interest*)	Rzeczywisty lub konceptualny obiekt, o którym oprogramowanie musi przetwarzać i/lub gromadzić dane, zidentyfikowany z perspektywy FUR i niezbędny do identyfikacji przepływów danych (w terminologii relacyjnych baz danych jest to encja).
Grupa danych (ang. *data group*)	Wyodrębniony, niepusty, nieuporządkowany i nieredundantny zestaw atrybutów danych, w którym każdy uwzględniony atrybut opisuje komplementarny (wzajemnie się uzupełniający) aspekt tego samego obiektu będącego przedmiotem zainteresowania.
Atrybut danych (ang. *data attribute*)	Najmniejsza porcja informacji w zidentyfikowanej grupie danych, mająca znaczenie z perspektywy FUR dla oprogramowania.
Pamięć trwała (ang. *persistent storage*)	Pamięć pozwalająca procesowi funkcjonalnemu na przechowywanie danych po jego zakończeniu i/lub umożliwiająca procesowi funkcjonalnemu odczyt danych zgromadzonych przez inny proces funkcjonalny bądź przez wcześniejsze wystąpienie tego samego procesu

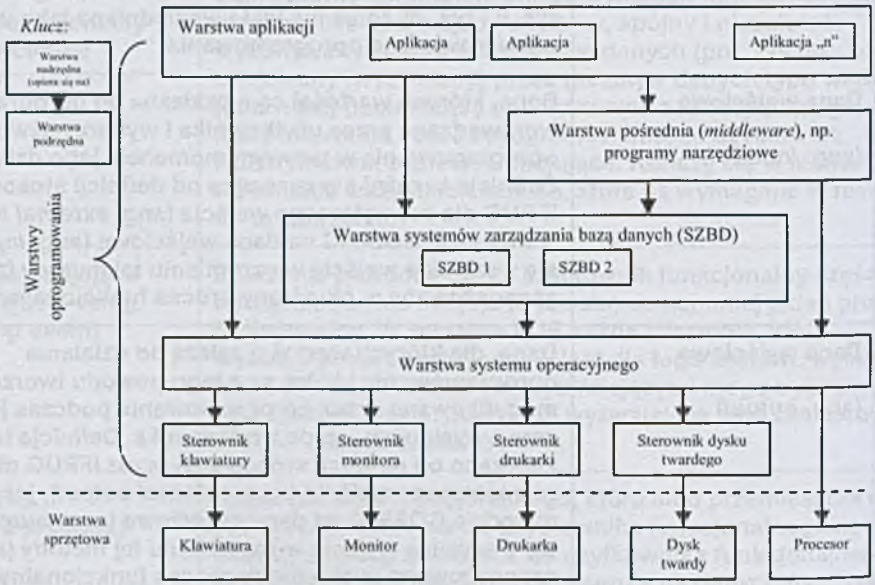
	funkcjonalnego, ewentualnie przez inny proces (np. z pamięci tylko do czytania). W modelu COSMIC pamięć trwała z punktu widzenia granic znajduje się po stronie oprogramowania (por. rys. 6 i rys. 7), toteż nie jest uwzględniana jako użytkownik wymiarowanego oprogramowania.
Dane wejściowe (ang. *input*)	**Dane, których wartości są niezależne od oprogramowania, wprowadzane przez użytkownika i wykorzystywane przez oprogramowanie w pewnym momencie jego działania. Definicja ta różni się znacząco od definicji stosowanej przez IFPUG dla zewnętrznego wejścia (ang. *external input*), jako że w metodzie COSMIC na dane wejściowe (ang. *input*) składają się wszystkie wejścia w rozumieniu tej metody (ang. *entries*) zaangażowane w określony proces funkcjonalny.**
Dane wyjściowe (ang. *output*)	**Dane, dla których wartości zależą od działania oprogramowania i które są z tego powodu tworzone lub modyfikowane przez oprogramowanie podczas jego działania przed wysłaniem ich do użytkownika. Definicja ta różni się znacząco od definicji stosowanej przez IFPUG dla zewnętrznego wyjścia (ang. *external output*), jako że w metodzie COSMIC na dane wyjściowe (ang. *output*) składają się wszystkie wyjścia w rozumieniu tej metody (ang. *exits*) zaangażowane w określony proces funkcjonalny.**

Źródło: Opracowanie własne na podstawie: [3, s. 14-15, 17, 19-21]

Dla zrozumienia reguł zawartych w CSCM istotne jest także rozróżnienie dwóch punktów widzenia na system oprogramowania, jako że oznaczają one inny kontekst dla wymiarowanej części produktu. Przy czym tylko drugi z nich jest niezbędny do celów FSM. Te dwa punkty widzenia to:

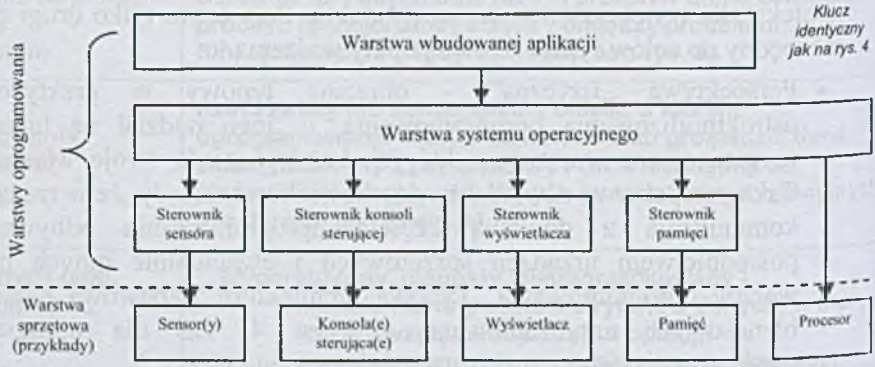
- Perspektywa „fizyczna" – obrazuje typowy w praktyce sposób ustrukturalizowania oprogramowania, tj. jego podział na hierarchicznie zorganizowane warstwy, z których każda posiada swoje własne funkcje. Taka perspektywa uświadamia użytkownikowi metody, że w rzeczywistości komunikacja z dowolną częścią oprogramowania odbywa się za pośrednictwem urządzeń sprzętowych i ewentualnie innych pośrednich warstw oprogramowania. Typową architekturę warstwową dla systemów biznesowych zaprezentowano na rys. 4, zaś dla oprogramowania wbudowanego w systemy czasu rzeczywistego na rys. 5.
- Perspektywa „logiczna" – stanowi abstrakcyjne ujęcie perspektywy „fizycznej", przyjęte dla celów wymiarowania rozmiaru funkcjonalnego produktu programowego. Pokazuje, że użytkownicy funkcjonalni wymiarowanej części oprogramowania wchodzą w interakcje z oprogramowaniem poprzez jego granice za pomocą przepływów danych typu wejście i wyjście, jak również, że oprogramowanie to przemieszcza dane do i z pamięci trwałej za pomocą przepływów danych typu zapisy i odczyty. W tym podejściu cały sprzęt i dodatkowe oprogramowanie, które umożliwia te interakcje jest ignorowane. Schematy

ilustrujące to spojrzenie przedstawiono na rys. 6 (dla aplikacji biznesowych) i rys. 7 (dla aplikacji czasu rzeczywistego).



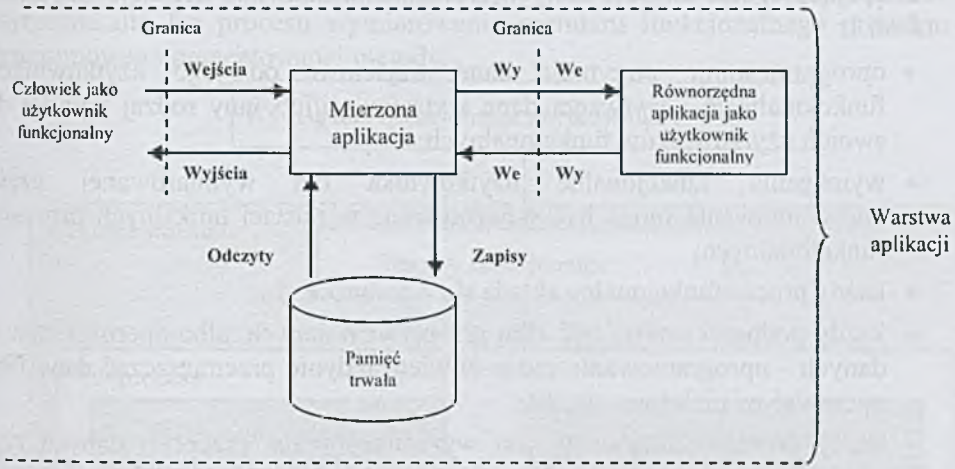
Rys. 4. Typowa architektura warstwowa dla systemów biznesowych w metodzie COSMIC – perspektywa „fizyczna"

Źródło: [6, s. 13].



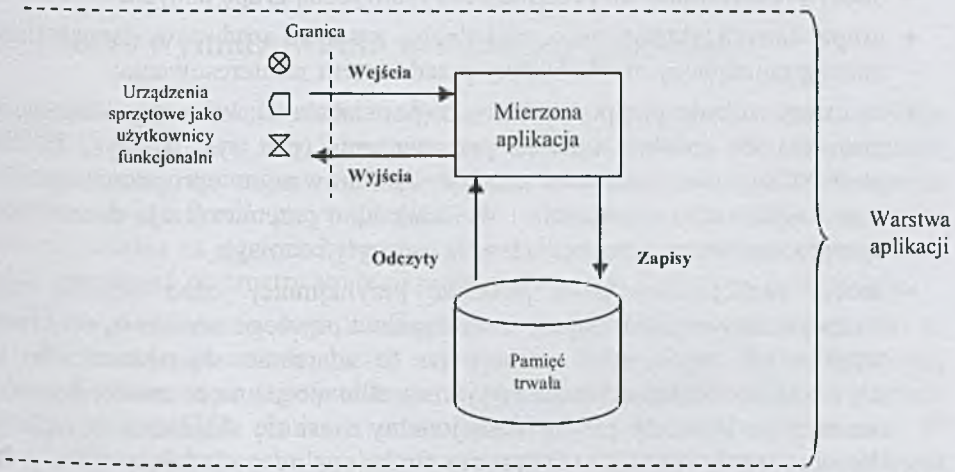
Rys. 5. Typowa architektura warstwowa dla oprogramowania wbudowanego w systemy czasu rzeczywistego w metodzie COSMIC – perspektywa „fizyczna"

Źródło: [6, s. 13]



Rys. 6. Warstwa aplikacji dla systemów biznesowych w metodzie COSMIC – perspektywa „logiczna"

Źródło: [6, s. 15]



Rys. 7. Warstwa aplikacji dla aplikacji wbudowanych czasu rzeczywistego w metodzie COSMIC – perspektywa „logiczna"

Źródło: [6, s. 15]

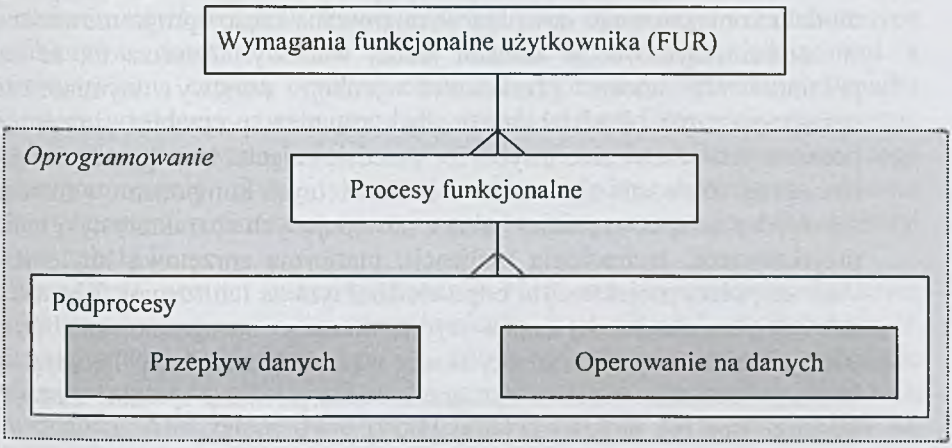
Po interpretacji FUR dla wymiarowanego oprogramowania za pomocą terminów modelu kontekstowego stosuje się wobec nich ogólny model oprogramowania (CGSM). Ma on na celu zidentyfikowanie komponentów funkcjonalnych, które będą podlegać pomiarowi. W modelu tym zakłada się, że dla oprogramowania, które może być wymiarowane za pomocą metody COSMIC, prawdziwe są

następujące ogólne zasady, których zrozumienie ułatwiają definicje zawarte w tabeli 1:

- oprogramowanie otrzymuje dane wejściowe od jego użytkowników funkcjonalnych i wytwarza dane wyjściowe i/lub inny rodzaj wyniku dla swoich użytkowników funkcjonalnych;
- wymagania funkcjonalne użytkownika dla wymiarowanej części oprogramowania mogą być odwzorowane w postaci unikalnych procesów funkcjonalnych;
- każdy proces funkcjonalny składa się z podprocesów;
- każdy podproces może być albo przepływem danych, albo operowaniem na danych - oprogramowanie może bowiem jedynie przemieszczać dane i/lub operować na nich (por. rys. 8);
- każdy proces funkcjonalny jest wyzwalany przez przepływ danych typu wejście pochodzący od użytkownika funkcjonalnego, który informuje proces funkcjonalny o identyfikacji przez użytkownika pewnego zdarzenia (inicjującego) – dlatego wszystkie FUR mogą być wyrażone jako lista zdarzeń oraz powiązane z nimi procesy funkcjonalne stanowiące reakcję oprogramowania na każde zdarzenie;
- każdy przepływ danych przemieszcza tylko jedną grupę danych;
- grupa danych składa się z unikalnego zestawu atrybutów danych, które opisują pojedynczy obiekt będący przedmiotem zainteresowania;
- są cztery rodzaje przepływu danych (por. tabela 1), które wyróżnia się na postawie ich źródła i miejsca przeznaczenia (por. rys. 6 i rys. 7): albo przekraczają one granice pomiędzy wymiarowanym oprogramowaniem a jego użytkownikami (wejścia i wyjścia), albo przemieszczają dane między oprogramowaniem a pamięcią trwałą (odczyty i zapisy);
- proces funkcjonalny musi posiadać przynajmniej jedno wejście, czyli przepływ danych informujący o wystąpieniu pewnego zdarzenia, oraz jedno wyjście lub zapis, czyli reakcję na to zdarzenie skierowane albo do użytkownika funkcjonalnego (wyjście), albo do pamięci trwałej (zapis) – oznacza to, iż każdy proces funkcjonalny musi się składać z co najmniej dwóch przepływów danych (proces funkcjonalny posiadający tylko jeden przepływ danych jest w praktyce bezużyteczny);
- podproces operowania na danych w obecnie obowiązującej wersji metody nie jest mierzony oddzielnie, a jego funkcjonalność jest uwzględniana przez podproces przepływu danych, z którym jest on związany (por. rys. 8)[1].

[1] Wynika to z faktu, że definicje niezbędne do pomiaru podprocesu typu operowanie na danych (np. formatowanie, walidacja danych, tworzenie atrybutów grupy danych) są ciągle przedmiotem dyskusji.

Zasady modelu kontekstowego i modelu oprogramowania stanowią podstawowe wytyczne dla faz procesu wymiarowania rozmiaru funkcjonalnego produktu programowego prezentowanej metody.



Rys. 8. Struktura wymagań funkcjonalnych użytkownika w metodzie COSMIC

Źródło: [6, s. 19]

5. Proces wymiarowania w metodzie COSMIC

Na proces wymiarowania rozmiaru funkcjonalnego oprogramowania w metodzie COSMIC składają się trzy fazy (por. rys. 1): przygotowania strategii pomiaru, odwzorowania wymagań funkcjonalnych użytkownika oraz właściwego wymiarowania. Przed przystąpieniem do właściwego wymiarowania osoba odpowiedzialna za jego przebieg musi uzgodnić ze sponsorem tego procesu oraz udokumentować parametry strategii pomiaru (por. rys. 9). Są one następujące:

- Cel wymiarowania. Jego zdefiniowanie ma decydujące znaczenie, jako że determinuje pozostałe, niżej wymienione parametry, które zazwyczaj wyznaczane są na jego podstawie w sposób iteracyjny, głównie ze względu na zmianę FUR oraz ich coraz większy poziom uszczegółowienia. W ramach definicji celu wymiarowania należy określić przyczyny podjęcia tego procesu oraz sposób wykorzystania jego rezultatów (por. tabela 1). Determinuje on także wymaganą dokładność wymiarowania, która zależy m.in. od etapu cyklu życia przedsięwzięcia, na którym jest ono przeprowadzane.
- Zakres pomiaru każdej części wymiarowanego oprogramowania. Ogólny zakres wymiarowanego oprogramowania wynika z celu tego procesu. W jego ramach ustanawia się tę funkcjonalność oprogramowania, która będzie podlegała wymiarowaniu. W zależności od celu pomiaru ów ogólny zakres może zostać podzielony na pewną liczbę oddzielnych części funkcjonalnych oprogramowania, z których każda powinna być mierzona osobno, tj.

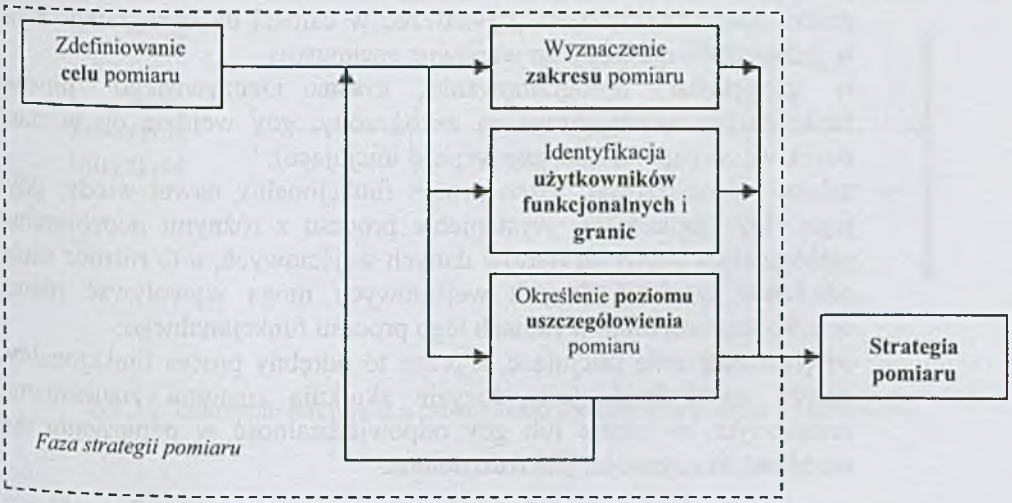
powinna posiadać swój własny zakres pomiaru. Taki podział jest konieczny, jeżeli ogólny zakres obejmuje oprogramowanie złożone z co najmniej dwóch warstw (por. rys. 4 i rys. 5), ponieważ zgodnie z jedną z zasad modelu kontekstowego dowolna wymiarowana część oprogramowania musi w całości rezydować w ramach jednej warstwy. Oznacza to, że zakres pomiaru nie może przekroczyć jednej warstwy wymiarowanego oprogramowania. Podział może być również potrzebny, jeżeli celem pomiaru jest szacowanie atrybutów przedsięwzięcia, którego rezultatem ma być oprogramowanie złożone z wielu oddzielnych komponentów różniących się między sobą co najmniej jedną z następujących charakterystyk: techniką projektowania, technologią realizacji, platformą sprzętową implementacji i/lub zespołem projektowym odpowiedzialnym za ich rozwój. Wyznaczenie zakresu pomiaru każdej części wymiarowanego oprogramowania wymaga więc przede wszystkim identyfikacji warstw (por. tabela 1), przy czym każda potencjalna warstwa oprogramowania powinna spełniać następujące warunki (por. rys. 4, 5, 6 i 7) [5, s. 21-22]:

- oprogramowanie w jednej warstwie wymienia dane z oprogramowaniem w innej warstwie za pomocą odpowiednich procesów funkcjonalnych;
- zależność hierarchiczna pomiędzy warstwami polega na tym, że oprogramowanie w danej warstwie może wykorzystywać usługi funkcjonalne oprogramowania z dowolnej warstwy podrzędnej, co oznacza, że oprogramowanie warstwy nadrzędnej w celu prawidłowego działania opiera się na usługach oprogramowania z warstw podporządkowanych, ale oprogramowanie z warstwy podporządkowanej (wraz ze swoimi warstwami podporządkowanymi) może funkcjonować bez usług oprogramowania z warstwy nadrzędnej;
- oprogramowanie w danej warstwie nie musi wykorzystywać wszystkich usług funkcjonalnych dostarczanych przez oprogramowanie warstwy podrzędnej;
- dane wymieniane między oprogramowaniem w dowolnych dwóch warstwach są zdefiniowane i interpretowane odmiennie na bazie odpowiednich FUR dla tych dwóch części oprogramowania, co oznacza, że dwie części oprogramowania identyfikują inne atrybuty danych i/lub podgrupy danych, chociaż musi istnieć część wspólna, tj. co najmniej jeden wspólnie zdefiniowany atrybut/podgrupa w celu możliwości interpretacji przez oprogramowanie warstwy odbierającej danych wysłanych przez oprogramowanie warstwy wysyłającej.

• Użytkownicy funkcjonalni i granice każdej wymiarowanej części oprogramowania. Waga tej fazy wynika z faktu, iż funkcjonalność będąca przedmiotem zainteresowania poszczególnych użytkowników funkcjonalnych danej aplikacji może się znacznie różnić i obejmować jedynie podzbiór całej wymaganej funkcjonalności. Użytkownicy funkcjonalni każdej wymiarowanej części oprogramowania mogą być

zidentyfikowani dzięki prześledzeniu przepływów danych wchodzących do i wychodzących z danej części oprogramowania (por. rys. 6 i rys. 7), wynikających z FUR dla tej części i przy uwzględnieniu celu pomiaru. Po ustaleniu użytkowników funkcjonalnych jako nadawców i zamierzonych odbiorców danych można łatwo określić granice pomiędzy nimi a wymiarowaną częścią oprogramowania, z którą się komunikują. Ponadto granice występują zawsze pomiędzy każdą zidentyfikowaną parą warstw, gdzie oprogramowanie jednej warstwy stanowi użytkownika funkcjonalnego oprogramowania drugiej warstwy, oraz między dowolnymi dwoma komponentami równorzędnymi tej samej warstwy, co wynika z ich definicji (por. tabela 1).

- Wymagany poziom uszczegółowienia FUR dla wymiarowania. Celem ustalenia poziomu uszczegółowienia wymagań funkcjonalnych użytkownika, na którym ma się odbyć wymiarowanie jest uniknięcie nieporozumień wynikających z odmiennych podejść wymiarujących do tych wymagań. Aby zatem dwie osoby odpowiedzialne za wymiarowanie tej samej części oprogramowania otrzymały zbliżone wyniki, muszą one przyjąć ten sam poziom uszczegółowienia FUR. Dlatego w metodzie COSMIC wprowadzono standardowy poziom ich uszczegółowienia: zgodnie z jedną z reguł modelu kontekstowego powinien on odpowiadać takiemu poziomowi, na którym zidentyfikowano procesy funkcjonalne i zdefiniowano dla nich przepływy danych[1].



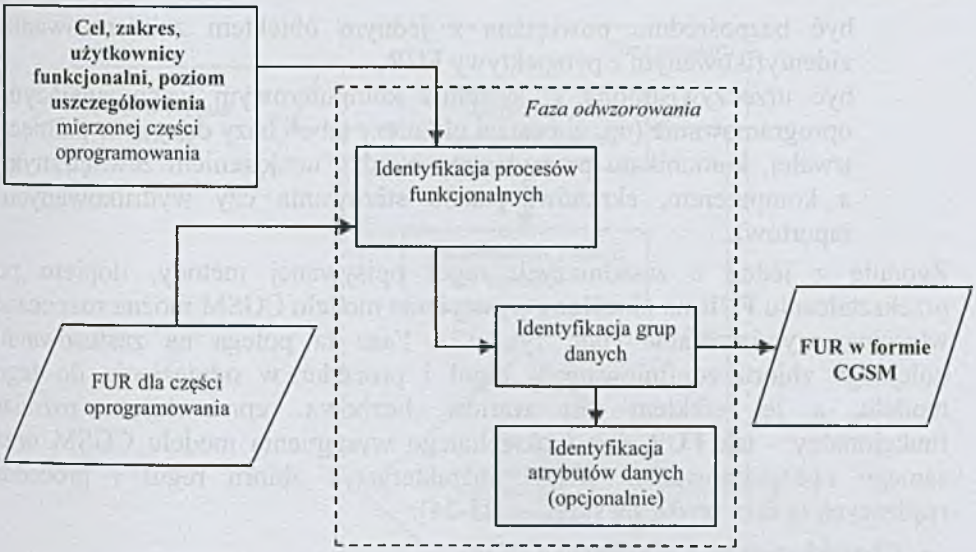
Rys. 9. Faza opracowania strategii pomiaru w metodzie COSMIC

Źródło: [5, s. 16]

[1] Problem ten jest dokładnie przedstawiony w: [2, s. 7-14].

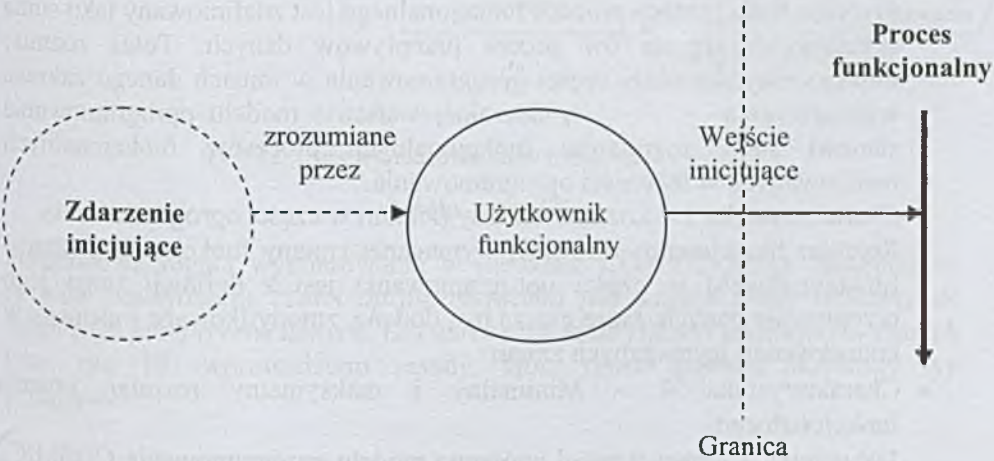
W kolejnej fazie procesu wymiarowania metodą COSMIC, fazie odwzorowania (por. rys. 10), wymagania funkcjonalne użytkownika dla każdej wymiarowanej części oprogramowania, wyznaczone w oparciu o model pre-wdrożeniowy (por. rys. 2) lub post-wdrożeniowy (por. rys. 3), są przekształcane przy uwzględnieniu modelu kontekstowego na określone wystąpienie ogólnego modelu oprogramowania. Wymagane są do tego następujące kroki [5, s. 35-36, 39-40]:

- Identyfikacja w otoczeniu użytkowników funkcjonalnych zdarzeń, na które oprogramowanie musi zareagować, tj. zdarzeń inicjujących, dzięki czemu może nastąpić identyfikacja procesów funkcjonalnych (por. rys. 11). Dla ułatwienia identyfikacji procesów funkcjonalnych wprowadzono następujące reguły:
  - proces funkcjonalny musi być wyodrębniony z co najmniej jednego możliwego do zidentyfikowania wymagania funkcjonalnego użytkownika wyznaczonego w obrębie uzgodnionego zakresu pomiaru;
  - proces funkcjonalny powinien być wywołany przez wystąpienie możliwego do zidentyfikowania zdarzenia inicjującego (wyzwalającego);
  - określone zdarzenie inicjujące może wyzwalać jeden lub więcej procesów funkcjonalnych, które są wykonywane równolegle, a dany proces funkcjonalny może być zainicjowany przez więcej niż jedno zdarzenie inicjujące;
  - proces funkcjonalny musi składać się z przynajmniej dwóch przepływów danych: wejścia oraz wyjścia lub zapisu;
  - proces funkcjonalny musi przynależeć w całości do oprogramowania w jednej i tylko jednej jego warstwie;
  - w przypadku oprogramowania czasu rzeczywistego proces funkcjonalny należy uznać za zakończony, gdy wejdzie on w stan oczekiwania (np. na następne wejście inicjujące);
  - należy zidentyfikować jeden proces funkcjonalny nawet wtedy, gdy jego FUR dopuszczają wystąpienie procesu z różnymi podzbiorami maksymalnej liczby atrybutów danych wejściowych, a te różnice i/lub odmienne wartości danych wejściowych mogą wywoływać różne ścieżki przetwarzania w ramach tego procesu funkcjonalnego;
  - odrębne zdarzenie inicjujące, a przez to odrębny proces funkcjonalny należy wyodrębnić, jeśli decyzje skutkują różnymi zdarzeniami rozłożonymi w czasie lub gdy odpowiedzialność w odniesieniu do określonych czynności jest rozdzielona.



Rys. 10. Faza odwzorowania w procesie wymiarowania rozmiaru funkcjonalnego metodą COSMIC

Źródło: [5, s. 32].



Rys. 11. Zdarzenie inicjujące a proces funkcjonalny w metodzie COSMIC

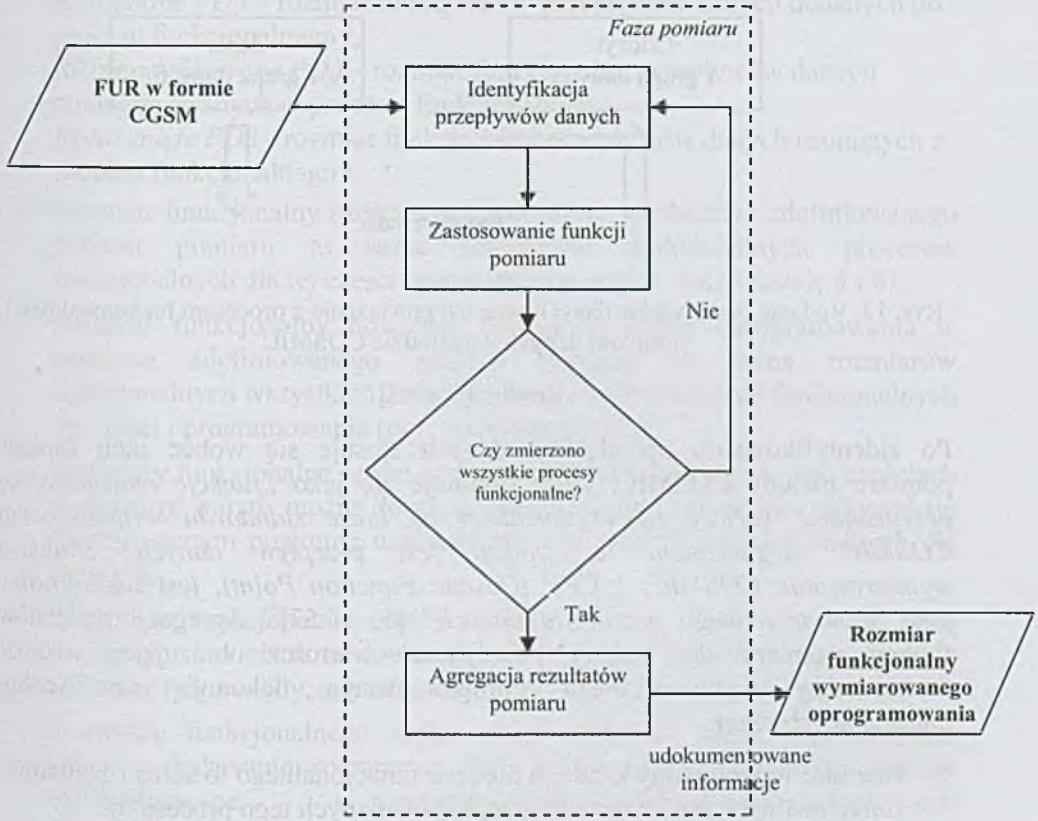
Źródło: [5, s. 35]

- Identyfikacja przepływów danych (wejść, wyjść, zapisów i odczytów) w ramach każdego procesu funkcjonalnego, co z kolei zależy od identyfikacji obiektów zainteresowania oraz przemieszczanych grup danych (por. tabela 1). Każda grupa danych powinna spełniać następujące warunki:
  - być unikalna i odróżnialna dzięki unikalnemu zestawowi atrybutów danych,

- być bezpośrednio powiązana z jednym obiektem zainteresowania zidentyfikowanym z perspektywy FUR,
- być urzeczywistniona w systemie komputerowym wspomagającym oprogramowanie (np. w postaci pliku czy tabeli bazy danych w pamięci trwałej, komunikatu przesyłanego między urządzeniem zewnętrznym a komputerem, ekranów, panelu sterowania czy wydrukowanych raportów).

Zgodnie z jedną z zasadniczych reguł opisywanej metody, dopiero po przekształceniu FUR na określone wystąpienie modelu CGSM można rozpocząć właściwe wymiarowanie (por. rys. 12). Faza ta polega na zastosowaniu kolejnego zbioru zdefiniowanych reguł i procedur w odniesieniu do tego modelu, a jej efektem jest wartość liczbowa reprezentująca rozmiar funkcjonalny - tak FUR, jak i określonego wystąpienia modelu CGSM oraz samego oprogramowania. Wśród charakterystyk zbioru reguł i procedur rządzących tą fazą wyróżnia się [6, s. 23-24]:

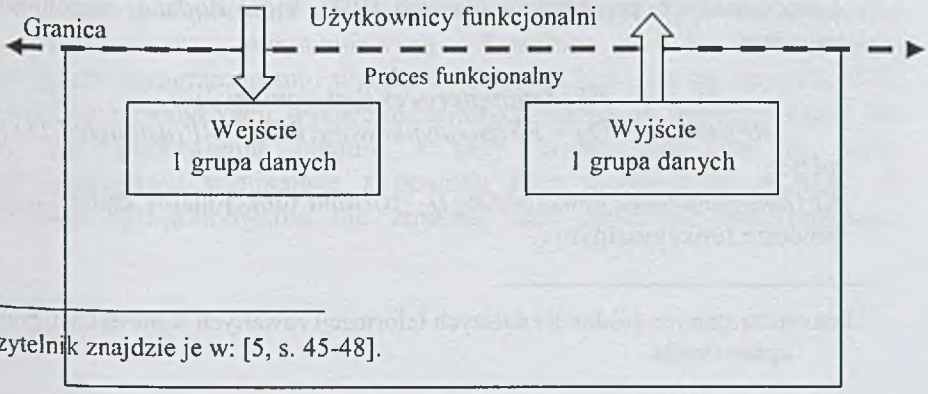
- Charakterystyka 1 – Jednostka miary.
  Standardowa jednostka pomiaru, tj. 1 CFP, jest zdefiniowana jako ekwiwalent pojedynczego przepływu danych.
- Charakterystyka 2 - Addytywność rozmiarów w obrębie danego zakresu pomiaru.
  Rozmiar funkcjonalny procesu funkcjonalnego jest zdefiniowany jako suma składających się na ów proces przepływów danych. Toteż rozmiar funkcjonalny dowolnej części oprogramowania w ramach danego zakresu wymiarowania w dowolnej warstwie modelu oprogramowania stanowi sumę rozmiarów funkcjonalnych procesów funkcjonalnych realizowanych w tej części oprogramowania.
- Charakterystyka 3 – Rozmiar zmiany (zmian) w części oprogramowania.
  Rozmiar funkcjonalny dowolnej wymaganej zmiany funkcjonalnej (zmian funkcjonalnych) w części oprogramowania jest z definicji sumą jego przepływów danych, które muszą być dodane, zmodyfikowane i usunięte w konsekwencji wymaganych zmian.
- Charakterystyka 4 – Minimalny i maksymalny rozmiar procesu funkcjonalnego.
  Jak wynika z jednej z zasad ogólnego modelu oprogramowania COSMIC, minimalny rozmiar funkcjonalny dla pojedynczego procesu funkcjonalnego to 2 CFP, jako że najmniejszy proces funkcjonalny musi posiadać wejście oraz wyjście lub zapis. Zmiana zaś może dotyczyć tylko jednego przepływu danych, dlatego minimalny rozmiar zmiany procesu funkcjonalnego wynosi 1 CFP. Nie istnieje natomiast górny limit rozmiaru funkcjonalnego dowolnego procesu funkcjonalnego, stąd nie istnieje również górny limit rozmiaru funkcjonalnego dla części oprogramowania.



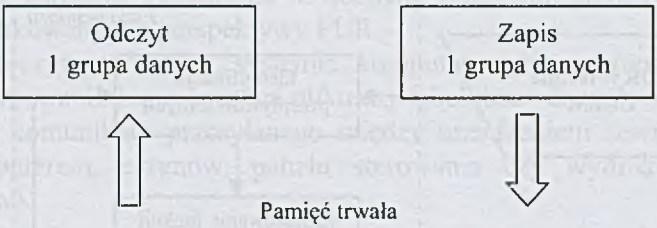
Rys. 12. Faza wymiarowania w metodzie COSMIC

Źródło: [5, s. 43]

Podstawowa reguła wymiarowania w metodzie COSMIC brzmi następująco: *rozmiar funkcjonalny części oprogramowania jest wprost proporcjonalny do liczby jego przepływów danych*. Dla ułatwienia identyfikacji przepływów danych (por. rys. 13) wprowadzono zasady, które musi spełniać określony typ przepływu[1].



[1] Czytelnik znajdzie je w: [5, s. 45-48].



Rys. 13. Rodzaje przepływów danych oraz ich powiązanie z procesem funkcjonalnym i grupami danych w metodzie COSMIC

Źródło: [5, s. 45]

Po zidentyfikowaniu przepływów danych stosuje się wobec nich funkcję pomiaru metody COSMIC, którą definiuje się jako *„funkcję matematyczną przypisującą wartość jej argumentowi na bazie standardu wymiarowania COSMIC. Argumentem tej funkcji jest przepływ danych. Standard wymiarowania COSMIC, 1 CFP (Cosmic Function Point), jest zdefiniowany jako rozmiar jednego przepływu danych"* [5, s. 57][1]. Agregacji rezultatów funkcji pomiaru do postaci pojedynczej wartości obrazującej rozmiar funkcjonalny wymiarowanego oprogramowania dokonuje się według następujących reguł:

1. Rozmiar funkcjonalny każdego procesu funkcjonalnego to suma rozmiarów funkcjonalnych pojedynczych przepływów danych tego procesu, tj.:

$$RF(proces\ funkcjonalny_i) = RF(wejścia_i) + RF(wyjścia_i) + RF(odczyty_i) + RF(zapisy_i),$$

gdzie:
$RF(proces\ funkcjonalny_i)$ – rozmiar funkcjonalny procesu funkcjonalnego i,
$RF(wejścia_i)$ - rozmiar funkcjonalny wejść w procesie funkcjonalnym i,
$RF(wyjścia_i)$ - rozmiar funkcjonalny wyjść w procesie funkcjonalnym i,
$RF(odczyty_i)$ - rozmiar funkcjonalny odczytów w procesie funkcjonalnym i,
$RF(zapisy_i)$ - rozmiar funkcjonalny zapisów w procesie funkcjonalnym i.

2. Rozmiar funkcjonalny zmian w procesie funkcjonalnym to suma rozmiarów funkcjonalnych przepływów danych (PD), które dodano, zmieniono lub usunięto:

$$RF(Zmiana(proces\ funkcjonalny_i)) = RF(dodane\ PD_i) + RF(zmodyfikowane\ PD_i) + RF(usunięte\ PD_i),$$

gdzie:
$RF(Zmiana(proces\ funkcjonalny_i))$ – rozmiar funkcjonalny zmian w procesie funkcjonalnym i,

[1] Pozycja ta stanowi źródło dla dalszych informacji zawartych w niniejszym punkcie opracowania.

RF(dodane PD_i) – rozmiar funkcjonalny przepływów danych dodanych do procesu funkcjonalnego *i*,
RF(zmodyfikowane PD_i) – rozmiar funkcjonalny przepływów danych zmodyfikowanych w procesie funkcjonalnym *i*,
RF(usunięte PD_i) - rozmiar funkcjonalny przepływów danych usuniętych z procesu funkcjonalnego *i*.

3. Rozmiar funkcjonalny części oprogramowania w obszarze zdefiniowanego zakresu pomiaru to suma rozmiarów funkcjonalnych procesów funkcjonalnych dla tej części oprogramowania (por. także zasadę 5 i 6).
4. Rozmiar funkcjonalny dowolnej zmiany w części oprogramowania w obszarze zdefiniowanego zakresu pomiaru to suma rozmiarów funkcjonalnych wszystkich zmian we wszystkich procesach funkcjonalnych tej części oprogramowania (por. także zasadę 5 i 6).
5. Rozmiary funkcjonalne części oprogramowania lub zmian w jego częściach w obszarze warstw można dodawać jedynie wtedy, gdy pomiar odbywał się na tym samym poziomie uszczegółowienia procesów funkcjonalnych dla ich FUR.
6. Rozmiary funkcjonalne części oprogramowania i/lub zmian w jego częściach w obszarze dowolnej jednej warstwy lub kilku warstw można dodawać tylko wtedy, gdy ma to sens ze względu na cel pomiaru.
7. Rozmiaru funkcjonalnego części oprogramowania nie można otrzymać poprzez dodawanie rozmiarów jego komponentów, chyba że zostaną wyeliminowane w nim udziały przepływów danych zachodzących między komponentami.
8. Jeśli metoda COSMIC zostanie zmodyfikowana w celu jej lokalnego wykorzystania, tj. w celu realizacji pomiaru pewnych aspektów, których nie pokrywa metoda standardowa, to jego rezultatu nie można dodawać do rezultatu pomiaru dokonanego zgodnie ze standardową metodą.

Część z powyższych reguł agregacji dotyczy pomiaru rozmiaru zmian funkcjonalnych w istniejącym oprogramowaniu. Przez zmiany funkcjonalne w metodzie COSMIC rozumie się dowolną kombinację działań polegających na dodaniu, zmodyfikowaniu i/lub usunięciu przepływów danych. W praktyce tego typu działania podejmowane są w ramach informatycznych przedsięwzięć doskonalących oraz utrzymaniowych. Potrzeba zmiany funkcjonalnej w istniejącym oprogramowaniu wynika zwykle z pojawienia się nowych FUR, z konieczności modyfikacji wyspecyfikowanych wymagań funkcjonalnych lub z potrzeby poprawienia błędów - przy czym jeżeli są to błędy w oprogramowaniu wynikające z powodu jego niezgodności z FUR, to modyfikacja oprogramowania nie zmienia jego rozmiaru funkcjonalnego,

natomiast w przypadku korekty błędów w FUR rozmiar funkcjonalny oprogramowania będzie zmieniony.[1]

6. Ocena metody COSMIC

Jak już nadmieniono, celem powstania metody COSMIC było rozszerzenie możliwości zastosowania koncepcji FSM na inne rodzaje oprogramowania niż systemy wspomagające zarządzanie, jednak także z ich uwzględnieniem. Cel ten udało się po części osiągnąć. Praktyka pokazuje bowiem, że opisywana metoda sprawdza się dla (por. p. 2):

- Aplikacji sterowanych danymi, tj. takich produktów programowych, których złożoność wynika przede wszystkim z konieczności zarządzania dużą ilością danych dotyczących zdarzeń mających miejsce w świecie rzeczywistym. Są one zwykle wykorzystywane do wspomagania zarządzania biznesowego.
- Systemów czasu rzeczywistego (sterowanych czasem), zadaniem których jest realizacja lub sterowanie zdarzeniami mającymi miejsce w świecie rzeczywistym. Właśnie w tej kategorii systemów, po licznych i zakończonych sukcesami testach, można obecnie zauważyć znaczny wzrost zainteresowania omawianą metodą.
- Rozwiązań hybrydowych, tj. łączących cechy obu wyżej wymienionych rodzajów oprogramowania, które dzięki omawianej metodzie mogą być wymiarowane w tej samej skali (np. systemy rezerwacji w czasie rzeczywistym biletów lotniczych).

Jednak obecna wersja metody COSMIC nie może być stosowana do wymiarowania rozmiaru funkcjonalnego oprogramowania bądź jego części, które:

- Charakteryzuje się skomplikowanymi algorytmami matematycznymi lub innymi wyspecjalizowanymi i złożonymi zasadami działania, jakie można spotkać na przykład w systemach ekspertowych, oprogramowaniu symulacyjnym, systemach samouczących się czy wspomagających prognozowanie pogody.
- Przetwarza zmienne ciągłe, jak dźwięki audio lub obrazy wideo, które można znaleźć na przykład w grach komputerowych lub oprogramowaniu instrumentów muzycznych.

Ograniczona stosowalność metody COSMIC wynika z faktu, że podproces operowania na danych (por. rys. 8) w obecnie obowiązującej wersji metody nie jest mierzony oddzielnie, a jego funkcjonalność uwzględnia się jedynie w sposób przybliżony - poprzez podproces przepływu danych, z którym jest on powiązany. Toteż metoda COSMIC jest właściwa dla wymiarowania

[1] Zasady obowiązujące przy wymiarowaniu zmian funkcjonalnych w istniejącym oprogramowaniu znajdzie Czytelnik w: [5, s. 59-61].

oprogramowania, w którym występuje dużo przepływów danych (ang. *movement-rich software*), co ma miejsce zwykle właśnie w aplikacjach biznesowych i systemach czasu rzeczywistego, natomiast na obecnym etapie rozwoju nie jest ona właściwa dla wymiarowania oprogramowania, w którym występuje dużo podprocesów operowania na danych (ang. *manipulation-rich software*) [6, s. 21]. Trzeba także zachować ostrożność przy wymiarowaniu bardzo małych części oprogramowania (obejmujących kilka przepływów danych), a zwłaszcza niewielkich zmian w oprogramowaniu, jako że w takim przypadku zasada przybliżonego uwzględniania funkcjonalności podprocesu operowania na danych w ramach odpowiedniego podproces przepływu danych nierzadko zawodzi. Rozważana metoda nie mierzy także wszystkich aspektów funkcjonalności, które mogą być uważane za istotne dla rozmiaru funkcjonalnego oprogramowania, np. wpływu na taki rozmiar liczby atrybutów danych przypadających na przepływ danych [5, s. 10].

Wśród zalet metody COSMIC należy także wymienić:

- reguły postępowania dla oprogramowania wielowarstwowego wyrażono *explicite*, co ułatwia jego właściwe wymiarowanie;
- ukierunkowanie na nowoczesne metody specyfikowania wymagań dla oprogramowania i jego konstrukcji (przede wszystkim na UML);
- pomoc w procesie specyfikacji jednoznacznych, mierzalnych wymagań;
- brak arbitralnych ograniczeń dla rozmiaru procesów funkcjonalnych;
- możliwość pomiaru rozmiaru oprogramowania z różnych perspektyw bazujących na procesach: zarówno z perspektywy użytkownika końcowego, jak i projektanta.

Do słabych stron metody COSMIC, oprócz braku pełnej uniwersalności, zalicza się:

- ujawniające się w praktyce nieporozumienia w identyfikacji warstw i grup danych [26, s. 288];
- ograniczone możliwości przeprowadzenia wczesnych w cyklu życia lub szybkich, ze względu na potrzebę oszczędności czasu i/lub nakładów pracy, przybliżonych obliczeń, co wynika z konieczności bazowania na specyfikacji FUR o dużym poziomie szczegółowości;
- stosunkowo niewiele źródeł poznawczych: przykładów, reguł postępowania w zróżnicowanych środowiskach, historycznych danych empirycznych, narzędzi wspomagających, specjalistów, możliwości szkoleń etc.

W związku z powyższym wśród prac zmierzających do zwiększenia użyteczności metody COSMIC należy wspomnieć przede wszystkim o dążeniach do:

- opracowania jednoznacznie interpretowanej definicji warstwy i kryteriów jej wyróżniania oraz definicji grupy danych;
- stworzenia zasad i procedur umożliwiających trafną ocenę rozmiaru funkcjonalnego oprogramowania w początkowych stadiach cyklu życia;

- odkrycia korelacji występujących pomiędzy wstępną oceną rozmiaru funkcjonalnego a rzeczywiście otrzymanym wynikiem;
- zbadania stopnia powtarzalności rezultatów dla tej samej specyfikacji FUR;
- budowy modeli umożliwiających przeliczenie liczby punktów funkcyjnych COSMIC na przewidywaną pracochłonność przedsięwzięcia, co jest na razie utrudnione ze względu na stosunkowo niewielkie zasoby danych historycznych;
- budowy repozytoriów i narzędzi gromadzących dane historyczne o rozmiarach produktów programowych wyprowadzonych za pomocą opisywanej metody;
- opracowania reguł konwersji rezultatów otrzymanych przy wykorzystaniu różnych metod FSM (zwłaszcza metody IFPUG) na rezultaty wyrażone w punktach funkcyjnych COSMIC i odwrotnie.

7. Metoda COSMIC a metoda IFPUG

Jak już nadmieniono, wszystkie FSMM wywodzą się od koncepcji *Function Point Analysis* stworzonej przez A. Albrechta, toteż istnieją między nimi pewne podobieństwa. Dla metody COSMIC oraz metody IFPUG - stanowiącej metodę pierwszej generacji FSM, jednak jak dotąd najpopularniejszą ze wszystkich metod wymiarowania rozmiaru funkcjonalnego oprogramowania - zalicza się do nich przede wszystkim:

- Wspólną koncepcję wymiarowania rozmiaru funkcjonalnego oprogramowania, opartą na podobnym rozumieniu celu wymiarowania, zakresu wymiarowania oraz definicji granic aplikacji [7]. Ponadto zasady obu metod bazują na zbliżonym, chociaż nieidentycznym, znaczeniu terminów powiązanych z danymi, gdzie odpowiednikami dla takich kluczowych pojęć metody IFPUG, jak: encja, plik i dana elementarna, są w metodzie COSMIC pojęcia: obiekt zainteresowania, grupa danych i atrybut danych. Zbieżna jest również koncepcja transformacji danych, tj. procesów elementarnych/funkcji transakcyjnych (IFPUG) oraz procesów funkcjonalnych (COSMIC), a także użytkowników w ujęciu ogólnym - jako odbiorców funkcjonalności wymiarowanego oprogramowania.
- Występowanie dwóch faz wymiarowania. W obu podejściach zakłada się występowanie fazy identyfikacji elementów, na bazie których wyznacza się rozmiar funkcjonalny, oraz fazy właściwego pomiaru, w której są one odwzorowywane na ów wyrażony liczbowo rozmiar [27, s. 113-114]. W metodzie IFPUG pierwsza z tych faz nie jest wprawdzie opisana *explicite* jako część procesu wymiarowania, jednak zakłada się w niej oparcie pomiaru na specyfikacji FUR, można także w tym celu wykorzystać modele danych, modele funkcji/procesów czy projekty okienek, ekranów, formularzy i raportów. Wobec tych elementów w fazie właściwego pomiaru stosuje się opisane *explicite* zasady tej metody. W metodzie COSMIC z

kolei faza pomiaru odbywa się jedynie na podstawie specyfikacji FUR, co prowadzi do wyodrębnienia różnych warstw oprogramowania. Następnie, po identyfikacji odpowiednich granic, wymagania funkcjonalne użytkownika są dekomponowane na zbiór procesów funkcjonalnych, z których każdy składa się z zestawu przepływów danych. W fazie właściwego pomiaru musi być zidentyfikowany każdy przepływ, na ich podstawie bowiem wyznacza się rozmiar funkcjonalny.

- Zbliżony sposób wyrażania wymagań funkcjonalnych użytkownika. W obu metodach FUR są wyrażane za pomocą podstawowych komponentów funkcjonalnych. W podejściu rozwijanym przez IFPUG wyróżnia się 5 rodzajów BFC: wewnętrzny plik logiczny, zewnętrzny plik komunikacyjny, zewnętrze wejście, zewnętrzne wyjście oraz zewnętrzne zapytanie, natomiast w metodzie COSMIC wyznaczono 4 rodzaje BFC: wejście, wyjście, odczyt i zapis. Nie ma jednakże pomiędzy nimi prostej analogii, jako że w metodzie COSMIC nie wymiaruje się *explicite* danych i nie wyróżnia ich jako rodzaju BFC, chociaż w obu metodach wymiaruje się logiczne transakcje (nieco odmiennie), na co wskazuje schematyczne porównanie obu metod zawarte w tabeli 2.
- Oba podejścia, chociaż w różny sposób, spełniają wymagania nałożone w normie ISO/IEC 14143 na metody FSM, w związku z czym oba uznano za międzynarodowe standardy takiego wymiarowania: metodę IFPUG zawężoną do nieskorygowanych punktów funkcyjnych (norma ISO/IEC 20926), natomiast metodę COSMIC w całości (norma ISO/IEC 19761).

Tab. 2. Porównanie podstawowych komponentów funkcjonalnych metody IFPUG i metody COSMIC

Metoda FSM	Dane	Rozmiar danych	Funkcje transakcyjne	Transa-kcje	Rozmiar transakcji
IFPUG	Plik wewnętrzny	Liczba danych elementarnych	Zewnętrzne wejście		Liczba danych elementarnych
	Plik zewnętrzny	Liczba rekordów elementarnych	Zewnętrzne wyjście		Liczba plików odniesienia
			Zewnętrzne zapytanie		
COSMIC	Nietrwałe	Część procesu funkcjonalnego	Proces funkcjonalny	Wejście	Liczba przepływów danych
				Wyjście	
	Trwałe			Odczyt	
				Zapis	

Źródło: [8, s. 146]

Podstawowe różnice między omawianymi metodami FSM dotyczą głównie:

- Zasad wymiarowania. Zasadniczą różnicą w tym obszarze jest uwzględnienie w metodzie IFPUG generalnych charakterystyk systemu, wyrażających wpływ wymagań technicznych i jakościowych na rozmiar funkcjonalny. Stanowi to przyczynę, z powodu której podejście to nie zostało w całości zaaprobowane przez ISO/IEC, jednakże branie ich pod uwagę w obliczeniach nie jest konieczne. Tego typu charakterystyki nie istnieją w metodzie COSMIC, gdzie wymiarowanie opiera się wyłącznie na wymaganiach funkcjonalnych użytkownika.
- Granic rozmiarów dla procesów. W metodzie IFPUG rozmiar wszystkich pięciu podstawowych komponentów funkcjonalnych jest arbitralnie ograniczony co do wielkości, toteż rozmiar oprogramowania zależy od ich liczebności. W podejściu COSMIC natomiast nie istnieje górna granica rozmiaru procesu funkcjonalnego, jako że rozmiar procesu jest zdeterminowany liczbą przepływów danych. Przy czym z perspektywy poziomu wymiarowania to przepływy danych metody COSMIC bardziej niż jej procesy funkcjonalne odpowiadają procesom elementarnym metody IFPUG (wejściom, wyjściom i zapytaniom), a ich rozmiar wynosi 1 CFP i nie zależy od liczby atrybutów danych ani plików, do których się odnoszą, co ma miejsce w metodzie IFPUG.
- Sposobu uwzględniania danych przechowywanych w systemie. W metodzie rozwijanej przez IFPUG dane uwzględniane są w obliczeniach w dwojaki sposób: osobno jako pliki wewnętrzne/zewnętrzne oraz jako pliki odniesienia wpływające na rozmiar procesu elementarnego (por. tabela 2). W przypadku metody COSMIC dane przechowywane w systemie uwzględniane są przy każdym przepływie danych typu odczyt lub zapis. Dlatego wykorzystanie metody IFPUG wymaga konstrukcji modelu danych, co nie jest niezbędne, chociaż bardzo przydatne w podejściu COSMIC [8, s. 145]. W metodzie IFPUG model danych stanowi także podstawę wczesnych szacunków, podczas gdy w podejściu COSMIC w celu aproksymacji wykorzystuje się model procesów.
- Zasobów danych historycznych. W obecnej wersji największego repozytorium z danymi historycznymi dotyczącymi wymiarowania rozmiaru funkcjonalnego oprogramowania, repozytorium International Benchmarking Standards Group (ISBSG), znajdują się dane dotyczące w niemal 85% produktów programowych wymiarowanych za pomocą metody IFPUG, a tylko w 3% tych mierzonych przy wykorzystaniu metody COSMIC - mimo że repozytorium to nie ogranicza się do przechowywania danych jedynie o aplikacjach biznesowych.

Ponadto w literaturze przedmiotu, w większości przypadków jednak firmowanej przez COSMIC, podkreśla się m.in. następujące cechy tej metody, które mają decydować o jej przewadze nad podstawową metodą pierwszej generacji FSM:

- Większy zakres zastosowania. Metodę IFPUG opracowano w celu wymiarowania systemów wspomagających zarządzanie, jednakże w

odniesieniu do jej obecnej wersji nie tylko rozwijająca ją organizacja, ale także ISO i IEC nie nakładają ograniczeń co do pomiaru innych domen funkcjonalnych (por. p. 2). Tymczasem często wysuwa się argument, iż metoda ta nie sprawdza się w przypadku wymiarowania systemów czasu rzeczywistego – w przeciwieństwie do metody COSMIC [27, s. 115]. Zdaniem autorki, wniosek taki jest zbyt daleko idący, zarówno z teoretycznego, jak i z praktycznego punktu widzenia, chociaż pomiar tego typu oprogramowania za pomocą metody IFPUG jest niewątpliwie bardziej skomplikowany w zestawieniu z metodą COSMIC i w związku z tym w efekcie może być mniej dokładny, co wynika z braku wyrażonych *explicite* reguł postępowania w przypadku oprogramowania wielowarstwowego. Zasady metody IFPUG uwzględniają jednakże systemy czasu rzeczywistego - tak w sposób bezpośredni, jak i pośredni, dotyczący kalkulacji rozmiaru funkcjonalnego [11]. Także w literaturze poświęconej tej metodzie można spotkać praktyczne przykłady jej wykorzystania do właściwego wymiarowania tego typu oprogramowania[1].

- Zgodność z analizą i programowaniem obiektowym. W tym przypadku argumentuje się, że skoro metoda COSMIC powstała znacznie później niż metoda IFPUG, to w przeciwieństwie do niej uwzględnia ona nowoczesne metody opisu FUR i konstrukcji systemów, zwracając przede wszystkim uwagę na podejście obiektowe[2]. Jednak nie dowodzi to w żaden sposób, że nie ma możliwości obliczenia rozmiaru funkcjonalnego przy zastosowaniu obiektowego podejścia do projektowania w oparciu o metodę IFPUG – na taką możliwość wskazują zasady metody, wskazówki jej dotyczące oraz przykłady praktyczne [11, s. 6-8][3].
- Szersza perspektywa wymiarowania. Za pomocą metody IFPUG dokonuje się pomiaru rozmiaru funkcjonalnego z perspektywy użytkownika końcowego, zaś za pomocą metody COSMIC z punktu widzenia użytkownika funkcjonalnego, który obejmuje oprócz użytkownika końcowego także projektantów, a ci dostrzegają inne aplikacje oraz urządzenia, które wchodzą w interakcje z wymiarowanym oprogramowaniem [27, s. 115][4]. Perspektywa ograniczona jedynie do

---

[1] International Function Point Users Group (IFPUG), https://www.ifpug.org/publications/case.htm, Case 4 (9.07.2008). Zainteresowany tym zagadnieniem Czytelnik znajdzie zasadnicze wskazówki dotyczące kalkulacji rozmiaru funkcjonalnego dla systemów czasu rzeczywistego i oprogramowania wbudowanego w: [23, s. 25, 29, 56 i 62].

[2] Common Software Measurement International Consortium (COSMIC), http://www.cosmicon.com/advantagecs.asp (8.07.2008).

[3] Por. także: International Function Point Users Group (IFPUG), https://www.ifpug.org/publications/case.htm, Case 3 (9.07.2008).

[4] Por. także: Common Software Measurement International Consortium (COSMIC), http://www.cosmicon.com/advantagecs.asp (8.07.2008).

użytkownika końcowego faktycznie niesie ze sobą niebezpieczeństwo pominięcia w obliczeniach takiej funkcjonalności, która jest niedostrzegalna dla użytkownika-człowieka, jednak pod warunkiem, iż założy się, że tylko użytkownik będący osobą może być odbiorcą funkcjonalności. Tymczasem uznanie metody IFPUG za zgodną ze standardem ISO/IEC 14143 oznacza, że stosowana przez nią obecnie definicja użytkownika jest zgodna z zawartą w tej normie definicją tego pojęcia, w której przez użytkownika rozumie się nie tylko osobę, ale także rzecz (np. inne aplikacje, urządzenia, sprzęt komputerowy), która wchodzi w interakcje z wymiarowanym oprogramowaniem.

- Możliwość szybszego dostarczenia wyników. Metoda COSMIC bywa uznawana za bardziej intuicyjną, zwięzłą i prostszą w zestawieniu z podejściem IFPUG, co w rezultacie powinno prowadzić do szybszego uzyskiwania rezultatów wymiarowania. Fakt ten jest nie bez znaczenia, gdyż stosowanie metody wymiarowania powinno być efektywne. Jednak nie potwierdzają tego badania, z których wynika, iż nie istnieją znaczące różnice pomiędzy szybkością wymiarowania za pomocą obu metod [8, s. 144]. Co więcej, nawet twórcy metody COSMIC przyznają, że gdy istnieje potrzeba szybkiego przeprowadzenia procesu wymiarowania przy niskiej jakości specyfikacji wymagań użytkownika, to prościej (i przez to szybciej) jest zastosować metodę IFPUG, co wynika z ograniczonego zakresu rozmiarów jej komponentów funkcjonalnych, które łatwiej prawidłowo przewidzieć[1]. W takiej sytuacji zastosowanie metody COSMIC będzie wymagało zaangażowania eksperta w celu uzyskania rezultatu na tym samym poziomie wiarygodności, a to spowoduje większą czasochłonność procesu wymiarowania. Zauważmy, że dotyczy to także możliwości wykorzystania obu metod w celach estymacyjnych, dla których metoda COSMIC wymaga dodatkowo odpowiedniego dostosowania (kalibracji) na poziomie organizacji.

Trudno zatem jednoznacznie przesądzać o przewadze metody COSMIC nad metodą IFPUG – obie mają swoje zalety i wady ujawniające się w określonych obszarach aplikacyjnych, obie mają swoich zwolenników i przeciwników. W najbliższej przyszłości najprawdopodobniej omawiana metoda FSM drugiej generacji nie zastąpi podejścia wywodzącego się bezpośrednio od koncepcji Albrechta. Jak bowiem uważa F. Vogelezang, *„FPA dowiodła, że jest wartościowym narzędziem dla projektantów aplikacji biznesowych i pozostanie takim przez wiele nadchodzących lat. (...) Dla oprogramowania aplikacyjnego, które może być całkowicie zmierzone za pomocą FPA, nie ma potrzeby z niej rezygnować"* [26, s. 288].

---

[1] Common Software Measurement International Consortium (COSMIC), http://www.cosmicon.com/advantagecs.asp (8.07.2008).

8. Podsumowanie

Jedną z zasadniczych przyczyn problemów w szacowaniu kluczowych atrybutów przedsięwzięć rozwoju systemów oprogramowania jest brak jednoznacznej miary rozmiaru produktu takich przedsięwzięć. Ze względu jednak na fakt, iż prawidłowe wymiarowanie rozmiaru produktu stanowi istotny czynnik skutecznej ich realizacji, celowe jest dążenie do poznania, weryfikacji w określonych warunkach projektowych oraz doskonalenia na bazie własnych doświadczeń tych podejść do jego wymiarowania, które są najbliższe spełnienia wymagań postawionych nie tylko przed miarami rozmiaru produktu programowego, ale także przed metodami szacowania kluczowych atrybutów przedsięwzięć. Badania wykazują, że są nimi podejścia służące do kalkulacji rozmiaru produktu w jednostkach jego funkcjonalności, tj. jego tzw. rozmiaru funkcjonalnego [12].

Duże znaczenie przypisywane funkcjonalności produktu programowego doprowadziło do powstania wystandaryzowanej definicji nie tylko wymagań funkcjonalnych użytkownika, ale także rozmiaru funkcjonalnego oprogramowania. Rozmiar funkcjonalny produktu programowego można oszacować i zmierzyć za pomocą odpowiednich metod wymiarowania. Metoda wymiarowania rozmiaru funkcjonalnego produktu programowego powinna opierać się na ściśle określonych zasadach, których zbiór zawarto w sześcioczęściowym standardzie ISO/IEC 14143.

Wśród znormalizowanych metod FSM znajdują się obecnie: metoda punktów funkcyjnych IFPUG (ISO/IEC 20926), metoda punktów funkcyjnych Mark II (ISO/IEC 20968), metoda punktów funkcyjnych NESMA (ISO/IEC 24570), metoda punktów funkcyjnych COSMIC (ISO/IEC 19761) oraz metoda FSM FiSMA (ISO/IEC 29881). Podejścia te różnią się możliwościami wymiarowania oprogramowania w odniesieniu do różnych domen funkcjonalnych. Dlatego przed wyborem określonej metody należy w pierwszym rzędzie ocenić jej adekwatność wobec rodzaju produktu, który ma podlegać wymiarowaniu, zgodnie z zasadami wyboru również zawartymi normie ISO/IEC 14143 i zgodnymi ze standardem ISO/IEC 15939. Zasadniczym celem wszystkich uznanych przez ISO/IEC metod FSM jest wymiarowanie rozmiaru produktu programowego poprzez ilościowe ujęcie funkcjonalności wymaganej przez użytkownika i dostarczonej użytkownikowi w sposób niezależny od technologii implementacji, możliwie prosty i spójny w obrębie różnych projektów i organizacji oraz stosunkowo wcześnie w cyklu życia projektu.

Celem powstania metody COSMIC (pierwotnie w wersji FFP, następnie COSMIC-FFP, a obecnie COSMIC), uważanej za pierwszą metodę FSM drugiej generacji, w całości uznaną przez ISO/IEC, było rozszerzenie możliwości zastosowania koncepcji FSM na te kategorie oprogramowania, których odbiorcą funkcjonalności nie jest człowiek lub nie jest nim tylko człowiek, w tym przede

wszystkim, chociaż nie tylko, na systemy czasu rzeczywistego. Na proces wymiarowania rozmiaru funkcjonalnego produktu przy wykorzystaniu tej metody składa się opracowanie strategii pomiaru, odwzorowanie wymagań funkcjonalnych użytkownika dla mierzonego oprogramowania w FUR wyrażone w formie ogólnego modelu oprogramowania oraz właściwe wymiarowanie, w rezultacie czego uzyskuje się rozmiar funkcjonalny wyrażony w punktach funkcyjnych COSMIC. Metoda ta sprawdza się nie tylko dla systemów czasu rzeczywistego, ale także dla aplikacji biznesowych, a przez to również dla rozwiązań hybrydowych. Jednak nie może być ona stosowana do wszystkich obszarów aplikacyjnych, a także nie mierzy wszystkich aspektów funkcjonalności, które mogą być uważane za istotne dla rozmiaru funkcjonalnego oprogramowania.

Do podobieństw między metodą IFPUG i metodą COSMIC należy zaliczyć przede wszystkim: wspólną koncepcję wymiarowania rozmiaru funkcjonalnego oprogramowania, opartą na podobnym rozumieniu celu wymiarowania, zakresu wymiarowania oraz definicji granic aplikacji; występowanie dwóch faz wymiarowania: fazy identyfikacji elementów wpływających na rozmiar funkcjonalny oraz fazy właściwego pomiaru, w której elementy te są odwzorowywane na ów wyrażony liczbowo rozmiar; zbliżony sposób wyrażania wymagań funkcjonalnych użytkownika, chociaż za pomocą innych BFC co do liczebności i rodzaju; a także spełnienie przez obie metody wymagań normy ISO/IEC 14143 (przez metodę IFPUG w części zawężonej do rozmiaru funkcjonalnego). Obie metody FSM charakteryzują się także dosyć wysokim stopniem skomplikowania, jednak koszty ich zastosowania warto potraktować jako inwestycję w doskonalenie procesów budowy oprogramowania w organizacji.

Podstawowe różnice między tymi dwiema metodami FSM różnych generacji dotyczą głównie: zasad wymiarowania, granic rozmiarów dla procesów, sposobu uwzględniania danych przechowywanych w systemie, sposobu wymiarowania przedsięwzięć polegających na modyfikacji istniejącego oprogramowania oraz zasobów danych historycznych, w których znaczną przewagę mają obecnie dane pozyskane dzięki wymiarowaniu rozmiaru funkcjonalnego za pomocą metody IFPUG. Różnice te są podstawową przyczyną trudności w konwersji wyników uzyskanych za pomocą obu metod.

Nierzadko w literaturze przedmiotu można spotkać mniej lub bardziej kontrowersyjne poglądy mające na celu uzasadnienie, iż metoda COSMIC posiada znaczną przewagę nad metodą IFPUG. Zdaniem autorki niektóre z nich są formułowane zbyt kategorycznie. O ile można się bowiem zgodzić, że metoda COSMIC jest bardziej adekwatna w odniesieniu do wymiarowania systemów czasu rzeczywistego niż metoda IFPUG, o tyle twierdzenie, iż ta druga w ogóle się w tym przypadku nie sprawdza nie znajduje potwierdzenia ani z perspektywy teoretycznej, ani praktycznej. Podobnie kontrowersyjny jest pogląd dotyczący braku zgodności metody IFPUG z analizą i programowaniem obiektowym, a także szerszej perspektywy wymiarowania, która wszak zależy

od przyjętej definicji użytkownika, a ta w obu metodach jest zgodna z definicją zawartą w normie ISO/IEC 14143. Aspekty te są możliwe do uwzględnienia za pomocą metody IFPUG, nawet jeżeli nie są w niej wyrażone *explicite*, chociaż niewątpliwie w sposób bardziej skomplikowany niźli w metodzie COSMIC. Badania nie potwierdzają także, że metoda COSMIC – uznana przez niektórych za bardziej intuicyjną, zwięzłą i prostszą - prowadzi do szybszego uzyskiwania rezultatów wymiarowania. Co więcej, gdy istnieje potrzeba szybkiego pomiaru rozmiaru oprogramowania przy niskiej jakości specyfikacji wymagań użytkownika, to prościej i przez to szybciej jest zastosować metodę IFPUG. Dotyczy to także możliwości wykorzystania obu metod w celach estymacyjnych. Trudno zatem jednoznacznie rozstrzygnąć o przewadze metody COSMIC nad metodą IFPUG.

Rozmiar funkcjonalny produktu programowego uzyskany za pomocą odpowiednich, uznanych przez ISO/IEC metod stanowi wystarczająco obiektywną i wiarygodną podstawę do wyznaczania kluczowych atrybutów przedsięwzięć rozwoju systemów oprogramowania, służąc jednocześnie racjonalizacji procesu zarządzania takimi przedsięwzięciami przez ich wykonawców, jak i decyzji biznesowych podejmowanych przez ich zleceniodawców. Metody estymacji oparte na takim rozmiarze charakteryzują się względnie dużą wiarygodnością. Nie należy jednak zapominać, że istotnym warunkiem ich dużej użyteczności jest dostępność odpowiednich, najlepiej własnych organizacyjnych danych historycznych.

LITERATURA

1. Albrecht A., J.: Measuring Application Development Productivity, Proceedings of IBM Application Development Symposium, Monterey, CA., Oct 14-17, 1979, s. 83-92.
2. Common Software Measurement International Consortium: The COSMIC Functional Size Measurement Method, Version 3.0, Advanced and Related Topics, COSMIC, Québec, December 2007.
3. Common Software Measurement International Consortium: The COSMIC Functional Size Measurement Method, Version 3.0, Documentation Overview and Glossary of Terms, COSMIC, Québec, September 2007.
4. Common Software Measurement International Consortium: The COSMIC Functional Size Measurement Method, Version 3.0, Guideline for Sizing Business Application Software (Version 1.1), COSMIC, Québec, May 2008.
5. Common Software Measurement International Consortium: The COSMIC Functional Size Measurement Method, Version 3.0, Measurement Manual, COSMIC, Québec, September 2007.
6. Common Software Measurement International Consortium: The COSMIC Functional Size Measurement Method, Version 3.0, Method Overview, COSMIC, Québec, September 2007.
7. Cuadrado-Gallego J., Rodríguez D., Machado F., Abran A.: Convertibility between IFPUG and COSMIC Functional Size Measurements, 8th International Conference on

Product-Focussed Software Process Improvement, PROFES 2007, Riga, Latvia, July 2007, s. 273-283.

8. Heeringen van H.: Changing from FPA to COSMIC. A transition framework, Proceedings of Software Measurement European Forum (SMEF) Conference, Roma 2007, s. 143-154.
9. IEEE Std 610.12-1990: IEEE Standard Glossary of Software Engineering Terminology, The Institute of Electrical and Electronics Engineers, Inc., New York, NY, 1990.
10. Ince D., C.: History and industrial application, in: Software Metrics: A Rigorous Approach, N.E. Fenton (ed.), Chapman & Hall, London 1991.
11. International Function Point Users Group: IFPUG Function Point Counting Practices Manual, Release 4.2, Part 1-4, IFPUG, Princeton Junction, NJ, January 2004.
12. International Software Benchmarking Standards Group: The ISBSG Report: Software Project Estimates – How accurate are they?, ISBSG, Hawthorn VIC, Australia, January 2005.
13. ISO/IEC 14143 Information Technology – Software measurement – Functional size measurement – Part 1-6, ISO, Geneva 1998-2007.
14. ISO/IEC 14143-1:2007 Information Technology – Software measurement – Functional size measurement – Part 1: Definition of concepts, second edition, ISO, Geneva 2007.
15. ISO/IEC 14143-6:2006 Information Technology – Software measurement – Functional size measurement – Part 6: Guide for use of ISO/IEC 14143 series and related International Standards, ISO, Geneva 2006.
16. ISO/IEC 15288:2008 Systems and software engineering -- System life cycle processes, ISO, Geneva 2008.
17. ISO/IEC 15939:2007 Systems and software engineering - Measurement process, ISO, Geneva 2007.
18. ISO/IEC 19761:2003 Software engineering – COSMIC-FFP – A functional size measurement method, ISO, Geneva 2003.
19. ISO/IEC 20926:2003 Software engineering - IFPUG 4.1 Unadjusted functional size measurement method - Counting practices manual, ISO, Geneva 2003.
20. ISO/IEC 20968:2002 Software engineering – Mk II Function Point Analysis - Counting practices manual, ISO, Geneva 2002.
21. ISO/IEC 24570:2005 Software engineering – NESMA functional size measurement method version 2.1 - Definitions and counting guidelines for the application of Function Point Analysis, ISO, Geneva 2005.
22. ISO/IEC 29881:2008 Information Technology – Software and systems engineering – FiSMA 1.1 functional size measurement method, ISO, Geneva 2008.
23. Longstreet D.: Function Point Analysis Training Course, Longstreet Consulting Inc., October 2004.
24. St-Pierre D., Maya M., Abran A., Desharnais J., M.: Adapting Function Points to Real-Time Software, "American Programmer" Vol. 10, No. 11, 1997, s. 32-43.
25. St-Pierre D., Maya M., Abran A., Desharnais J., M., Bourque P.: Full Function Points: Counting Practices Manual, Technical Report 1997-04, Université du Québec à Montréal, Montréal 1997.

26. Vogelezang F.: COSMIC Full Function Points. The next generation of functional sizing, Proceedings of Software Measurement European Forum (SMEF) Conference, Rome, March 2005.
27. Xunmei G., Guoxin S., Hong Z.: The Comparison between FPA and COSMIC-FFP, Proceedings of Software Measurement European Forum (SMEF) Conference, Roma 2006, s. 113-114.

Rozdział 13

Budowa efektywnie działającej bazy danych dla informatycznych systemów zarządzania przedsiębiorstwem

Konrad Sztumski
Politechnika Częstochowska
konrad@gmail.com

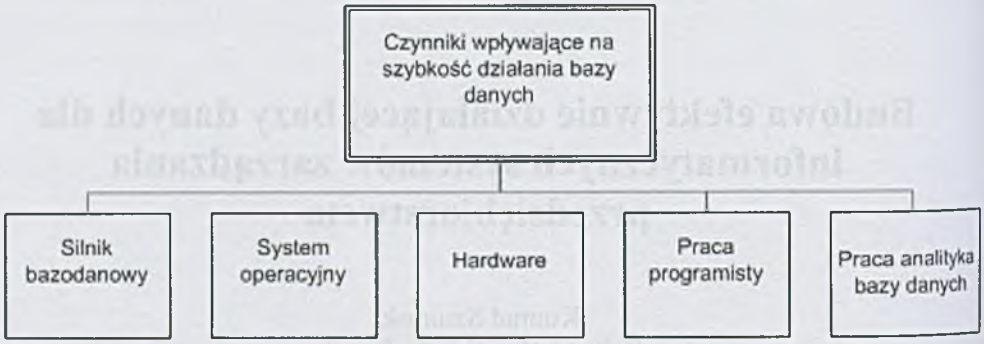
Streszczenie

W rozdziale przedstawiono tematykę efektywności baz danych oraz systemów informatycznych zarządzania.

1. Uwagi wstępne

W ostatnich latach Informatyczne Systemy Zarządzania cieszą się bardzo dużą popularnością. Stąd rosnące zainteresowanie nimi przez różne firmy. Firmy, które chcą stworzyć odpowiednie oprogramowanie muszą brać pod uwagę zarówno aktualne trendy w branży informatycznej jak i wymagania, jakie stawiają im odbiorcy oprogramowania. Bowiem wszystkie programy przeznaczone do obsługi firmy muszą umożliwiać gromadzenie i przechowywanie danych. Producenci wykorzystują bazy danych przede wszystkim do gromadzenia informacji. Wprawdzie już od wielu lat bazy danych wykorzystywane są w Informatycznych Systemach Zarządzania, ale stale dokonujący się postęp w dziedzinie wiedzy z informatyki, doskonalenie silników bazodanowych i osprzętu, wzrost naszych możliwości oraz umiejętności korzystania z informatyki i z oprogramowania stwarzają coraz większe możliwości przyspieszenia rozbudowy i unowocześniania oprogramowań opartych na bazie danych. Często oprogramowanie potrzebne do zarządzania przedsiębiorstwem musi umożliwiać przechowywanie bardzo dużej liczby danych, które dają się przetwarzać przez system informatyczny. Dlatego baza danych musi działać efektywnie i szybko. Chodzi o to, by użytkownikowi końcowemu zapewnić jak najbardziej sprawny dostęp do danych. Szybkość

wykonywania poszczególnych operacji zapewnić można na kilka sposobów pokazanych na poniższym rysunku.



Rys. 1. Czynniki wpływające na szybkość działania bazy danych

Z powyższego schematu widać, że na efektywne, szybkie działanie naszego systemu bazodanowego ma wpływ kilka czynników, za które w jednakowym stopniu odpowiadają: administrator bazy, projektant i programista. Postaram się bardziej szczegółowo przybliżyć i omówić te czynniki, jakimi powinni kierować się tak projektanci, jak i programiści.

- Wybór silnika bazodanowego.
  Ważne jest dobranie odpowiedniego silnika bazodanowego do naszych potrzeb
  i systemu operacyjnego oraz sprzętu, jakim będziemy dysponować, ze względu na to, że różne silniki bazodanowe wykazują różne zachowania w danych zastosowaniach.
- Optymalizacja działania systemu operacyjnego.
  Polega ona na wyłączeniu niepotrzebnych procesów, modułów, instalacji i odpowiedniej konfiguracji firewalli programów antywirusowych. Chodzi o to, by zadbać o jak najlepszą wydajność komputera, na którym zainstalowana jest baza
  i silnik bazodanowy.
- Optymalizacja ze względu na wykorzystywany hardware następuje poprzez modyfikację sprzętu (wymianę na wydajniejszy sprzęt, co wymaga wydatku na zakup nowego sprzętu), dobór architektury i parametrów pracy podzespołów, stosownych dla określonego silnika bazodanowego, system operacyjny
  i przewidywaną potrzebną moc obliczeniową, która daje możliwość obsługi zakładanej przez nas ilości informacji z uwzględnieniem odpowiedniego zapasu.
- Optymalizacja programistyczna.
  Przekłada się ona głównie na przyspieszeniu wyświetlania zapytań w języku SQL

z bazy i na wyborze odpowiedniego środowiska programistycznego (języka programowania), zawierającego interesujące nas biblioteki i komponenty, które umożliwiają i ułatwiają tworzenie oprogramowania współpracującego z bazą danych.

- Optymalizacja analityczna.
  Po określeniu wymagań, jakie stawiane są przed oprogramowaniem, tworzy się projekt bazy danych oraz relacji pomiędzy poszczególnymi polami tabel.
- Do tego wszystkiego może dojść jeszcze odpowiednie ustawienie łącza internetowego (managerów pasma) lub sieci Ethernet w zależności od architektury bazy, od wymagań określonych przez użytkownika i od wymagań samego systemu, który pracuje na naszej bazie danych lub metody pozyskiwania przez aplikację danych z bazy.

W rozdziale zajmę się jedynie sposobami optymalizacji w aspekcie analitycznym
i programistycznym.

Pierwszym krokiem przy tworzeniu bazy danych jest zebranie danych funkcjonalnych od zamawiającego przedsiębiorstwa. Na ich podstawie należy stworzyć schemat relacji pomiędzy danymi obiektami będącymi odwzorowaniem rzeczywistych procesów biznesowych na relacje pomiędzy obiektami bazy. Schemat ten nazywa się diagramem związków encji.

Projektując diagram należy zwrócić uwagę na sposoby gromadzenia danych w tabelach. Aktualnie najpopularniejszym sposobem jest stosowanie relacyjnych baz danych. Cechują się tym, że sposób dostępu do danych i łączenia ich ze sobą jest taki, żeby użytkownik nie musiał widzieć, jak te dane pobrane zostały do komputera[1].

2. Normalizacja danych

Normalizacja danych jest to czynność mająca na celu zmianę układu danych i relacji pomiędzy nimi na takie, które będą eliminować efekty niespójności, redundancji (nadmiarowości), relacji wieloznacznych i problemów przy aktualizacji danych. Niedoświadczony projektant może utworzyć projekt w taki sposób, iż zgromadzi wszystkie obiekty w jednej bazie.

Tab. 1. Przykładowa tabela kontrahentów

Firma	Adres	Produkt	Ilość
Hortex	ul. Mszczonowska 2, 02-337 Warszawa	Sok jabłkowy	500
Dar Natury	Domaniewska 41/7, 02-672 Warszawa	Woda	400
Tymbark	Tymbark 156, 34-650	Sok wiśniowy	155
Hortex	Mszczonowska 2, 02-337 Warszawa	Żurek	260
Dar Natury	ul. Domaniewska 41/7, 02-672 Warszawa	Woda	320
Tymbark	34- 50 Tymbark 156	Sok wiśniowy	155

Pola powyższej tabeli zostały zaprojektowane w nieodpowiedni sposób:
Dodanie kilku zamówień danemu kontrahentowi powoduje, że dane dotyczące firmy powtarzają się w wielu miejscach, a co za tym idzie, zajmują niepotrzebnie miejsce na dysku. Aby dodać zamówienie, trzeba ciągle podawać komplet informacji o kontrahencie. Dane teleadresowe powinny być przetrzymywane w osobnej tabeli,
a każdy ze składników adresu powinien stanowić osobne pole. Tutaj brak powoduje, że adres kontrahenta może zostać wpisany na różne sposoby lub z błędem znakowym. Natomiast jeden standardowy uchroni nas przed błędami. Przy aktualizacji danych kontrahenta trzeba będzie aktualizować każdą krotkę[1] zawierającą jego dane. Usuwając dane o zamówieniu, usuwamy dane kontrahenta. Pola „produkt" i „ilość" powinny zostać przeniesione do nowej tabeli, którą możemy sobie nazwać „zamówienie". Zanim jednak będziemy mieli możliwość znormalizowania powyższej bazy, potrzebne będzie jeszcze wprowadzenie terminu „klucz główny" i „klucz obcy". Klucz główny - kolumna (lub grupa kolumn), zawierająca wartości unikalne dla każdego rekordu znajdującego się w tabeli. W podanej powyżej tabeli kluczem głównym mogłaby być np. para kolumn (Nazwisko, Przedmiot). Ta para wartości jest inna w każdym wierszu. W praktyce często stosuje się inne rozwiązanie. Do tabeli dodaje się dodatkową kolumnę, w której wartości zwiększają się o jeden automatycznie wraz z każdym dodanym rekordem (zazwyczaj typ danych to autonumber lub autoincrement, ew. kolumna musi mieć nadany taki atrybut). Klucz obcy - klucz w tabeli, który odwołuje się do klucza głównego w innej tabeli. Klucz obcy podobnie jak i klucz główny może składać się z kliku kolumn. Dzięki cesze unikalności klucza głównego wymienionej powyżej, mamy pewność, że klucz obcy wskaże nam dokładnie jeden rekord w innej tabeli[2].

[1] Wiersz tabeli

Kontrahenci

nip	Firma	Ulica	Nr domu	Nr mieszk	Miejscowość	Kod
1	Hortex	Mszczonowska	2		Warszawa	02-337
2	Dar Natury	Domaniewska	41	7	Warszawa	02-672
3	Tymbark	Tymbark	156		Tymbark	34-650

Produkty

Id_prod	Produkt
1	Sok jabłkowy
2	Woda
3	Sok wiśniowy
4	Żurek
5	Woda
6	Sok wiśniowy

Zamówienie

nip	Id_produktu	Ilość
1	1	500
2	2	400
3	3	155
1	4	260
2	5	320
3	6	155

Rys 1. Propozycja poprawnie znormalizowanych tabel

Te tabele poprawiają nam wszystkie niedogodności wymienione wcześniej. W tabeli zamówienie identyfikatorami kontrahenta i produktu są liczby, co powoduje,
że można je szybciej porównywać. Można by się pokusić jeszcze np. na rozdzielenie w tabeli „kontrahenci firmy" i „adres", aby ułatwić edycję. Niestety, niekoniecznie może to być potrzebne lub wpłynąć korzystnie na szybkość wyszukiwania danych w bazie. Normalizacja ma wyodrębnionych pięć *postaci* teoretycznych i na postawie ich oraz wymagań funkcjonalnych tworzonej bazy musimy określić potrzebny nam jej stopień.

1. Pierwsza postać normalna **1NF** (*ang. normal form*) – postać najsłabsza; wymaga jedynie, aby dziedziny atrybutów były elementarne (nierozkładalne, atomowe, najprostsze), np. liczby całkowite, daty, łańcuchy, a nie np. listy liczb lub zbiory dat. Rekordy w 1NF są stałej długości.
2. Druga postać normalna **2NF** – jeżeli każdy atrybut Y, który nie jest kluczem zależy funkcyjnie od klucza (a nie od podzbioru atrybutów stanowiących klucz) – po wyznaczeniu wszystkich zależności funkcyjnych.
3. Trzecia postać normalna **3NF** – gdy nie istnieją żadne zależności przechodnie (nie-trywialne).

Postać normalna **Boyce-Codda** (najmocniejsza) – zależności funkcyjne muszą mieć następującą postać: jeżeli $X \rightarrow A$ i atrybut A nie jest zawarty w X, to X jest kluczem lub zawiera klucz.

5. Czwarta postać normalna 4NF – jeżeli zawsze wtedy, kiedy zbiór atrybutów X określa wartościowo Y, to zachodzi jeden z następujących warunków: Y jest puste lub zawiera się w X, suma zbiorów X i Y jest pełnym zbiorem atrybutów lub, wreszcie, X zawiera klucz.
6. Piąta postać normalna 5NF – jeżeli nie istnieje rozkład odwracalny na zbiór mniejszych tabel[3].
W praktyce nie zawsze potrzebne będzie dojście do 5NF. Zwykle normalizacja kończy się na 3NF ze względu na koszt czasu, jaki aplikacja musi poświęcić na przeszukanie bazy.

3. Denormalizacja

Czasami przy bardzo dużej bazie danych i ogromie wykonywanych zapytań możemy przeprowadzić proces denormalizacji. Zmiana ta zaowocuje przyspieszeniem wykonywania się tych zapytań. Realizować ją najlepiej na dwa sposoby:

- poprzez zapis w bazie danych wartości wyliczonych na podstawie innych danych z bazy;
- poprzez zapis klucza obcego nie tylko w bezpośrednio powiązanej tabeli, ale także w "sąsiednich" tabelach[2].

4. Indeksy

Zwykle przeszukując dużą bazę danych z milionami rekordów czas oczekiwania na wykonanie podzapytania byłby bardzo długi. Dane z tabel fizycznie zapisane są w postaci pliku z danymi na dysku. Pliki te czytane są blokami danych. Chcąc wybrać z tabeli szukane przez nas dane, nasz silnik bazodanowy zacznie czytać kolejne bloki pamięci. Zastosowanie indeksów na konkretnym polu spowoduje, że informacja o bloku pamięci, w którym można znaleźć interesujące nas dane zostanie zapisana do pliku. Indeksy stosuje się głownie przy wykonywaniu zapytań SELECT wyszukujących za pomocą WHERE konkretnych pól oraz przy złączeniach tabel[4]. Zatem nasuwa się pytanie: jak stworzyć indeks? W zależności od silnika bazodanowego i typu indeksu tworzy się je na różne sposoby. Nie jestem w stanie opisać wszystkich typów i implementacji przy użyciu silników bazodanowych. Podam tylko jeden przykład. Dla zaspokojenia ciekawości najlepiej sprawdzić to w dokumentacji używanego środowiska bazodanowego.

Przykładowo w środowisku Oracle służy do tego polecenie:

CREATE INDEX nazwa_indexu ON nazwa_tabeli(pola)[1]

Kiedy warto stosować indeksy?
Najlepsze efekty osiąga się przy wybieraniu małej liczby rekordów z dużego zbioru. Zazwyczaj przyjmuje się, że wtedy opłaca się stosować indeksy, gdy z tabeli czytamy nie więcej niż 15% rekordów. Minusem stosowania indeksów jest do trzech razy dłuższe modyfikowanie danych. Spowodowane jest to przez aktualizację informacji w indeksach.

5. Optymalizacja podzapytań SQL

Do przyspieszenia wykonywania zapytań w SQL trzeba tak zmodyfikować polecenie SELECT, aby pracowało na możliwie najmniejszej ilości danych. Nie używajmy raczej

SELECT * from,

tylko określmy potrzebne nam pola. Sytuacje, kiedy potrzebujemy wyświetlić cały rekord z tabeli zdarzają się bardzo rzadko.
Innym sposobem jest filtrowanie danych zawartych już w bazie danych. Za sprawą normalizacji często interesujące dane mamy rozproszone w kilku tabelach. Szybsza i wygodniejsza jest selekcja danych z kilku tabel naraz niż wybieranie z pojedynczych i późniejsza ich obróbka. Za pomocą poleceń specyfikującego kryteria doboru wierszy WHERE. HAVING, ORDER BY możemy otrzymać już interesujące nas dane z jednej lub większej ilości tabel (jak w przykładzie poniżej).

```
SELECT firma, ilosc
FROM kontrahenci, produkty
WHERE kontrachenci.nip=zamowienie.nip;
GROUP BY firma
HAVING ilość>500;
```

Wybieramy z dwóch tabel nazwę firmy i ilość zamówionego towaru, Szukamy tych które zamawiają powyżej 500 sztuk i sortujemy wyniki po nazwach firm. Przedstawiłem tutaj dość proste metody, które mogą się przydać na początkowym etapie pracy z bazami danych przy ich projektowaniu i programowaniu. Zwracam jednak uwagę na to, że opisane tu metody nie zawsze

[1] Jest wiele rodzajów indeksów w Oracle i do poznania wszystkich odsyłam na stronę < http://www.cs.put.poznan.pl/bbebel/sbd_2/18Indeksy.pdf> lub stronę producenta.

znajdują zastosowanie w jakimś konkretnym przypadku i nie zawsze uda nam się optymalizować za pomocą wszystkich wymienionych tu metod.

LITERATURA

1. Hotka D.: Oracle 9i w przykładach. MIKOM, Warszawa 2003, s. 23
2. Internet: http://www.poradnikwebmastera.com/artykuly/bazy_danych/optymalizacja_bazy_danych.php
3. Forkiewicz M.: http://www.zie.pg.gda.pl/md/bazy_danych/przyklady_normalizacja.pdf
4. Krokiewicz M.: Optymalizacja bazy danych, SDJ 9/2008, Warszawa 2008, s. 23.
5. Internet: http://www.zie.pg.gda.pl/md/bazy_danych/przyklady_normalizacja.pdf
6. Krokiewicz M.: Optymalizacja bazy danych, Software Developer Journal
7. Bębel B.:http://www.cs.put.poznan.pl/bbebel/sbd_2/18Indeksy.pdf

Rozdział 14

Zastosowanie normy PN ISO/IEC 12207 do oceny wdrożenia systemu informatycznego

Leszek Grocholski
Instytut Informatyki Uniwersytetu Wrocławskiego
Leszek.Grocholski@ii.uni.wroc.pl

Andrzej Niemiec
Prim Sp. z o.o.
ani@prim.com.pl

Streszczenie

W rozdziale omówiono praktyczne aspekty wykorzystania normy PN ISO/IEC 12207 do oceny wdrożenia systemu informatycznego. Norma może być stosowana dla umów dotyczących oprogramowania, w których nie określono metody zarządzania projektem i wykonawca posiada certyfikat zarządzania systemem jakości zgodny z norma ISO/IEC 9001:2001. Rozdział zawiera opis wykorzystania w/w normy do oceny wdrożenia w szczególności do rozstrzygania sporów nt. przebiegu, zakończenia czy efektów wdrożenia. Ocena taka może mieć następujące zastosowania:

- polubowne rozstrzyganie sporów między zamawiającym a dostawcą,

- przeprowadzanie audytów zabezpieczających interes zamawiającego.

Norma PN ISO/IEC 12207 może również mieć zastosowanie do postępowań sądowych w celu analizy, „zachowania należytej staranność" w sensie artykułu 527 kodeksu cywilnego. Potwierdzony wyrokiem sądowym brak „zachowania należytej staranności" oznacza nie tylko przegranie procesu, zwrot wydatków na wdrożenie systemu i wypłatę odszkodowania ale w przypadku projektu realizowanego jako zmówienie publiczne również wykluczenie na 3 lata z przetargów prowadzonych w oparciu o ustawę Prawo o zamówień publicznych. Norma PN

ISO/IEC 12207 może również znaleźć zastosowanie do audytu procesów produkcji oprogramowania z punktu widzenia zgodności z wymaganą przez NATO normą AQAP 21101.

1. Wstęp

Rozdział to wynik doświadczenia zdobytego podczas oceny wdrożenia systemu informatycznego w jednej ze spółek komunalnych. Zamówienie podlegało ustawie Prawo zamówień publicznych. Przedmiotem umowy były:

1) dostawa, instalacja i wdrożenie systemu informatycznego do obsługi i zarządzania usługami świadczonymi w obiekcie klienta;

2) dostawa i montaż urządzeń niezbędnych do prawidłowego funkcjonowania systemu informatycznego określonego w pkt. 1.

Zakres prac do wykonania w celu realizacji punktu 1 obejmował:

1) dostawę i zainstalowanie systemu informatycznego w sieci komputerowej klienta,

2) dostawę, instalację i uruchomienie urządzeń niezbędnych do prawidłowego działania systemu informatycznego,

3) przeszkolenie personelu klienta w zakresie uruchomienia i eksploatacji zainstalowanego systemu informatycznego i urządzeń.

W punkcie 4 umowy na stronach od 1 do 7 wyspecyfikowano wymagania funkcjonalne systemu informatycznego. Wymagania funkcjonale zostały określone przez podanie nazwy i krótkiego opisu zamawianych funkcji systemu, np.:

„ Sprzedaż – ta część systemu powinna:

umożliwiać korektę faktur,

umożliwiać sprzedaż rabatową,

wydruk raportu kasowego oraz analizy sprzedaży."

W paragrafie 4 pkt. 2 Wykonawca gwarantuje, że przedmiot umowy jest wolny od wad i braków.

Podczas wdrożenia wystąpiły klasyczne problemy:

realizacja trwała już 2 razy dłużej niż to określono w umowie.

nie spełniono części wymagań funkcjonalnych.

wystąpiły błędy wykonania.

Cel ekspertyzy

Zamawiający chciał zbadać po której stronie leży wina. Jeżeli po stronie wykonawcy to opinia miała przyczynić się do:

zakończenia 1 etapu, odstąpieniu od 2 drugiego i udzieleniu gwarancji,

zapłaceniu kar umownych.

2. Metoda oceny

Jako podstawową metodę oceny realizacji umowy przyjęto analizę czy realizacja umowy przebiegała z ogólnie znaną wiedzą na temat zasad realizacji dowolnych przedsięwzięć oraz szczególnych zasad dotyczących przedsięwzięć programistycznych. Analiza polegała na badaniu czy ogólnie przyjęte zasady przebiegu standardowych procesów realizacji przedsięwzięć były przestrzegane przez wykonawcę.

Jest oczywiste, że o realizacji umowy decyduje jej zarządzanie dlatego analizę realizacji umowy przez Wykonawcę wykorzystano jako dziedzinę wiedzy zarządzenia projektami[4,5], w tym zarządzanie projektem informatycznym.

Ocena realizacji umowy polega na porównywaniu procesów realizowanych przez wykonawcę z ogólnie znanymi modelowymi procesami. W trakcie przeprowadzania oceny badano, czy dany proces jest realizowany, a jeśli tak, to czy jest realizowany w przyjęty w inżynierii oprogramowania i systemów zasadami.

Jest to ogólnie znana metoda stosowana w takich standardach oceny dojrzałości procesów jak CMM, CMMI, ISO/IEC 15504 – SPICE, Bootstrap czy Trillium.

Mówiąc inaczej analiza polegała na badaniu czy ogólnie przyjęte zasady zarządzania standardowymi procesami realizacji przedsięwzięć oraz zasady dotyczące zarządzania projektem, projektowania, instalowania, wytwarzana i wdrażania oprogramowania były przestrzegane przez wykonawcę.

W świecie a w szczególności w Unii Europejskiej przyjęto, że gwarancją zabezpieczenie interesów klienta jest posiadanie przez wykonawcę certyfikaty jakości ISO. Wykonawca posiada certyfikat jakości ISO 9001 [1]. W ocenie przebiegu wdrożenia zbadano czy realizacja umowy przebiegała zgodnie z ze zobowiązaniami wynikającymi z posiadania takiego certyfikatu jakości.

Norma jakości PN ISO 9001 dla firm informatycznych została uzupełniona normą PN ISO 90003 [1], która to w odniesieniu do wytwarzania oprogramowania zaleca stosowanie normy PN ISO /IEC 12207. Podobnie standard NATO AQAP 2110 w odniesie do produkcji oprogramowania zaleca stosowanie właśnie normy PN ISO/IEC 12207.Wykonawca - powszechnie ogłasza posiadanie certyfikatu jakości PN ISO/IEC 9001 oraz AQAP 2110 oraz jak wynika z umowy zajmuje się produkcją oprogramowania. Należy zatem przyjąć, że wykonawca, jeżeli tylko zachowuje „należytą staranność zawodową" to musi znać zalecaną przez normy jakości ISO 90003 i AQAP 2110 normę PN ISO/IEC 12207.

Powszechnie uznaje się, że normy ISO stanowią kompendium najlepszych praktyk zawodowych a stosowanie się do ich zaleceń świadczy o zachowaniu

należytej staranności zawodowej. W szczególności ogólna norma ISO/IEC 9001 i normy bardziej szczegółowe: 90003 i 12207 są uważane za zbiór najlepszych praktyk dot. zarządzania projektem na poziome ogólnym oraz specyficznych dla oprogramowania procesów.

Badanie zgodności realizowanych przez wykonawcę procesów z wytycznymi normy PN ISO/IEC 12207 miało na celu stwierdzenie czy wykonawca „zachował należytą staranność". Stwierdzenie „niezachowania należytej staranności" wg kodeksu cywilnego i ustawy Prawo zamówień publicznych może mieć poważne konsekwencje i dlatego ma kluczowe znaczenie do rozstrzygania sporów na drodze polubownej i sądowej. Poniżej przedstawiono podstawowe artykuły aktów prawnych dot. „zachowania należytej staranności".

Kodeks cywilny

Art.354 § 1. Dłużnik powinien wykonać zobowiązanie zgodnie z jego treścią i w sposób odpowiadający jego celowi społeczno-gospodarczemu oraz zasadom współżycia społecznego, a jeżeli istnieją w tym zakresie ustalone zwyczaje - także w sposób odpowiadający tym zwyczajom.
§ 2. W taki sam sposób powinien współdziałać przy wykonaniu zobowiązania wierzyciel.

Art. 355 § 1. Dłużnik obowiązany jest do staranności ogólnie wymaganej w stosunkach danego rodzaju (należyta staranność).
§ 2. Należytą staranność dłużnika w zakresie prowadzonej przez niego działalności gospodarczej określa się przy uwzględnieniu zawodowego charakteru tej działalności.

Art. 471. Dłużnik obowiązany jest do naprawienia szkody wyniklej z niewykonania lub nienależytego wykonania zobowiązania, chyba że niewykonanie lub nienależyte wykonanie jest następstwem okoliczności, za które dłużnik odpowiedzialności nie ponosi.

Art. 472. Jeżeli ze szczególnego przepisu ustawy albo z czynności prawnej nie wynika nic innego, dłużnik odpowiedzialny jest za niezachowanie należytej staranności.

Art. 473. § 1. Dłużnik może przez umowę przyjąć odpowiedzialność za niewykonanie lub za nienależyte wykonanie zobowiązania z powodu oznaczonych okoliczności, za które na mocy ustawy odpowiedzialności nie ponosi.

§ 2. Nieważne jest zastrzeżenie, iż dłużnik nie będzie odpowiedzialny za szkodę, którą może wyrządzić wierzycielowi umyślnie.

„Zachowanie należytej staranności" w ustawie Prawo zamówień publicznych

Art. 24.

Z postępowania o udzielenie zamówienia wyklucza się wykonawców, którzy w ciągu ostatnich 3 lat przed wszczęciem postępowania wyrządzili szkodę nie wykonując zamówienia lub wykonując je nienależycie a szkoda ta nie została

dobrowolnie naprawiona do dnia wszczęcia postępowania, chyba że niewykonanie lub nienależyte wykonanie jest następstwem okoliczności, za które wykonawca nie ponosi odpowiedzialności.

Jeżeli zatem wykażemy przed sądem niezachowanie przez wykonawcę należytej staranności i wyrządzenie szkody to prawdopodobnie wykonawca poniesie karę. Realna perspektywa poniesienia przez wykonawcę takiej kary poparta dowodem – oceną wdrożenia – może uchronić zamawiającego przed odbiorem nienadającego się do eksploatacji systemu i w zdecydowany sposób zdyscyplinować wykonawcę. Pozwala również mieć nadzieję, że wykonawca będzie starał się zachować należytą staranność i zakończy wdrożenie sukcesem.

3. Omówienie normy PN ISO/IEC12207

Norma jakości PN ISO 9001 dla firm informatycznych została uzupełniona normą PN ISO 90003, która to w odniesieniu do wytwarzania oprogramowania zaleca stosowanie normy PN ISO/IEC 12207 – procesy życia oprogramowania. Wykonawca - powszechnie ogłasza posiadanie certyfikatu jakości PN ISO/IEC 9001 oraz AQAP. Jak wynika z umowy zobowiązał się dostarczyć dedykowane oprogramowania. Standard NATO AQAP w odniesie do produkcji oprogramowania zaleca stosowanie właśnie normy PN ISO/IEC 12207. Należy zatem przyjąć, że jeżeli tylko wykonawca zachowuje „należytą staranność zawodową" to zna zalecenia norm jakości, której certyfikat posiada i czyli zna i stosuje normę PN ISO/IEC 12207.

Norma ISO/IEC PN 12207 – procesy życia oprogramowania – podaje zalecenia dotyczące podstawowych procesów. Poniżej wyspecyfikowano procesy mające znaczenie dla realizacji omawianej umowy.

Procesy organizacyjne:

zarządzanie przedsięwzięciem,
zapewnienie infrastruktury,
doskonalenie,
szkolenie.

Procesy podstawowe

zakup,
dostawa,
wytwarzanie,
eksploatacja,
utrzymanie.

Procesy wspomagające, wyróżniamy tutaj:

dokumentowanie,
weryfikacja,

walidacja,
zarządzanie konfiguracją,
rozwiązywanie problemów,
zapewnienie jakości,
wspólne przeglądy,
audyt.

W procesie wytwarzania wyróżniamy następujące aktywności:

określenie wymagań i przyjęcie założeń,
utworzenie koncepcji i projektowanie architektury,
projektowanie na poziomie szczegółowym,
realizacja (kodowanie i testowanie),
testowanie kwalifikacyjne,
instalacja oprogramowania,
wspieranie akceptacyjne oprogramowanie.

4. Ocena zgodności z normą PN ISO/IEC 12207

Ocena polegała na badaniu czy opisane w normie PN ISO/IEC 12207 ogólnie zasady przebiegu standardowych procesów i aktywności życia oprogramowania były przestrzegane przez wykonawcę. Poniżej przedstawiono omówienie, w kontekście wytycznych normy PN ISO/IEC 12207, realizacji przez wykonawcę wybranych procesów związanych z realizacją umowy

Zarządzanie projektem

Podstawą zarządzania nie tylko wg normy PN ISO/IEC 12207 jest opracowanie planu i harmonogramu realizacji projektem. Plan obejmuje spis oraz opis etapów realizacji, kolejność ich wykonania, niezbędne do realizacji zasoby i odpowiedzialne za nie osoby. Harmonogram podaje dla poszczególnych procesów daty rozpoczęcia i zakończenia poszczególnych aktywności i ich składowych - kroków .

Zarządzanie polega na planowaniu procesów i aktywności, sprawdzaniu czy realizowane są zgodnie z planem, kosztami i harmonogramem oraz podejmowaniu decyzji o przejściu do kolejnego etapu lub ew. wprowadzaniu do planu modyfikacji.

Zarządzanie obejmuje działania i zadania kierownictwa związane z realizacją określonych procesów (np. nabywania, wytwarzania, eksploatacji ...).

W skład procesu zarządzania wchodzą: inicjacja i definicja zakresu zarządzanego procesu, planowanie, wykonanie i kontrola, przeglądy i ocena oraz zakończenie.

Plan realizacji i harmonogram oraz ich modyfikacje powinny być przedstawiane wykonawcy do wglądu i ew. zaakceptowania.

Wykonywanie procesów i aktywności przedsięwzięcia musi być monitorowane. Nie wystarczy przedsięwzięcie zaplanować ale trzeba sprawdzać jak jest wykonywane i ew. wprowadzać korekty.

Realizacja przez wykonawcę.

Mimo, że realizacja umowy zgodnie z umowa miała trwać dłużej niż 3 miesiące. Wykonawca nie przedstawił:

szczegółowego planu realizacji umowy, w tym planów realizacji konkretnych

procesów,

szczegółowego harmonogramu.

Bardziej szczegółowo planowany były jedynie następny etap.

Dostawa

Wytyczne normy PN ISO/IEC 12207

Proces dostawy obejmuje aktywności i zadania wykonawcy. Proces może być inicjowany decyzją o przygotowaniu oferty w odpowiedzi na ofertę przetargową lub podpisaniem i zakontraktowaniem u nabywcy dostarczenia systemu, produktu programowego lub usługi programowej. Proces trwa począwszy od określenia procedur i zasobów potrzebnych do zarządzania i zapewnienia realizacji projektu, włączając w to opracowanie koncepcji, dokumentacji oraz planów wykonania, skończywszy na przekazaniu systemu, produktu programowego lub usługi programowej nabywcy.

Wykonawca zarządza procesem dostawy na poziomie przedsięwzięcia postępując zgodnie z procesem zarządzania, który jest powołany do życia w tym właśnie procesie. Jeśli nie postanowiono w umowie, to wykonawca określa lub wybiera model cyklu życia oprogramowania odpowiedni dot. zakresu, rozmiaru i złożoności przedsięwzięcia. Wykonawca wybiera i odwzorowuje na model cyklu życia procesy, aktywności i zadania normy PN ISO/IEC 12207.

Wykonawca ustala infrastrukturę, dostosowuje proces zarządzania do przedsięwzięcia; wreszcie zarządza procesem na poziomie organizacyjnym postępując zgodnie z procesem doskonalenia i procesem szkolenia.

Realizacja przez wykonawcę

Proces dostawy był realizowany w sposób tylko częściowo zaplanowany. Plany i częściowe harmonogramy dotyczyły tylko kolejnego kroku. Analiza wykazała brak albo tylko bardzo elementarne zarządzanie podstawowymi procesami niezbędnymi do efektywnego zarządzania realizacją umowy.

Określenie wymagań i przyjęcie założeń

Wytyczne normy PN ISO/IEC 12207

Zamówienie, (ew. umowa) zawiera opis podstawowych cech dot. produktu:

wymagań funkcjonalnych,

wymagań technicznych,

założeń.

Przystępując do wytwarzania obowiązkiem wykonawcy jest określenie wymagań lub wsparcie aktywności prowadzących do ich określenia. Specyfikacja wymagań ma być udokumentowana. Specyfikacja wymagań powinna być zaakceptowana przez zamawiającego w procesie wspólnych przeglądów.

Realizacja przez wykonawcę

Wykonawca nie określił dokładnie wymagań zamawiającego.

W szczególności wykonawca nie dokonał analizy wszystkich wymagań określonych w umowie. Źle określił wymagania dotyczące czytników sterujących urządzeniami kontroli dostępu czego rezultatem była niemożliwość efektywnego korzystania z fragmentu systemu służącego do zarządzania dostępem.

Wykonawca nie przeprowadził szczegółowej analizy wymagań dotyczących funkcji systemu zapisanych umowie, np.:

panelu internetowego,

internetowej rezerwacji usług,

statystyk.

Zarzut wykonawcy o niedostarczeniu przez zamawiającego szczegółowych opisów wymagań w szczególności przebiegu odpowiednich procesów biznesowych jest bezpodstawny. Przeprowadzenie dokładnej analizy wymagań jest obowiązkiem wykonawcy.

Opracowanie koncepcji i projektowanie architektury

Wytyczne normy PN ISO/IEC 12207

Wykonawca powinien przedstawić do akceptacji koncepcje wraz z opisem architektury produktu, który ma być wytworzony. Koncepcja przed rozpoczęciem szczegółowego projektowania musi być zatwierdzona przez zamawiającego.

Realizacja przez wykonawcę

Nie przedstawiono dokumentu zwierającego koncepcję realizacji przez system wymagań określonych w umowie oraz opisu architektury systemu.

Wykonawca w przypadku pewnych wymagań posługiwał się koncepcją prototypu, tzn. przedstawił takowy i sugerował, że tak system powinien działać.

Sposób dziania pro typu nie był opisany w żadnym dokumencie i pisemnie zaakceptowany przez zamawiającego.

Projektowanie

Wytyczne normy PN ISO/IEC 12207

Celem projektowania jest opracowanie projektu stanowiącego wzór wg którego należy wykonać przedmiot zamówienia. Powinien on być tak wykonany aby na jego podstawie można było, bez zadawania zamawiającemu pytań wykonać przedmiot zamówienia. Projekt powinien, co najmniej zawierać opis realizacji wymagań funkcjonalnych i niefunkcjonalnych. W praktyce oznacza to opracowanie opisu zawierającego listy elementów składowych i rysunki, które pokazują jakie elementy dostarczonego produktu realizują poszczególne wymagania funkcjonalne. Taki spis należy uzupełnić opisem i schematami połączeń składowych produktu i dokumentem stwierdzającym jakie normy techniczne i przepisy zostały uwzględnione podczas projektowania produktu.

Realizacja przez wykonawcę.

Nie przedstawiono pełnego projektu zamówionego systemu. Dostarczona dokumentacja dotyczyła systemu o częściowo podobnej funkcjonalności, który nie realizuje wszystkich wymagań zamawiającego. Zawierała natomiast opis funkcji, które nie dotyczą zmówienia – np. zarządzanie łóżkami Dostarczona dokumentacja techniczna systemu informatycznego jest niepełna. W szczególności:

nie przedstawiono, które fragmenty systemu i jak realizują odpowiednie wymagania (funkcje menu),

nie przedstawiono projektu bazy danych.

Do dokumentacji systemu nie wprowadzono zmian zaistniałych w procesie wytwarzania.

5. Ocena realizacji umowy

Konsekwencją posiadania certyfikatu jakości jest obowiązek wykonawcy umożliwienia zamawiającemu dokonanie audytu mającego na celu sprawdzenie czy realizacja umowy przebiega zgodnie z procedurami zapisanymi w dokumentacji systemu jakości. Audytor powinien mieć zapewniony dostęp do odpowiednich zapisów dot. realizacji umowy. W przypadku omawianej umowy zamawiający posiada certyfikat systemu zarządzania jakością ISO 9001 w zakresie świadczenia usług a wykonawca w ramach umowy miał dostarczyć system, który miał umożliwić świadczenie takowych.

Wykonawca powinien przedstawić zamawiającemu wynikające z posiadania certyfikatu jakości zapisy dotyczące realizacji umowy. Umożliwia to stwierdzenie czy Wykonawca wywiązał się w stosunku do zamawiającego ze

zobowiązań wynikających z posiadania certyfikatu systemu jakości ISO 9001, w tym udzielonych gwarancji dążenia do osiągnięcia spełnienia wymagań klienta.

Oprócz analizy zgodności przebiegu standardowych procesów wytwarzania oprogramowania opisanych w normie ISO IEC 12207 były przestrzegane przez wykonawcę, ocena realizacji umowy przebiegała również na analizie czy ogólne zasady:

- zarządzania systemem zapewnienia jakości wg PN ISO 9001,
- zarządzania przedsięwzięciem programistycznym [4],
- etyki zawodowej [5]

były przestrzegane przez wykonawcę.

Przeprowadzona analiza umożliwiła sformułowanie następujących wniosków dotyczących realizacji umowy:

- Wykonawca nie przestrzegał ogólnie znanych zasad dot. zarządzania projektami.
- Wykonawca nie uwzględnił specyfiki realizacji przedsięwzięć programistycznych.
- Dostarczony system informatyczny posiada wady uniemożliwiające samodzielną eksploatację przez zamawiającego.
- Wykonawca nie przestrzegał zasad etyki zawodowej.

Przedstawione powyżej wnioski umożliwiają stwierdzenie, że wykonawca nie zachował należytej staranności.

6. Podsumowanie

W rozdziale, na przykładzie przeprowadzonego rzeczywistego audytu, omówiono wykorzystanie normy PN ISO/IEC 12207 do oceny wdrożenia systemu informatycznego. Zaproponowane podejście może mieć zastosowanie do oceny wdrożenia oprogramowania przez firmy, które posiadają certyfikat zgodności systemu zarządzania z normą PN ISO 9001. Firma, która posiada taki certyfikat, jeżeli tylko dochowuje należytej staranności, powinna znać i stosować wytyczne dotyczące realizacji procesów życia oprogramowania opisane w normie PN ISO/IEC 12207. W omawianym przykładzie uzasadniono, że wdrażająca oprogramowanie firma nie stosowała, wynikających z normy PN ISO/IEC 12207, większości zasad zarządzania projektem obejmującym procesy życia oprogramowania. co zostało uznane za niezachowanie należytej staranności. Taki wynik oceny wdrożenia umożliwił polubowne zakończenie sprawy. Wykonawca uznał, że wina leży po jego stronie i zgodził się zapłacić

kary umowne i zamienić umowę serwisu na udzielenie gwarancji. Obie strony zgodziły się odstąpić od realizacji drugiego etapu umowy.

LITERATURA

1. PN ISO 9001 i PN ISO/IEC 90003
2. PN ISO/IEC 12207
3. Kerzner H.: Advanced Project Management edycja polska, Wydawnictwo Helion 2005.
4. McGary R, Wysocki R.K.: Efektywne Zarządzanie projektami, Wydawnictwo Helion 2005.
5. Somerville I.: Podstawy inżynierii oprogramowania, PWNT 2004, wyd.3.

Rozdział 15

Ocena efektywności systemu informatycznego zaprojektowanego przy aktywnym wsparciu systemu ekspertowego

Zbigniew Buchalski
Politechnika Wrocławska
zbigniew.buchalski@pwr.wroc.pl

Streszczenie

W rozdziale przedstawiono pewną koncepcję informatycznego wsparcia procesu projektowania systemu informatycznego dla różnych grup użytkowników. Podano opis systemu ekspertowego SYSTEMINF, który wspiera ten proces, jego strukturę oraz poszczególne etapy budowy systemu. Implementacja komputerowa systemu SYSTEMINF zbudowana została w oparciu o szkieletowy system ekspertowy EXSYS Corvid. Zaprezentowano przebieg procesu wnioskowania przy wykorzystaniu systemu SYSTEMINF oraz wyniki badań testujących ten system.

1. Wstęp

Ostatnie lata przyniosły gwałtowny rozwój specjalistycznych systemów komputero-wych zawierających w sobie wiedzę ekspercką [2, 3, 4, 5, 6, 7]. Dziedzina systemów ekspertowych obejmuje obszar zagadnień technicznych oraz metodologicznych, zmie-rzających do użycia komputerów przy rozwiązywaniu złożonych problemów decyzyj-nych [1, 7, 8].

Systemy ekspertowe można obecnie spotkać prawie na każdym kroku począwszy od medycyny poprzez technikę do podejmowania skomplikowanych decyzji finanso-wych. Przestały one być już wyłącznie domeną naukowców i laborantów zajmujących się badaniami w dziedzinie sztucznej inteligencji, stały się powszechnie wykorzystywane. Zastosowanie systemu ekspertowego może przynieść znaczne korzyści firmie, która pokusi się o jego wdrożenie.

Rozwój techniki mikroprocesorowej doprowadził do tworzenia systemów

eksperto-wych na relatywnie tanim i ogólnie dostępnym sprzęcie komputerowym, dzięki czemu możliwy jest gwałtowny wzrost wykorzystania systemów ekspertowych w praktyce Są one z powodzeniem stosowane w roli systemów diagnostycznych, doradczych, prognozujących, klasyfikujących i monitorujących.

Proces podejmowania decyzji przy projektowaniu systemu informatycznego jest zorganizowanym, realizowanym na zasadzie algorytmu zestawem czynności, którego za-daniem jest precyzyjne określenie warunków i ograniczeń sytuacji decyzyjnych oraz dokonanie wyboru optymalnego wariantu. Sprawność i skuteczność podejmowania decyzji jest kluczowym czynnikiem sukcesu każdego przedsięwzięcia. Istotną rolę we wspomaganiu procesu decyzyjnego odgrywa zaprezentowany w niniejszym rozdziale pewien system ekspertowy o nazwie SYSTEMINF, który wspiera projektanta systemu informatycznego w doborze odpowiednich urządzeń w celu stworzenia takiego systemu, który odpowiadałby potrzebom użytkowników tego systemu.

2. Czynniki wpływające na kształt systemu informatycznego

Współczesny świat każdego dnia zalewa nas tysiącami informacji przede wszystkim za pośrednictwem różnorodnych mediów elektronicznych. Każda w zasadzie gałąź nauki i techniki została trwale związana z nieustannie rozwijającą się komputeryzacją. Dziś nie potrafilibyśmy wyobrazić sobie życia bez otaczającej nas techniki informatycznej. W miejscu pracy i przy wykonywaniu codziennych czynności posługujemy się wieloma elementami systemu informatycznego, jak i korzystamy z wielu nowoczesnych technologii informatycznych. Dodatkowo co chwila pojawiają się nowe rozwiązania informatyczne.

System informatyczny – jest to zbiór powiązanych ze sobą elementów, którego fun-kcją jest przetwarzanie danych przy użyciu techniki komputerowej. Na systemy infor-matyczne składają się takie elementy, jak:

- sprzęt, czyli komputery oraz urządzenia służące do przechowywania danych,
- urządzenia służące do komunikacji pomiędzy sprzętowymi elementami systemu,
- urządzenia służące do komunikacji między użytkownikami a systemem informa-tycznym,
- urządzenia służące do odbierania danych ze świata zewnętrznego (np. czujniki elektroniczne, kamery, skanery),
- urządzenia służące do wpływania systemów informatycznych na świat zewnętrzny - elementy wykonawcze (np. silniki sterowane

komputerowo, roboty przemysłowe, sterowniki urządzeń mechanicznych, itp.),

- urządzenia służące do przetwarzania danych nie będące komputerami,
- oprogramowanie,
- elementy organizacyjne, czyli procedury korzystania z systemu informatycznego, instrukcje robocze, itp.,
- elementy informacyjne, bazy wiedzy – ontologie dziedzin, w których używany jest system informatyczny, np. podręcznik księgowania w przypadku systemu finansowo-księgowego.

Czynniki, które wpływają na wygląd systemu informatycznego i są istotne przy jego projektowaniu można podzielić na trzy następujące grupy:

Przeznaczenie danego systemu informatycznego.

Z pracą biurową (edycja tekstów, drukowanie, arkusze kalkulacyjne, itp.) poradzi sobie praktycznie każdy komputer, nawet o najsłabszych parametrach. Podobnie jest z przeglądaniem stron WWW czy odtwarzaniem multimediów. Jeżeli komputer ma wy-konywać bardziej wymagające prace, np. pracować jako serwer lub sprzęt do obróbki wideo i audio – potrzebne są wydajniejsze podzespoły, co pociąga za sobą większe wydatki. Potrzebna jest odpowiednia równowaga między wydajnością a kosztami. Należy odpowiednio dobierać pojedyncze elementy zestawu komputerowego tak, aby spełniał on odpowiednie wymogi a jednocześnie nie przekroczył środków finanso-wych przeznaczonych na jego zakup.

Różnorodność urządzeń i podzespołów w danej kategorii.

Wszystkie podzespoły występują w kilku wersjach, których ilość jest często bardzo duża i utrudnia znacznie wybór właściwej. Taka różnorodność wynika z parametrów tychże elementów, które sprawiają, że jedne są bardziej wydajne od innych, różnią się też rodzajem zastosowania i oczywiście ceną.

Mnogość producentów podzespołów i różnorodność wśród podzespołów tego sa-mego typu.

Firm produkujących podzespoły jest mnóstwo. Każda posiada swoją markę na rynku elektronicznym, która mówi o jakości oferowanych wyrobów. Jedno urządzenie może cechować się różnorakimi parametrami technicznymi takimi, jak: ilość, rodzaj i szybkość pamięci, różne zegary taktujące układ (np. karty graficzne), szybkość transferu i wielkość bufora pamięci (dyski twarde), które wpływają na wydajność i jednocześnie koszt elementu.

3. Struktura systemu SYSTEMINF

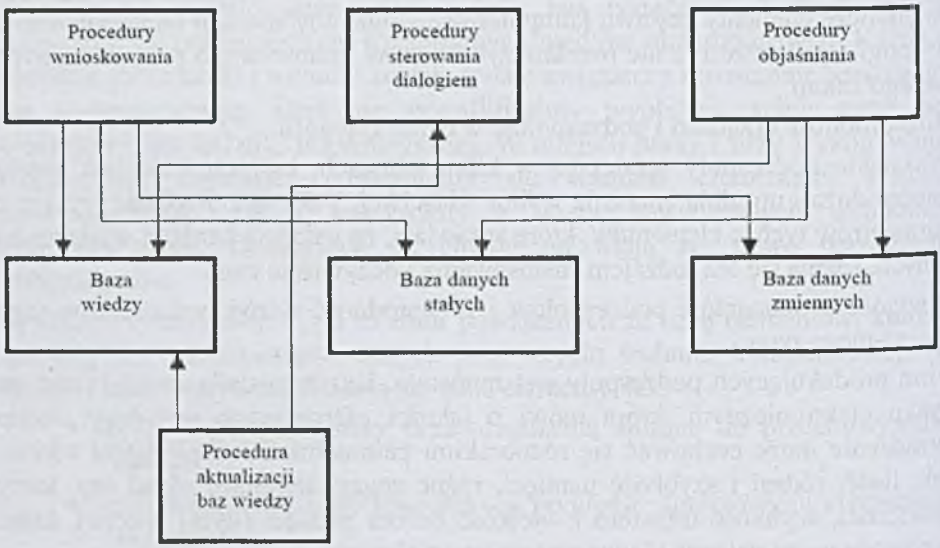
Podstawowym komponentem każdego systemu ekspertowego jest baza wiedzy. Wiedza jest pojęciem podstawowym dla różnego rodzaju procesów decyzyjnych i procesów wnioskowania zarówno przez człowieka, jak i przez komputer.

Baza wiedzy systemu SYSTEMINF składa się z kilku części. Podstawową część stanowi zbiór reguł z dziedziny wiedzy dotyczącej systemów informatycznych. Drugą część bazy wiedzy stanowi zbiór faktów. Pomiędzy zbiorem reguł i zbiorem faktów zachodzą określone relacje.

W strukturze systemu SYSTEMINF możemy wyróżnić następujące elementy:

- baza wiedzy (np. zbiór reguł i faktów),
- baza danych (np. dane obiektu, wyniki pomiarów, hipotezy, itd.),
- procedury wnioskowania – maszyna wnioskująca,
- procedury objaśniania – objaśniają strategie wnioskowania,
- procedury sterowania dialogiem – procedury wejścia/wyjścia umożliwiają formu-łowanie zadań przez użytkownika i przekazywanie rozwiązania przez system eks-pertowy,
- procedury umożliwiające rozszerzenie oraz modyfikację wiedzy – pozyskiwanie wiedzy.

Uwzględniając te elementy strukturę systemu ekspertowego można przedstawić w postaci następującego schematu:



Rys. 1. Podstawowe elementy systemu SYSTEMINF

4. Szkieletowy system ekspertowy EXSYS Corvid

Szkieletowy system ekspertowy EXSYS Corvid jest produktem firmy EXSYS Inc. EXSYS Corvid jest narzędziem przeznaczonym do szybkiego tworzenia systemów ekspertowych na stronach HTML. Jego funkcjonalność jest

zdeterminowana trzema nowymi rozwiązaniami, które według producenta mają zrewolucjonizować tworzenie systemów ekspertowych w sieci. Są to: struktura obiektowa (ang. object structure), bloki logiczne (ang. Logic Blocks) oraz aplety Java.

Struktura obiektowa to nic innego, jak obiektowe podejście do zmiennych przy zacho-waniu podejścia strukturalnego w projektowaniu systemów decyzyjnych. Corvid wy-korzystuje znany z Visual Basic'a model programowania, czyli połączenie obiekto-wości z podejściem strukturalnym.

W poprzednich wersjach EXSYS logika była reprezentowana przez drzewa logiczne i indywidualne reguły IF/THEN. Niestety, dużo systemów wymaga wielu drzew decyzyjnych, co przysparza kłopotów przy organizacji logiki. Bloki logiczne wprowadzone w EXSYS Corvid mają za zadanie wspomóc projektanta systemu ekspertowego właśnie przy organizacji logiki. Mogą nimi być różne zestawy reguł lub cała baza wiedzy. Pozwala to na uporządkowanie wiedzy w bloki, które zachowują się jak obiekty.

Corvid pozwala na przesyłanie systemu poprzez aplety Javy, natomiast komunikacja z użytkownikiem realizowana jest poprzez strony HTML. Aplikacje tworzone w Corvi-dzie przesyłane są w ok. 100kB apletach. Twórca systemu ekspertowego wybiera tyl-ko pytanie i sposób jego zadania. Korzystając z możliwości HTML'a wyniki wniosko-wania można umieszczać w odpowiednio sformatowany sposób, dzięki czemu system staje się czytelniejszy.

Aplet Corvid'a zapewnia dużą funkcjonalność dla większości systemów ekspertowych. Jeśli jednak zajdzie potrzeba na dodanie nowej funkcji, to istnieje możliwość komunikowania się z innymi apletami na stronie. Jeśli chodzi o pobieranie lub obliczanie danych dostępnych tylko z serwera, to można wykorzystać skrypty CGI, ASP i JSP. Funkcjonalność sieciowa jest najmocniejszą stroną tego produktu wykorzystywaną przez największe firmy.

EXSYS Corvid tak, jak każdy szkieletowy system ekspertowy, pozwala na automa-tyzację pracy w dziedzinie zapotrzebowania na wiedzę. Na pewno na plus należy zaliczyć Corvid'owi to, że zminimalizował w pełni działający w Internecie system ekspertowy do ok. 100KB pamięci, przy czym aplet ściągany jest tylko raz. Wystarczy niewielka ilość kodu HTML i można stworzyć funkcjonalny system ekspertowy. Jednak nie przeceniajmy wartości tegoż produktu. Niemożliwa jest realizacja sytemu ekspertowego od początku do końca w oparciu o edytor tekstu, tak jak można to zro-bić w języku programowania LISP. Dla zaawansowanych użytkowników i programis-tów tworzenie skomplikowanych projektów będzie poszukiwanie wielu opcji w dużej liczbie okien. Nie sprzyja to bynajmniej oszczędzaniu czasu. Trzeba jednak przyznać, iż zastosowane rozwiązania są zarówno nowatorskie, jak i przejrzyste.

5. Opis systemu SYSTEMINF

Implementacja komputerowa systemu SYSTEMINF powstała w oparciu o narzędzie informatyczne EXSYS Corvid omówione w poprzednim punkcie. Proponowany system ma za zadanie wspomóc projektanta systemu informatycznego w doborze odpowiednich elementów tego systemu w zależności od potrzeb jego użytkowników, jednocześnie uwzględniając koszt realizacji tego przedsięwzięcia. Wybór tej dziedziny zastosowania systemu ekspertowego można uzasadnić ciągłym zapotrzebowaniem na sprzęt komputerowy i urządzenia dodatkowe, który nieustannie ulega modernizacjom, zmianom pod względem możliwości, wydajności i zastosowań. Dobór właściwych elementów systemu informatycznego nie jest prosty, a niekiedy wręcz niemożliwy dla przeciętnego człowieka bez odpowiedniej wiedzy i doświadczenia bądź bez konsultacji ze specjalistą. Narzędzie jakim jest system ekspertowy wydaje się idealnym rozwiązaniem tego problemu.

Budowa systemu ekspertowego SYSTEMINF została podzielona na kilka etapów. Proces ten był długotrwały i pracochłonny. Główny nacisk położony został na dobór podzespołów komputerowych oraz urządzeń peryferyjnych systemu informatycznego.

Aby określić charakterystykę problemu budowy systemu informatycznego należało przeanalizować dostępne podzespoły i urządzenia komputerowe oraz zbadać ich użyteczność w zależności od zastosowania.

Wydzielono wstępnie kilka abstrakcyjnych grup zastosowań systemów informatycznych na podstawie doświadczeń własnych twórcy systemu ekspertowego, na podstawie analiz for dyskusyjnych, portali internetowych poświęconych sprzętowi komputerowemu oraz literatury. Najczęściej występujące zastosowania systemów informatycznych, to:

- komputer biurowy (podstawowe zastosowanie to arkusze kalkulacyjne, edytory tekstu i aplikacje typowo biurowe, jak np. Płatnik),
- komputer domowy (zastosowanie podobne jak komputer biurowy oraz zwiększone wymagania co do wydajności ze względu na pracę z wieloma aplikacjami na raz),
- komputer multimedialny (dodatkowe zastosowania multimedialne, jak oglądanie filmów, słuchanie muzyki, gry komputerowe, przeglądanie stron WWW, itp.),
- komputerowe centrum rozrywki (posiadający wszystkie nowinki technologiczne, bardzo wydajny i szybki komputer do najnowszych wymagających gier 3D, dający sobie radę z obróbką audio i wideo, itd.),
- serwery (magazyny danych posiadające rozbudowane macierze dyskowe, wieloprocesorowe).

Analiza urządzeń i podzespołów pod względem ich podziału i przypisania do danych grup zastosowań doprowadziła do wyciągnięcia szeregu wniosków. Podstawowe wnioski to:

- ogromna różnorodność dostępnych urządzeń i podzespołów składowych zestawu komputerowego zwłaszcza w obrębie danej kategorii (np. wiele modeli kart gra-ficznych różniących się niekiedy tylko nazwą),
- mnogość zastosowań podzespołów danej kategorii – trudność z jednoznacznym przyporządkowaniem elementu do danej grupy zastosowań,
- brak sformalizowanej wiedzy dotyczącej dużej gamy zastosowań różnych pod-zespołów.
- serwery (magazyny danych posiadające rozbudowane macierze dyskowe, wieloprocesorowe).

Wstępna analiza pokazała dużą złożoność problemu budowy odpowiedniego systemu informatycznego. Wydzielono początkowo podstawowe grupy podzespołów ze wzglę-du na ich funkcje w systemie komputerowym i podzielono na grupy modeli takie, jak:

- procesor (np. gr1: AMD Sempron i Intel Celeron),
- płyta główna (np. gr1: płyta główna dla procesorów Intel Celeron / Pentium4 (socket 775, chipset 945)),
- pamięć RAM (np. gr1: wielkość 512MB),
- karta graficzna (np. gr1: karta graficzna typu GeForce 7100 / Ati Radeon X300),
- karta dźwiękowa (np. gr1: zintegrowana na płycie głównej),
- dysk twardy (np. gr1: pojemność 80GB),
- zasilacz (np. gr1: moc 350W – 450W),
- napęd optyczny.

Następnie ustalono grupy zastosowań mających decydujące znaczenie dla wyboru odpowiedniego typu podzespołu. W przypadku wyboru grupy określającej rodzaj procesora wyróżniono:

- zastosowanie komputera do gier,
- zastosowanie komputera do multimediów,
- złożoność obliczeniowa,
- obróbka wideo,
- typ procesora: dwurdzeniowy.

Ostatecznym etapem budowy systemu SYSTEMINF było zbudowanie drzewa decy-zyjnego i zapis reguł w bazie wiedzy. Informacje na temat wymagań dostarcza użyt-kownik systemu podczas sesji z programem odpowiadając na kolejne pytania kwestio-nariusza takie, jak:

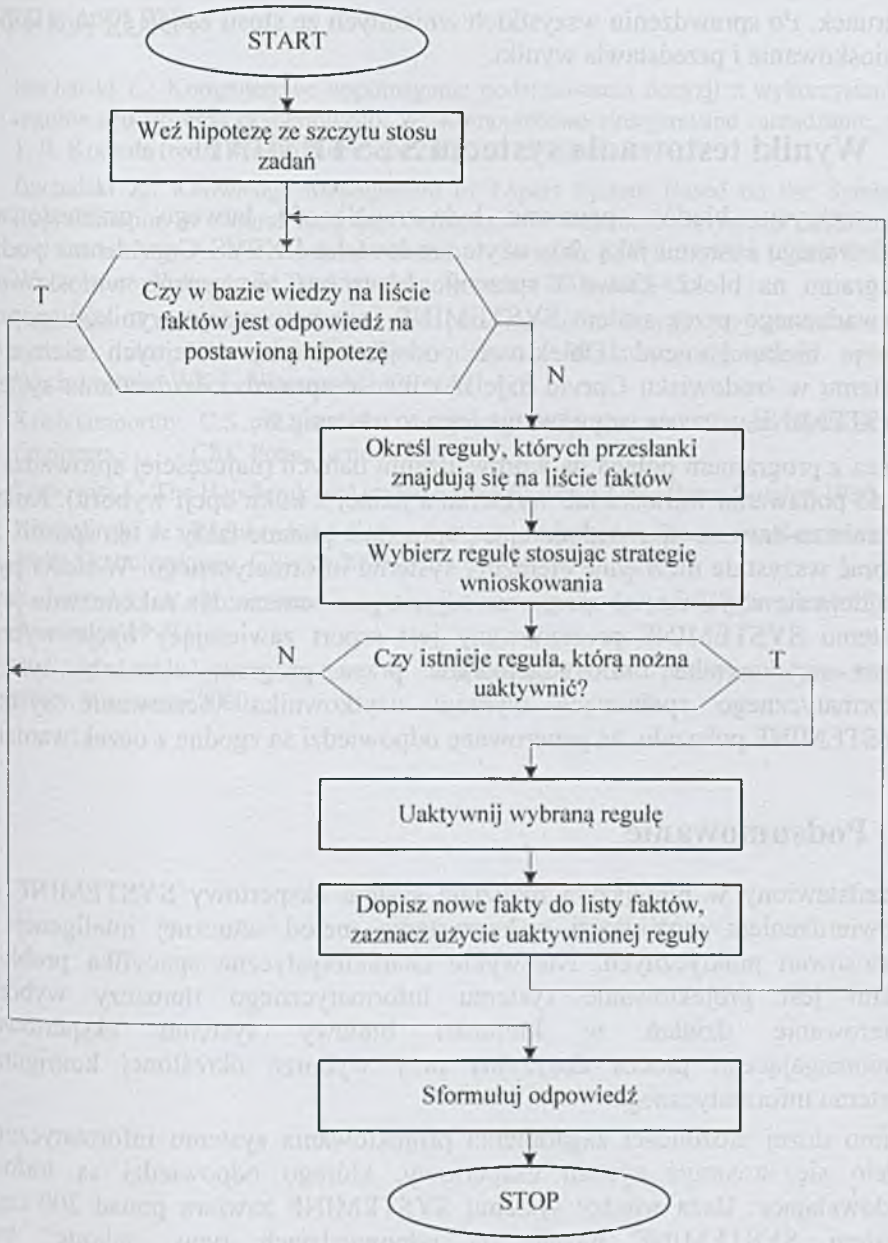
- czy użytkownik planuje obróbkę video (kodowanie / dekodowanie / nagrywanie / zgrywanie),
- czy używane będą złożone obliczeniowo aplikacje np. w pracy laboratoryjnej, przemysłowej, biura projektów,
- czy użytkownik planuje odtwarzać filmy, przeglądać strony WWW, słuchać muzyki,
- czy potrzebny jest procesor dwurdzeniowy.

6. Przebieg procesu wnioskowania

Środowisko EXSYS Corvid pozwala programiście na wygodne i szybkie zaimplemen-towanie algorytmów wnioskowania dla budowanego systemu ekspertowego. W omawianym systemie SYSTEMINF zastosowane zostało wnioskowanie mieszane.

Poprzez odpowiednie funkcje takie, jak DERIVE CONF oraz DERIVE [nazwa zmien-nej] umieszczone zostają wszystkie zmienne typu CONF oraz zmienne innych typów wybrane przez użytkownika na stosie zadań. Następnie w procesie wnioskowania w systemie SYSTEMINF algorytm wnioskowania w przód rozpoczyna się od umieszczenia hipotezy na stosie zadań. Następnie system przegląda listę faktów w bazie wiedzy, sprawdzając czy nie ma tam odpowiedzi na postawioną hipotezę. Jeżeli znajduje się tam już fakt, który daje się dopasować do hipotezy, to następuje zakończenie procesu wnioskowania i jest generowany odpowiedni komunikat.

W przypadku, gdy po przejrzeniu całej bazy faktów system nie może dać odpowiedzi na postawioną hipotezę, podejmowane są kroki w wyniku których generowane są no-we fakty. Uruchamiane są reguły, których przesłanki są prawdziwe. Wyznacza się zbiór reguł możliwych do zastosowania w danym etapie wnioskowania. Wybierana i uaktywniana jest jedna z reguł. Proces wnioskowania jest kontynuowany tak długo, aż zostanie osiągnięty cel lub gdy nie można uaktywnić więcej reguł. Algorytm wniosko-wania w przód przedstawiony został na poniższym schemacie:



Rys. 2. Schemat algorytmu wnioskowania w przód

W przypadku, gdy w procesie analizowania danej reguły w części warunkowej pojawi się nowa hipoteza system automatycznie podejmuje akcje wnioskowania wstecz próbując wykazać jej prawdziwość. Sprawdzane są kolejne reguły zawierające nową hipotezę bądź też system zadaje pytanie użytkownikowi o

warunek. Po sprawdzeniu wszystkich zmiennych ze stosu zadań system kończy wnioskowanie i przedstawia wyniki.

7. Wyniki testowania systemu SYSTEMINF

W wykryciu błędów pomocna była możliwość łatwego przetestowania budowanego systemu jaką daje użyte środowisko EXSYS Corvid oraz podział programu na bloki. Łatwo i sprawnie kontroluje się sposób wnioskowania prowadzonego przez system SYSTEMINF oraz prezentacje wyników poprzez edycje bloku komend. Obiektowe podejście zapisu kolejnych elementów systemu w środowisku Corvid daje możliwość sprawdzania działania systemu SYSTEMINF w czasie stopniowego jego rozrastania się.

Praca z programem polega na wprowadzaniu danych (najczęściej sprowadza się to do podawania wartości lub wybierania jednej z kilku opcji wyboru). Kolejne pytania za-dawane są uwzględniając poprzednio podane fakty w ten sposób, aby dobrać wszystkie niezbędne elementy systemu informatycznego. W treści pytań znajdują się wyjaśnienia, opisy oraz zdjęcia pomocnicze. Na zakończenie pracy systemu SYSTEMINF prezentowany jest raport zawierający opcje wybrane przez użytkownika oraz sugerowane przez program elementy systemu informatycznego spełniające wymogi użytkownika. Testowanie systemu SYSTEMINF pokazało, że generowane odpowiedzi są zgodne z oczekiwaniami.

8. Podsumowanie

Przedstawiony w niniejszym rozdziale system ekspertowy SYSTEMINF jest potwierdzeniem możliwości wykorzystania metod sztucznej inteligencji do zastosowań praktycznych. Niezwykle charakterystyczna specyfika problemu jakim jest projektowanie systemu informatycznego tłumaczy wybór i skierowanie działań w kierunku budowy systemu ekspertowego wspomagającego proces decyzyjny przy wyborze określonej konfiguracji systemu informatycznego.

Mimo dużej złożoności zagadnienia projektowania systemu informatycznego udało się stworzyć system ekspertowy, którego odpowiedzi są trafne i zadowalające. Baza wiedzy systemu SYSTEMINF zawiera ponad 200 reguł. System SYSTEMINF bazuje na odpowiedziach typu „tak/nie" bądź pochodzących z niewielkiego zbioru dyskretnego.

Przedstawiony w niniejszym rozdziale system SYSTEMINF na tym etapie budowy i złożoności potrafi dość sprawnie utworzyć szkielet konfiguracji i oszacować koszt budowy systemu informatycznego. Rozbudowanie systemu o kolejne reguły, a także połączenie z bazami danych podzespołów pozwoliłoby prawdopodobnie na budowę konfiguracji systemu informatycznego na poziomie zbliżonym do takich, jakie zaproponowałby ekspert.

LITERATURA

1. Buchalski Z.: Komputerowe wspomaganie podejmowania decyzji z wykorzystaniem regułowego systemu ekspertowego. W: Komputerowo zintegrowane zarządzanie, tom 1, R. Knosala (red.), WNT, Warszawa 2004, s.156-164.
2. Buchalski Z.: Knowledge Management of Expert System Based on the Symbolic Representation of Natural Language Sentences. W: Information Systems Architecture and Technology, L. Borzemski, A. Grzech, J. Świątek, Z. Wilimowska (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2006, s.75-85.
3. Buchalski Z.: Zarządzanie wiedzą w podejmowaniu decyzji przy wykorzystaniu systemu eks-pertowego. W: Bazy danych. Struktury, algorytmy, metody. Wydawnictwo WKiŁ, Warszawa 2006, s.471-478.
4. Krishnamoorthy C.S., Rajeev S.: Artificial Inteligence and Expert Systems for Engineers. CRC Press, London 1994.
5. Liebowitz J.: The Handbook of Applied Expert Systems. CRC Press, London 1996.
6. Niederliński A.: Regułowo-modelowe systemy ekspertowe. Pracownia Komputerowa Jacka Skalmierskiego, Gliwice 2006.
7. Radzikowski W.: Komputerowe systemy wspomagania decyzji. Wydawnictwo PWE, Warszawa 1990.
8. Zieliński J.: Inteligentne systemy w zarządzaniu. Teoria i praktyka. Wydawnictwo PWN, Warszawa 2000.

Rozdział 16

Wybrane aspekty wdrażania i ewolucji rozległych systemów automatycznego pomiaru energii elektrycznej AMR/AMM w warunkach krajowych

Paweł Piotrowski
Politechnika Warszawska
pawel.piotrowski@ien.pw.edu.pl

Streszczenie

W rozdziale zaprezentowano charakterystykę wykorzystywanych w praktyce architektur systemów AMR (Automatic Meter Reading) oraz AMM (Advanced Meter Management) oraz typowe funkcje dostępne w systemach. Przedstawiono dostępne na rynku krajowym systemy AMR/AMM oraz korzyści wynikające ze stosowania systemów. Przeanalizowano ponadto problemy związane z wdrażaniem systemów AMR/AMM w warunkach krajowych. Zakończenie stanowią wnioski związane z przyszłością systemów AMR/AMM.

1. Wprowadzenie

AMR (*Automatic Meter Reading*) to system umożliwiający zdalny odczyt liczników energii elektrycznej odbiorców zasilanych na niskim napięciu [2]. Celem tego typu systemów jest obniżenie kosztów odczytu oraz zwiększenie jego częstotliwości. Jest to więc system raczej pasywny, w niewielkim stopniu wpływający na zachowania odbiorców i nie zmieniający w zbyt dużym stopniu procesów rynkowych. System taki daje głównie korzyści firmie wykonującej odczyty energii elektrycznej.

AMM (*Advanced Meter Management*) to natomiast zautomatyzowany system opomiarowania będący rozwinięciem systemu AMR [3]. Główne funkcje tych systemów to zdalne zarządzanie obciążeniem, zdalne odłączanie odbiorców energii elektrycznej, realizacja rozwiązań przedpłatowych oraz zdalna rekonfiguracja.

W niektórych przypadkach spotkać można się także z terminami – systemy AM (*Advanced Metering*) lub AMI (*Advanced Meter Infrastructure*) Sugerują one systemy, które posiadają znacznie większe możliwości w zakresie dostępnych funkcji, rozbudowane możliwości oprogramowania a także integracji z innymi systemami np. systemem bilingowym.

Wykorzystywanie i ciągły rozwój systemów zdalnego odczytu energii elektrycznej trwa od lat 70-tych ubiegłego wieku [1]. Systemy tego typu mogą mieć zasięg lokalny np. fabryka lub zakład przemysłowy lub rozległy (obszar terytorialny zasilany ze spółki dystrybucyjnej lub wielu spółek). W tym drugim przypadku systemy służą przede wszystkim do odczytu zużycia energii elektrycznej u odbiorców komunalno-bytowych i przeprowadzane są przez spółkę dystrybucyjną. Na uwagę zwraca wysoki stopień skomplikowania budowy i obsługi takiego systemu (bardzo duża liczba rozproszonych liczników). Najbardziej rozwiniętym rynkiem systemów AMR są Stany Zjednoczone (ponad 20% wszystkich punktów pomiarowych jest związanych z systemami zdalnego odczytu energii elektrycznej). W Europie wprowadzenie systemów AMR zachodzi nieco wolniej. Jedną z dominujących na rynku europejskim jest szwajcarska firma Landis+Gyr, która dostarczyła ponad 1000 systemów AMR (najwięcej wdrożeń miało miejsce w państwach skandynawskich w których systemy AMR obejmują są ponad 2 mln. liczników). W Turcji liczba liczników również wynosi około 2 mln. Natomiast największe w Europie wdrożenie przeprowadzane jest obecnie we Włoszech w firmie energetycznej ENEL, posiadającej 80% włoskiego rynku energii elektrycznej. System wdrażany jest z pomocą firmy Echelon od roku 2001 i obecnie jest w fazie końcowej. Od strony informatycznej wykonawcą jest firma IBM (oprogramowanie komunikacyjne systemu, mechanizmy wymiany i przepływu danych, oprogramowanie IBM WebSphere). Dodać należy, że Firma IBM w chwili obecnej wdraża również system w kilku innych europejskich i amerykańskich spółkach dystrybucyjnych [2]. Do przesyłu informacji od licznika do lokalnego koncentratora wykorzystano linie niskiego napięcia. Liczba już zainstalowanych liczników w systemie wynosi ponad 30 mln. Współpracują one z ponad 350 tysiącami lokalnymi koncentratorami. Na jeden koncentrator przypada więc niecałe 100 liczników. Z uwagi na bardzo dużą liczbę przetwarzanych danych powstał zespół centrów zarządzania – na jedno centrum zarządzania przypada około 3 mln. liczników. W sumie istnieje w systemie 15 centrów zarządzania – 4 krajowe oraz 11 lokalnych. System obecnie ma zdolność odczytu około 700 tysięcy liczników dziennie.

Szacuje się natomiast, że na całym świecie liczba liczników w systemach AMR wzrasta o około 15-20%.

W Polsce, proces ten ma do tej pory charakter raczej pilotażowy (z uwagi na koszty wprowadzania tego typu systemów). Szacuje się najczęściej, że pojedynczy punkt pomiarowy w systemie to koszt około 1000 zł. Przykładowo firma APATOR S.A wykonała pilotażowe wdrożenie systemu AMR dla spółki

dystrybucyjnej ENERGA Gdańsk. System jest zainstalowany w różnych rozproszonych terytorialnie punktach (osiedla mieszkaniowe, domki jednorodzinne). Jest to jednak rozwiązanie jedynie lokalnego odczytu na niewielkim obszarze. Inkasent dokonuje odczytów w tzw. trybie „obchodzeniowym" - gromadzi odczyty poprzez komunikację specjalnego mobilnego czytnika z grupą niezbyt odległych liczników (100-200m) wykorzystując komunikację radiową (obecnie jest to w Polsce najbardziej popularna wersja systemu AMR). Natomiast aby w pełni wykorzystać zalety systemu należałoby zainstalować zespół koncentratorów z których dane przesyłane byłyby przy wykorzystaniu np. technologii GSM do centralnego punktu gromadzącego odczyty - bazy odczytowej mającej komunikację z systemem bilingowym oraz innymi systemami.

2. Architektura systemów AMR/AMM

Pomiar zużycia energii elektrycznej odbywa się poprzez specjalistyczny licznik energii elektrycznej wyposażony w interfejs komunikacyjny do odbioru i wysyłania sygnałów lub standardowy licznik wyposażony w odpowiednią przystawkę komunikacyjną i połączoną z licznikiem przez wyjście impulsowe [4]. Do komunikacji wykorzystywany jest najczęściej zdefiniowany w roku 1996 przez producentów liczników i dostawców energii, zbiór protokołów warstwy aplikacyjnej DLMS (Device Language Message Specification), co tłumaczyć można jako: specyfikacja komunikatów w języku urządzenia. W roku 2002, DLMS oznaczony jako IEC 62056 został zaadoptowany na potrzeby urządzeń do odczytu zużycia gazu, wody oraz energii elektrycznej. DLMS stanowi logiczne połączeniem pomiędzy licznikiem a systemem gromadzącym dane. Odpowiada ponadto za autoryzację oraz prawa dostępu. Prezentacja danych warstwy aplikacyjnej opiera się na standardzie ASN.1 (*Abstract Syntax Notation One*). Jest to standard służący do opisu struktur przeznaczonych do reprezentacji, kodowania, transmisji i dekodowania danych i dostarcza zbiór formalnych zasad pozwalających na opis struktur obiektów w sposób niezależny od konkretnych rozwiązań sprzętowych.

DLMS wykorzystywany może być w różnych mediach (m.in.: standardy PLC, PSTN, GSM, HHU, GPRS, Ethernet). Standard skutecznie działa w kanałach o ograniczonej przepustowości, podejście obiektowe umożliwia dodanie do standardu DLMS nowych parametrów oraz umożliwia utworzenie nowych modeli taryf przeznaczonych dla określonych grup klientów. Możliwa jest zatem elastyczna zmiana modelu danych oraz profili komunikacyjnych w zależności od nowych wymagań. Bezpieczeństwo przesyłanych danych zapewniają typowe standardy kryptograficzne (AES, TDES). Szczegółowy opis standardu DLMS dostępny jest na stronie http://www.dlms.com/.

Językiem wymiany danych może być oprócz ASN.1 także SML (*Sensor Markup Language*) czyli XML o profilu sprzętowym (opracowany w roku 1997). Rys.1 przedstawia przykładowe porównanie notacji w języku XML oraz ASN.1.

ASN.1

```
GET-Request-Normal ::= SEQUENCE
{
    invoke_id_and priority ::= bit string
    cosem_attribure_descriptor
    {
        class_id ::= unsigned16
        instance_id ::= octet-string
        attribute_id ::= Integer8
    }
}
```

XML

```
- <GetRequestNormal>
    <InvokeIdAndPriority Value="81" />
  - <AttributeDescriptor>
      <ClassId Value="0008"/>
      <InstanceId Value="0000010000FF" />
      <AttributeId Value="02"/>
  </AttributeDescriptor>
</GetRequestNormal>
```

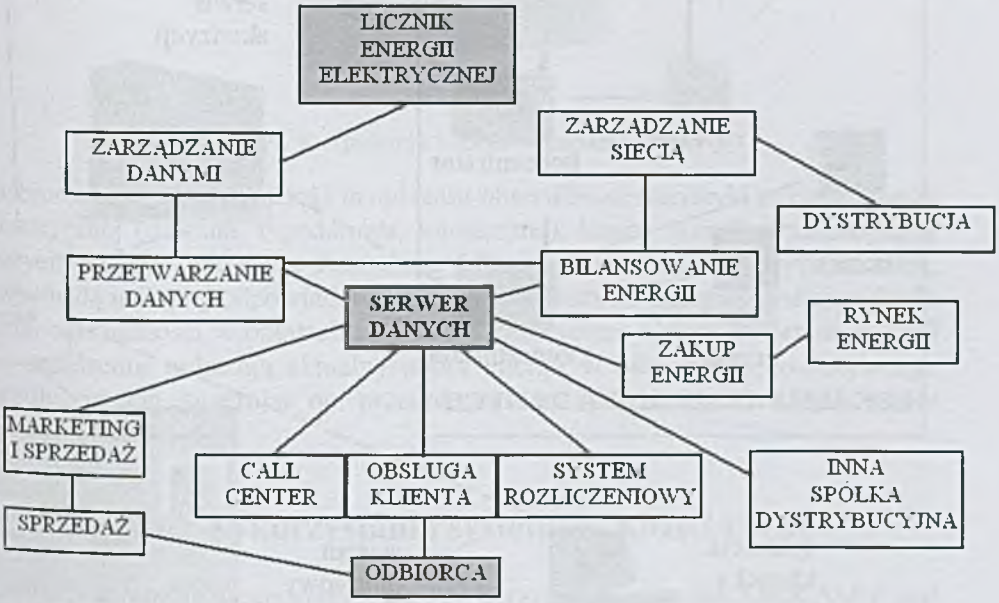
Rys. 1. Przykład zapisu informacji w notacji ASN.1 oraz XML

Licznik może być dostosowany do integracji z systemami AMR dzięki wykorzystaniu wymiennych modułów komunikacyjnych (np. moduł PLC z obsługą DLMS lub moduł GSM/GPRS z obsługą DLMS).

Dane o zużyciu energii przesyłane są od liczników do specjalnych koncentratorów (przekaźników - komputerów przemysłowych) wykorzystując transmisję przewodową – transmisja liniami niskiego napięcia, średniego napięcia - technologia PLC o częstotliwości sygnałów od 3 do 95 kHz w technologii Lonworks lub bezprzewodową (modem GSM z transmisją pakietową GPRS, technologia GSM z wykorzystaniem do przesyłania danych wiadomości tekstowych SMS, modem radiowy do lokalnej transmisji radiowej o zasięgu od 50 metrów w budynku do kilku kilometrów w terenie otwartym (częstotliwość sygnału wynosi najczęściej kilkaset MHz), wykorzystanie infrastruktury lokalnej sieci bezprzewodowej WLAN (konieczne specjalne moduły konwersji)).

Zadaniem koncentratora synchronizowanego zegarem satelitarnym jest gromadzenie danych z lokalnie rozproszonej (od kilkudziesięciu do ponad stu) grupy liczników oraz synchronizowanie ich zegarów. W takiej konfiguracji licznik stanowi urządzenie typu slave (podrzędne) natomiast koncentrator urządzenie typu master (nadrzędne). Standardowe aktywne urządzenia sieciowe typu hub lub router nie sprawdziły się niestety w roli koncentratorów w systemach AMR [4]. Koncentratory występują w wersjach stacjonarnych oraz mobilnych (zestaw na wyposażeniu inkasenta, który odczytuje liczniki). W sytuacjach gdy powstaje nadmierne tłumienia lub istnieją duże zakłócenia pomiędzy licznikami a koncentratorem, stosuje się dodatkowo repeatery, służące do regeneracji (wzmocnienia) transmitowanego sygnału. Zadaniem serwerów

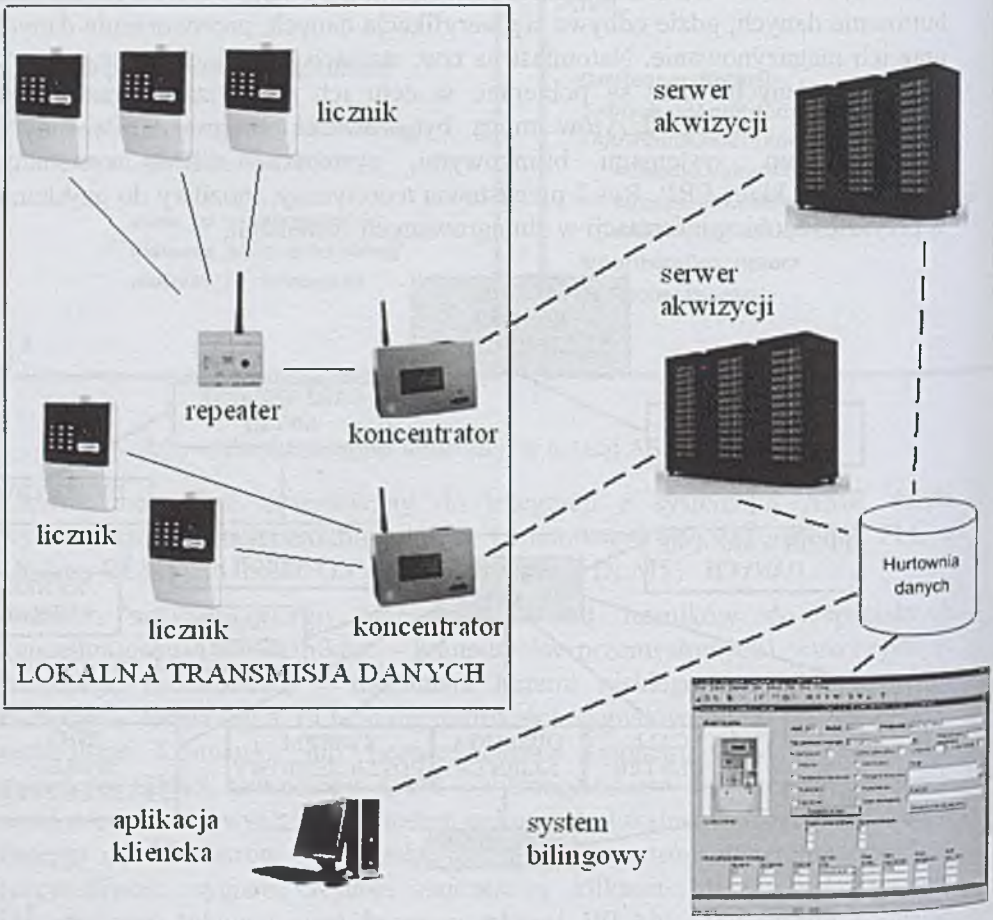
akwizycji z bazami telemetrycznymi (komputery przemysłowe typu serwer o dużej mocy i niezawodności) jest pobieranie danych z odległych koncentratorów w ustalonych odstępach czasu wykorzystując technologie: przewodową (sieć Ethernet, łącze komutowane PSTN) lub bezprzewodową (HSCSD, GPRS, transmisja radiowa WiMAX). Bardzo rozbudowane systemy AMR wykorzystują hurtownie danych, gdzie odbywa się weryfikacja danych, przetworzenie danych oraz ich magazynowanie. Natomiast na tzw. stacjach klienckich połączonych z hurtownią danych, dane są pobierane w celu ich wizualizacji oraz analiz. Przetworzone dane z odczytów mogą być również integrowane z innymi systemami np. systemami bilingowymi, systemami CRM, systemami biznesowymi klasy ERP. Rys.2 przedstawia teoretyczny, możliwy do uzyskania w przyszłości obieg informacji w zintegrowanych systemach.



Rys. 2. Struktura przetwarzania danych w zintegrowanych systemach (opracowano na podstawie[6])

Lokalna transmisja danych (licznik-koncentrator) wykorzystuje najczęściej technologie bezprzewodowe (67%) oraz technologię PLC (25%). Sporadycznie wykorzystuje się także linie telefoniczne (6%). Inne metody transmisji lokalnej mają udział 1,6% [7]. Stany Zjednoczone, które jako pierwsze już w latach 70-tych ubiegłego wieku wdrażały systemy AMR, mają bardzo elastyczne podejście do sposobu transmisji - jest ono uzależnione od warunków na danym obszarze (ukształtowanie terenu, sytuacja techniczna, stan transformatorów, stan sieci niskiego napięcia, infrastruktura, lokalizacja budynków) [8]. Wykorzystywana w praktyce jest zarówno transmisja wykorzystująca technologię PLC jak również technologie mobilne czyli komunikacja radiowa lub transmisja GSM. Natomiast

we Włoszech wdrażany w firmie ENEL system AMR/AMM wykorzystuje w połączeniach lokalnych wyłącznie technologię PLC co znacznie utrudnia wdrażanie całego systemu. Rysunek 3 przedstawia typowe elementy rozbudowanego systemu AMR.



Rys. 3. Struktura systemu AMR

Patrząc na system AMR od strony klienta (odbiorcy energii elektrycznej), bardzo ciekawym rozwiązaniem jest możliwość analizy zużycia energii elektrycznej wykorzystując do tego celu dane z systemu przy wykorzystaniu odpowiedniej aplikacji na komputerze odbiorcy (połączenie przez internet) lub specjalnego urządzenia zainstalowanego w domu (np. australijski ecoMeter (rys.4)– terminal do dwukierunkowej komunikacji).

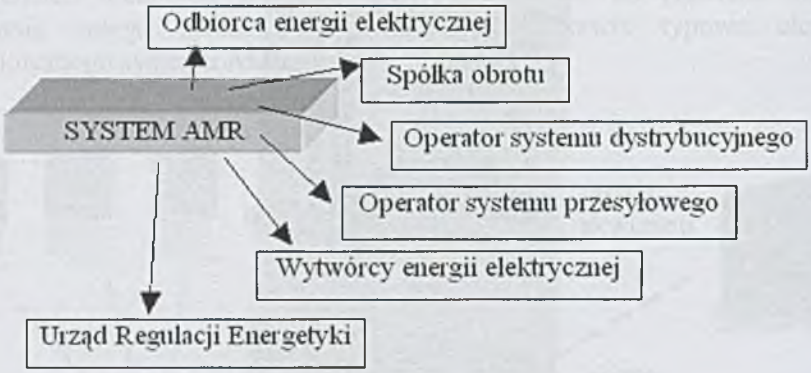
Rys. 4. Aplikacja klienta – ecoMeter [5]

Odbiorca energii może dzięki urządzeniu obserwować statystyki zużycia energii elektrycznej (dzienna, tygodniowa, miesięczna), koszty energii oraz prognozy zużycia. Dzięki temu może elastycznie dopasowywać dla siebie taryfy, na takie, które będą najlepiej odpowiadały profilom jego zużycia energii uzyskując dzięki temu oszczędności w opłatach za energię elektryczną. Cztery kolory diod LED na urządzeniu wskazują aktualny pobór energii w odniesieniu do aktywnego profilu zużycia, sugerując np. przekroczenie szczytu obciążenia taryfy, pobór energii w normie itd.

3. Korzyści z wykorzystania systemów AMR/AMM

Liczba podmiotów mogących oczekiwać korzyści z systemów AMR/AMM jest dość duża, od najmniejszego podmiotu czyli odbiorcy energii elektrycznej aż do Urzędu Regulacji Energetyki (rys.5).

Odbiorca energii elektrycznej dzięki systemowi ma m.in. [1, 12]: dostęp do informacji o historii zużycia energii elektrycznej, możliwość zmiany taryfę na taką, która jest lepiej dostosowana do odbiorcy i zapewni mu mniejsze opłaty za zużycie energii, krótsze przerwy w zasilaniu dzięki automatycznej sygnalizacji zaniku napięcia u odbiorcy w systemie komputerowym dostawcy energii, ułatwioną zmianę dostawcy energii, możliwość płacenia za energię elektryczną w formie przedpłatowej, miesięczne płatności za energię oparte na rzeczywistym zużyciu energii a nie na prognozach opartych na zużyciu energii przez odbiorcę w okresach wcześniejszych, w przyszłości możliwość sterowania zdalnym włączaniem oraz wyłączaniem domowych odbiorników energii elektrycznej w celu zmniejszenia zużycia energii elektrycznej.



Rys. 5. Podmioty mogące odnieść korzyści z systemów AMR

Spółka obrotu a także Operator systemu dystrybucyjnego może odnosić korzyści w postaci m.in. [1, 12]: obniżenia kosztów odczytu, skrócenia cyklu rozliczeniowego klienta, automatyzacji rozliczeń, finansowania zgodnie z rzeczywistą sprzedażą, lepszych prognoz (dzięki aktualnym i częściej odczytywanym danych o zużyciu energii przez odbiorcę), możliwości segmentacji odbiorców i ustalania nowych taryf i cenników bardziej dla nich dopasowanych, ułatwienia tworzenia umów bazujących na ograniczeniu maksymalnego poboru energii, dużej elastyczności regulacji dostawy energii (zdalne odłączanie odbiorców, natychmiastowe lub czasowe ograniczenie ich mocy), zdalne sterowanie zabezpieczeniami, znacznego ułatwienia identyfikacji uszkodzeń w sieci oraz skrócenia przerw w zasilaniu, bilansowania wybranych obszarów sieci wykorzystujące jednoczesny pomiar z różnych liczników, znacznego ułatwienia wykrywania nielegalnego poboru energii dzięki skutecznemu bilansowaniu sieci energetycznej czyli lepszej kontroli strat handlowych, nowoczesnej i bezpiecznej archiwizacji danych o odczytach, wspomagania decyzji inżynierskich mających na celu np. modernizację lub budowę nowej stacji elektroenergetycznej oraz możliwości pomiaru parametrów jakościowych energii elektrycznej oraz wizualizacji danych.

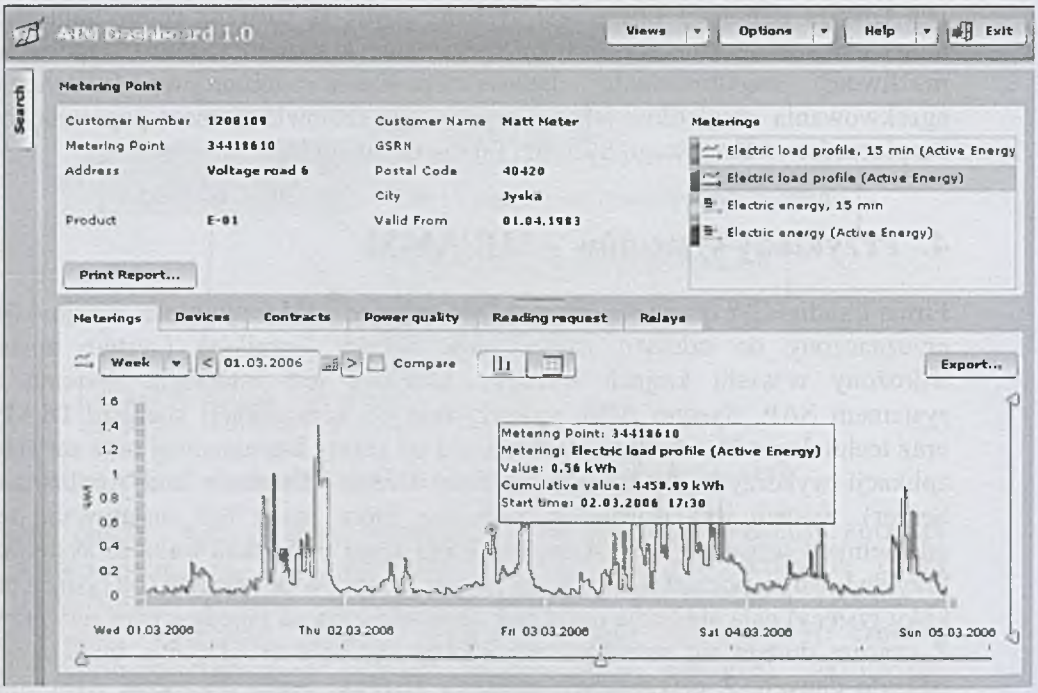
Potencjalne zalety dla Operatora systemu przesyłowego to możliwość wykorzystania małych odbiorców jako elementu regulacji Krajowego Systemu Energetycznego, uzyskując dzięki temu znaczne spłaszczenie krzywej zużycia energii, chroniąc równocześnie sieć energetyczną przed awariami wynikającymi z przeciążeń sieci [12]. Ważna jest także możliwość poprawy jakości prognozowania ruchu sieci mogąca ograniczyć koszty wynikające z niezbilansowania sieci energetycznej.

Wytwórcy energii elektrycznej mogą odnieść korzyści z systemów AMR w zakresie ograniczenia krótkotrwałych uruchomień lub odstawień bloków (stabilizacja systemu energetycznego) oraz polepszania jakości prognozowania produkcji energii elektrycznej [12].

Dla Urzędu Regulacji Energetyki systemy AMR są również cenne ponieważ mogą ułatwiać zmianę sprzedawcy (wspieranie konkurencyjności), zapewniać możliwość monitorowania jakości zasilania odbiorców (ułatwienie egzekwowania standardów jakościowych) oraz stanowić element poprawiający bezpieczeństwo Krajowego Systemu Energetycznego [12].

4. Przykłady systemów AMR/AMM

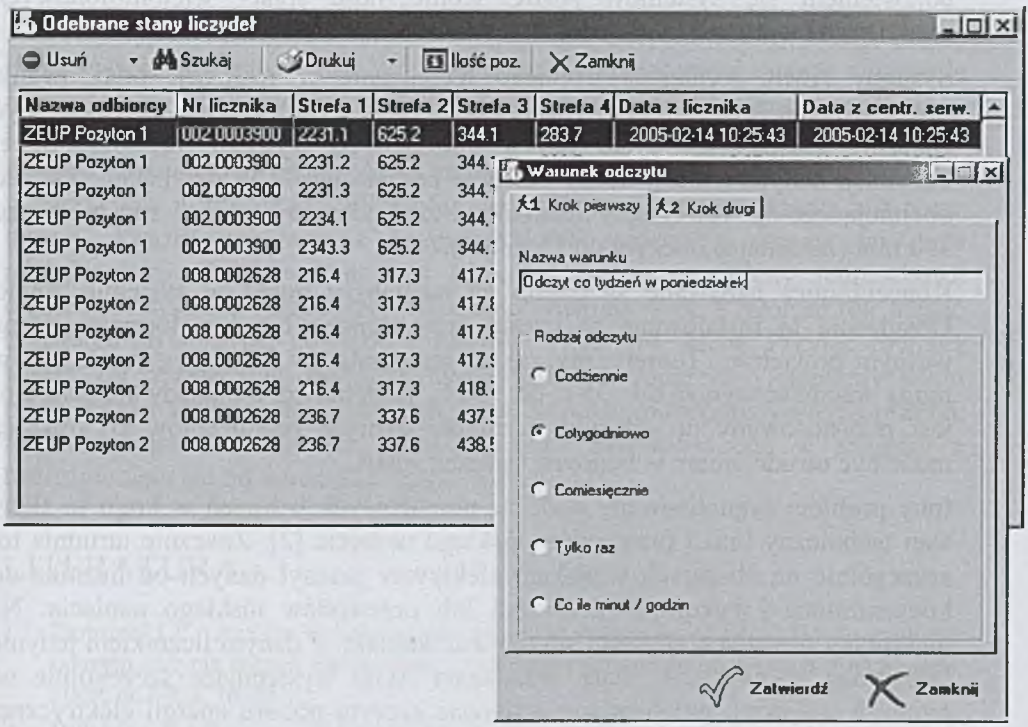
Firma Landis+Gyr oferuje system AMM o nazwie AIM (najnowsza wersja 4.0), przeznaczony do odczytu nawet setek tysięcy liczników (system został wdrożony w wielu krajach Europy). Możliwa jest integracja systemu z systemem SAP. System AIM wykorzystuje do komunikacji standard DLMS oraz technologię PLAN PLC. Architektura od strony bazodanowej oraz serwera aplikacji wykorzystuje rozwiązanie firmy Oracle (Database and Application Server). System jest złożony z modułów, które mogę być instalowane na oddzielnych serwerach. Uzyskuje się dzięki temu dużą skalowalność systemu. Przykładowo na początku przy małej liczbie liczników do odczytu (od tysiąca do kilku tysięcy) cała aplikacja może być zainstalowana na pojedynczym serwerze. Z czasem dodaje się serwery komunikacji w celu zwiększenia wydajności odczytu danych. Z ciekawszych aplikacji systemu warto wymienić AIM Site Manager oraz AIM Dashboard Web Application (rys.6). Pierwsza z nich wykorzystywana może być na urządzeniach mobilnych (np. PDA) i służy do automatycznej konfiguracji liczników oraz do zarządzania informacją (połączenie z głównym systemem odbywa się w trybie on-line). Druga aplikacja sieciowa umożliwia działowi obsługi klienta i menedżerom przedsiębiorstwa energetycznego uzyskanie szybkiego wglądu w system AIM klienta. Aplikacja przekazuje również dane pomiarowe i sprzętowe, dotyczące liczników, podłączonych do systemu AIM. Aplikacja usprawnia wiele istotnych czynności wykonywanych przez przedsiębiorstwo energetyczne, zapewniając m.in.: lepszą wydajność obsługi klientów, zmieniających miejsce zamieszkania oraz dostawcę, szybszą reakcję na zapytania dotyczące fakturowania, łatwość śledzenia przerw w dostawie energii oraz informacji, dotyczących jakości dostarczanej energii, zwiększa efektywność procesu ściągania długów oraz zapewnia większą różnorodność usług, oferowanych przez operatora sieci energetycznej [6].



Rys. 6. Widok profili obciążeń w aplikacji klienta AIM Dashboard 1.0 systemu AIM firmy Landis+Gyr [6]

System ANT AMR (odczyt zużycia energii elektrycznej, gazu, ciepła, wody) ma w ofercie Firma ANT Sp. z o.o. [9]. System współpracujące z oprogramowaniem dedykowanym ANT STUDIO 3.1. Dane przesyłane są technologią GPRS do wewnętrznej bazy danych SQL Lite, z możliwością ich eksportu do innych baz danych np. Oracle, MySQL. Do systemu podłączony może być każdy licznik z wyjściem impulsowym. Ciekawostką w ofercie firmy jest specjalna nakładka ze skanerem OCR na licznik tradycyjny starego typu, umożliwiająca zeskanowanie stanu licznika, rozpoznanie widocznych liczb i ich konwersję do zapisu cyfrowego w postaci pliku z danym.

System zdalnego odczytu SHOOK G (System Handlowej Obsługi Odbiorców Komunalnych GSM) do obsługi odbiorców komunalnych ma w ofercie firma Pozyton Sp. z o.o. [14]. Elementy systemu stanową 1 i 3 fazowe liczniki wyposażone w moduł komunikacji bezprzewodowej GSM, centralny system wykonywania odczytów wyposażony w oprogramowanie SHOOK G (rys.7) oraz modemy GSM. Możliwy jest zdalny odczyt danych pomiarowych przez klienta przy pomocy telefonu komórkowego. Jedno stanowisko komputerowe z oprogramowaniem SHOOK G może obsługiwać do kilkunastu tysięcy odbiorców. Dane standardowo odczytywane są raz w miesiącu. Oprogramowanie umożliwia analizę odbiorców, wykonywanie bilansów zużycia energii dla grup odbiorców za zadany okres czasu. Możliwy jest także eksport danych do specjalistycznych systemów bilingowych.



Rys. 7. Zdalny odczyt danych o zużyciu energii – oprogramowanie SHOOK G [14]

Firma APATOR S.A. również posiada w ofercie system AMR o nazwie AMRsystem [11]. Akwizycja danych i ich przetwarzanie odbywa się w systemie zarządzania informacją (SIPO- Sieciowa Informacja Pomiarowo-Odczytowa). System został wdrożony w wielu rejonach Polski. Firma ma również w ofercie coraz popularniejszy system o nazwie LEWsystem Apator (przedpłatowy system pomiaru i rozliczania mediów energetycznych), który wykorzystuje w Polsce już ponad 200tys. osób.

5. Problem wdrażania systemów AMR w Polsce

W kraju trwa obecnie proces związany z konsolidacją spółek dystrybucyjnych [2]. Ma miejsce również proces rozdziału pomiędzy operatorami systemów dystrybucyjnych a spółkami obrotu. Dodatkowo od 1 lipca 2007 roku została wprowadzona zasada TPA (Third Part Access) dająca możliwość zmiany dostawcy energii elektrycznej. Zachodzą wiec w kraju procesy, które nie sprzyjają dużym inwestycjom w systemy AMR. Brak jest również jednolitych standardów związanych z podejściem do zagadnienia systemów AMR. Trudno jest przewidywać kto miałby być inwestorem w systemy AMR. Dużym

problemem mogą być również potencjalne redukcje pracowników związane z pojawieniem się systemów AMR. Konieczność spłaty wielomilionowych kredytów na wdrożenie systemu AMR jest również istotnym problemem.

Systemy AMR wymagają trudnego technicznie w realizacji bilansowania rozproszonego olbrzymiego zbioru odbiorców. Odczyt danych musi być znacznie częstszy niż tylko raz w miesiącu co wymaga zwielokrotnionej transmisji danych z dużej liczby punktów pomiarowych. W przypadku systemu obejmującego np. 100 tysięcy liczników, liczba odczytów wynosi rocznie ponad 350 mln (zakładając odczyty do 15 minut).

Koncentratory uznawane są często za najsłabszy punkt w systemie AMR. Urządzenia te instalowane są często przy słupach lub transformatorach na wolnym powietrzu. Teoretyczny zakres temperatur w której wg producentów mogą pracować wynosi od –35 C do +50 C. Dodatkowo, nie każdy koncentrator jest przystosowany do odczytu liczników różnych producentów co również może być utrudnieniem w budowie systemu AMR.

Inny problem sygnalizowany podczas pilotażowych wdrożeń w kraju to słaby stan techniczny linii i przewodów niskiego napięcia [2]. Znacznie utrudnia to, szczególnie na obszarach wiejskich efektywny przesył danych od licznika do koncentratora z wykorzystaniem linii lub przewodów niskiego napięcia. Na niektórych obszarach udawało się uzyskać kontakt z danym licznikiem jedynie kilka razy w miesiącu. Stare urządzenia AGD występujące szczególnie na terenach wiejskich uruchamiane w okresie szczytu poboru energii elektrycznej również w praktyce znacznie utrudniały komunikację z licznikami i był to realnie występujący problem, trudny do rozwiązania. W niektórych miejscach w kraju linie napowietrzne ulegają częstym uszkodzeniom (zerwanie spowodowane np. wiatrem) a same końcówki przewodów połączone metalowymi złączkami znacznie utrudniają przesył danych. Przy nieprawidłowych połączeniach aluminium ulega korozji utrudniając komunikację z licznikami. W niektórych przypadkach nawet zwiększanie siły sygnały (z 60 do 70 dB) nie pozwalało na uzyskanie skutecznego połączenia z licznikiem. Często na terenach wiejskich występowało również zjawisko sprzężeń pojemnościowych (sygnał wysłany do licznika jedną fazą powodował wygenerowanie odpowiedzi licznika po innej fazie). Powstaje więc również finansowy problem modernizacji fragmentów sieci energetycznej aby możliwe było skuteczne przesyłanie danych w systemie AMR

Dodatkowo, w przypadku podjęcia decyzji o redundancji systemu AMR (większa dostępność i bezpieczeństwo systemu) wielokrotnie zwiększą się nakłady finansowe.

6. Podsumowanie

Krajowy rynek rozległych systemów zdalnego odczytu energii AMR rośnie o kilkanaście procent rocznie. Wdrażanie tego typu systemów to duże wyzwanie informatyczne ale przede wszystkim trudna po pokonania bariera ekonomiczna. Duże korzyści ze stosowania systemów zdalnego odczytu sprawiają jednak, że w czasie kilkunastu lat systemy AMR staną się zapewne powszechnym standardem w energetyce w Polsce [1]. Firm oferujących kompletne rozwiązania jest coraz więcej. Koszty związane z bezprzewodową transmisją danych na duże odległości również z rok na rok są coraz mniejsze. Ceny jednostkowe elementów systemu oraz koszty wdrażania powinny więc z roku na rok ulegać obniżeniu co sprzyjać powinno rozwojowi systemów AMR. Przyjąć również należy, że bezpieczniej budować system AMR w oparciu o elementy wielu producentów ale kompatybilne, co pozwoli uniknąć monopolistycznych praktyk ze strony dostawcy sprzętu (nieuzasadnione zawyżanie cen w związku z uzależnieniem się od konkretnej marki sprzętu) [2].

LITERATURA

1. Piotrowski P.: Analiza techniczno-ekonomiczna systemów AMR/AMM do zdalnego odczytu zużycia energii elektrycznej u odbiorców komunalno-bytowych. Elektro-info, 2009, nr.7-8, s. 70-76.
2. Billewicz K.: Systemy AMR – polskie problemy. Przegląd Elektrotechniczny, 2007, nr.12.
3. Piotrowski J.: Nowoczesna technologia wspiera wolny rynek energii. Wokół energetyki, kwiecień 2007
4. Billewicz K.: System automatycznego odczytu liczników energii elektrycznej AMR. Pomiary Automatyka Robotyka, 2007, nr.7-8.
5. http://www.ecometer.com.au
6. http://poland.landisgyr.com
7. http://www.freescale.com/files/dsp/doc/reports_presentations/ECS04P31.pdf
8. http://www.gigawat.net.pl/article/articleview/912/1/72
9. http://www.ant-iss.pl
10. http://www.pozyton.com.pl/
11. http://www.apator.eu
12. Piotrowski J.: Nowoczesna technologia wspiera wolny rynek energii. Wokół energetyki, 2007, nr.7

Rozdział 17

Ocena efektywności zastosowania systemu GIS w spółce dystrybucji energii elektrycznej

Piotr Helt
Politechnika Warszawska
piotr.helt@ien.pw.edu.pl
Sławomir Noske
ENERGA Operator S.A.
slawomir.noske@elblag.energa.pl

Streszczenie

W opracowaniu przedstawiono zagadnienia związane z wykorzystaniem Systemu Informacji Geograficznej (GIS) w przykładowej spółce dystrybucyjnej. Podkreślono celowość dokonania szerokiej analizy struktury i organizacji zakładu energetycznego w procesie wdrażania systemu. Przedstawiono strukturę logiczną systemu GIS dla omawianej spółki dystrybucyjnej. Opisano zrealizowany zestaw funkcji obliczeniowo-analitycznych, a także określono możliwe do realizacji kierunki rozwoju systemu. Dokonano oceny wpływu wprowadzenia systemu GIS na działalność analizowanej spółki, oszacowano także efekty wynikające ze stosowania takiego systemu. W rozdziale uwzględniono także powiązania systemu GIS z innymi systemami informatycznymi, eksploatowanymi w analizowanej spółce.

1. Wprowadzenie

W procesie przemian następujących w polskiej energetyce stworzone zostały silne grupy energetyczne, wydzielone zostały przedsiębiorstwa dystrybucyjne, postępuje proces liberalizacji rynku energii elektrycznej, następuje prywatyzacja energetyki.

Tak rewolucyjne zmiany i perspektywa nowej przyszłości wymagają od przedsiębiorstw dystrybucyjnych stworzenia strategii w zakresie zarządzania

majątkiem sieciowym. Zarządzanie majątkiem sieciowym wymaga pozyskania jak najpełniejszych informacji. Lecz nie sama wiedza gromadzona w przedsiębiorstwie, ale to w jaki sposób zostanie wykorzystana do osiągania strategicznych celów, będzie istotnym czynnikiem w osiąganiu konkurencyjności firmy. Znaczącym fragmentem tej wiedzy jest informacja techniczna o sieciach energetycznych, jeszcze do niedawna gromadzona w paszportach linii, utrzymywana w formie wielu planów i schematów papierowych. Coraz częściej informacja ta gromadzona jest we wdrażanych systemach GIS (Geographic Information System - System Informacji Geograficznej) - systemach informacyjnych służących do wprowadzania, gromadzenia, przetwarzania oraz wizualizacji danych geograficznych.

W przedsiębiorstwach dystrybucyjnych w różny sposób definiowana jest rola systemów GIS. W skrajnych przypadkach może być realizowany jako:

- Graficzny obraz sieci z podstawowymi danymi technicznymi – system paszportyzacji sieci elektroenergetycznej. Pozostałe dane techniczne, eksploatacyjne i inne wykorzystywane w zarządzaniu majątkiem sieciowym przechowywane są w innych systemach komputerowych.
- System, w którym oprócz danych związanych z paszportyzacją sieci (dane techniczne, topologia, topografia) gromadzone są także dane eksploatacyjne (pomiary, wyniki oględzin, ocena stanu technicznego elementów sieci) i jest on stałym narzędziem, na którym oparta jest eksploatacja sieci.

Biorąc pod uwagę czynnik biznesowy, można wyodrębnić dwa podstawowe obszary rozwiązań informatycznych w przedsiębiorstwach sieciowych:

1. Obsługa klienta.
2. Utrzymanie infrastruktury sieciowej (zaliczając tu również eksploatację i rozwój sieci), gwarantującej zapewnienie wymaganego poziomu usług.

Ze względu na występujący w obu obszarach rozwiązań aspekt przestrzenny, właściwym narzędziem do wspomagania procesów biznesowych są systemy GIS, przy czym ich funkcje powinny znacznie wykraczać poza stosunkowo prosty system paszportyzacji.

2. Struktura systemu GIS w spółce dystrybucyjnej

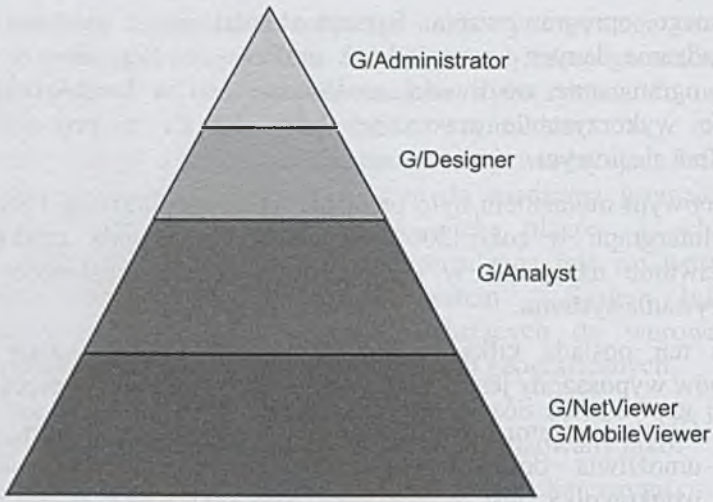
Obecna spółka dystrybucyjna ENERGA-Operator SA Oddział w Elblągu, jako jedno z pierwszych przedsiębiorstw energetycznych w Polsce zdecydowała się na budowę systemu informacji przestrzennej GIS. W 1997 roku zostały rozpoczęte prace projektowe związane z budową systemu informatycznego a w kolejnym roku powstała aplikacja GIS. Przy tworzeniu bazy danych głównie wykorzystano własne zasoby osobowe i techniczne, niezbędnym elementem tych działań było zawarcie umów na udostępnienie map wektorowych z państwowych zasobów geodezyjnych. Barierą rozwoju wspomagania działalności dystrybucyjnej była ograniczona funkcjonalność analityczna

ówczesnego oprogramowania. System składał się z modułu służącego do wprowadzania danych i przeglądarek graficznych. Narzędzia te miały jednak bardzo ograniczone możliwości analityczne co w konsekwencji utrudniało szerokie wykorzystanie stworzonej bazy danych w procesie zarządzania majątkiem sieciowym.

Przełomowym momentem było przejście na nową platformę GIS - G/Electric firmy Intergraph w roku 2004. Wprowadzone zostały moduły G/Electric przedstawione na rys.1., w efekcie pojawiły się możliwości pełniejszego wykorzystania systemu.

System ten posiada kilka poziomów funkcjonalności, każdy z wyższych poziomów wyposażony jest w funkcjonalność rozwiązania poprzedzającego:

- G/Administrator pozwala na pełną administrację systemu. To narzędzie umożliwia dostosowywanie systemu do indywidualnych potrzeb użytkownika. Jest to możliwe dzięki budowaniu aplikacji poprzez tak zwane metadane, bez ingerencji w aplikacje G/Electric. Niesie to wiele korzyści dla użytkowników miedzy innymi: łatwe a tym samym tanie dokonywanie modyfikacji systemu z dostosowaniem do indywidualnych potrzeb użytkowników, korzystanie z rozwijanego światowego rozwiązania G/Electric zapewniające rozwój funkcjonalności systemu,
- G/Designer – stanowiska które umożliwiają wprowadzanie i modyfikacje gromadzonych w systemie danych.
- G/Analist – aplikacja inżynierska o identycznym interfejsie jak G/Designer, lecz bez możliwości wprowadzania i modyfikacji danych. Zaleta tego rozwiązania jest rozbudowany system analityczny i szerokie narzędzia pozwalające bardzo prosto tworzyć złożone wydruki. W bardzo krótkim czasie na stanowiskach pracy gdzie często korzysta się z GIS (komórki Rozwoju i Dokumentacji, Eksploatacji, Ruchu) aplikacja ta „wyparła" przeglądarki graficzne. Łatwość obsługi, możliwości analityczne i prezentacji danych, szybkość działania zadecydowały o sukcesie tego rozwiązania.
- G/Netviewer – przeglądarka graficzna zapewniająca on-line dostęp do danych z szerokimi możliwościami analitycznymi.
- G/Mobileviewer – przeglądarka graficzna pracująca w trybie off-line.



Rys. 1. Moduły systemu G/Electric

Do roku 2005 zostały wdrożone następujące moduły systemu:

- G/Electric, system paszportyzacji sieci oparty o wektorowe mapy terenu,
- moduł AWARIE służący do rejestracji i analizy awarii w sieciach energetycznych,
- moduł PRZYŁĄCZANIE ODBIORCÓW wspomagający proces przyłączania odbiorców do sieci elektroenergetycznej.

Pierwszym przykładem wykorzystania G/Mobileviewer są wozy pomiarowe (rys.2.), wyposażone w aparaturę przy pomocy której dokonywane są pomiary i lokalizacje uszkodzeń w liniach kablowych. Dodatkowy laptop z zainstalowanym G/Mobileviewer daje dostęp do informacji o sieci energetycznej. Sieci w obszarze miast odwzorowane są na pełnych mapach wektorowych. G/Electric zapewnia dostęp do informacji o trasach linii kablowych, o kolizjach z uzbrojeniem podziemnym, daje możliwość dokonania pomiarów np. długości linii kabla czy też odległości linii kablowej od innych elementów infrastruktury. Poprawia się w ten sposób komfort pracy zarówno w trakcie lokalizacji uszkodzeń jak i przy robotach ziemnych (informacje o kolizjach), ulega skróceniu czas lokalizacji uszkodzeń. Bardzo ważne jest przekazywanie przez pracowników obsługujących wozy pomiarowe informacji o ewentualnych różnicach, zaobserwowanych między danymi w systemie komputerowym a rzeczywistością. Taka weryfikacja jest bardzo cenna i poprawia jakość bazy danych.

Rys. 2. Praca na wozie lokalizacyjnym z wykorzystaniem systemu GIS

Kolejny przykład to testowany obecnie system wspomagający prace zespołów pogotowia energetycznego. System ten to specjalny „laptop terenowy" połączony z GPS (rys. 3.). Cały system wyposażony jest w oprogramowanie G/Mobileviewer umożliwiające dostęp do informacji o sieciach energetycznych (wraz ze schematami złącz i stacji transformatorowych). W G/Electric do każdego przyłącza lub złącza dołączone są także informacje o odbiorcach zasilanych z tego punktu (są to informacje pobrane z systemu bilingowego). Dzięki temu wyszukując w systemie odbiorcę można zlokalizować budynek czy punkt poboru. Jest to szczególnie ważne na obszarach wiejskich, gdzie do tej pory nawet bardzo dobra znajomość terenu nie zawsze wystarczała, aby szybko i sprawnie dotrzeć do urządzeń energetycznych lub odbiorcy zgłaszającego reklamację. Dodatkowe oprogramowanie to AUTOMAPA dzięki której po określeniu w G/Electric współrzędnych wybranego elementu sieci lub miejsca zamieszania odbiorcy, można wyznaczyć trasę dojazdu.

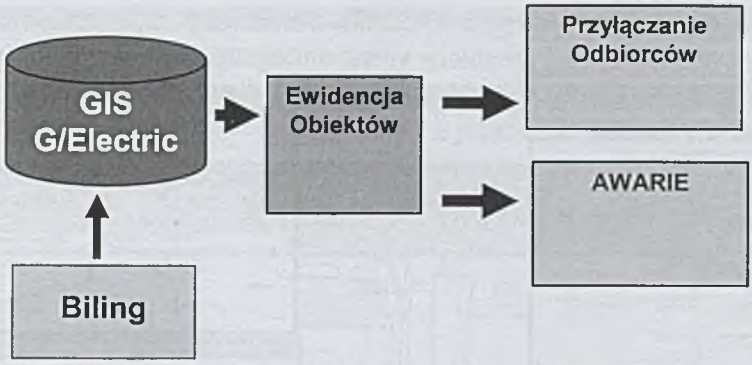
Rys. 3. Laptop zamontowany w samochodzie Pogotowia Energetycznego

3. Paszportyzacja sieci

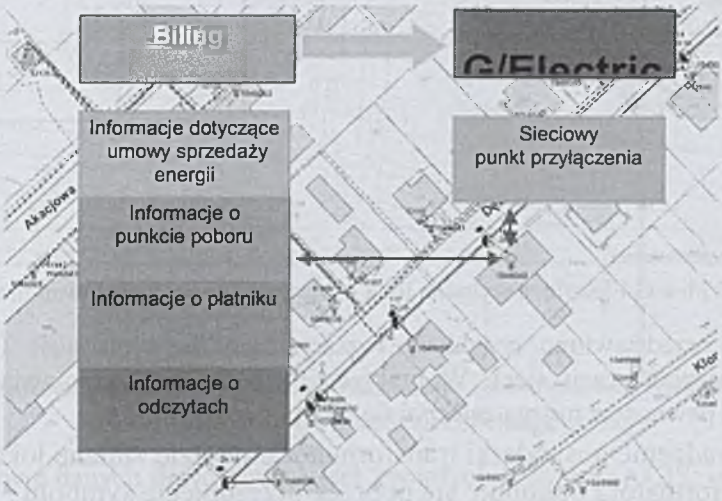
Paszportyzacja sieci jest na ogół pierwszą wdrażaną i wykorzystywaną funkcją systemu GIS w przedsiębiorstwach sieciowych. W jednej bazie danych znajdują się informacje o wszystkich elementach sieci energetycznej, dla wszystkich poziomów napięć. W oparciu o inwentaryzację sieci w terenie, posiadane dokumenty, pomiary lokalizacji w terenie z wykorzystaniem GPS oraz dane otrzymane z prac eksploatacyjnych została zbudowana baza danych opisująca sieć w zakresie danych technicznych, lokalizacji urządzeń i topologii. Inwentaryzacja linii kablowych została wykonana w porozumieniu z Urzędem Miejskim: spółka dystrybucyjna wykorzystując wóz pomiarowy określała trasy kabli a przedsiębiorstwo geodezyjne na zlecenia miasta dokonywało inwentaryzacji. Istotnym elementem tych danych są obiektowe schematy w których każde z urządzeń posiada swój symbol, dane techniczne i miejsce w topologii. Aby zapewnić użyteczność tych danych wprowadzono procedury zapewniające bieżącą aktualizację danych. Utrzymywanie aktualnej bazy danych i zapewnienie każdemu z pracowników technicznych dostęp do nich na stanowisku pracy (odpowiednia do potrzeb aplikacja) w znaczący sposób poprawiło efektywność i jakość pracy. Przełomowym momentem dla wdrożenia GIS było uznanie przez pracowników tej bazy danych jako najdokładniejszej zbiór łatwo dostępnych informacji. Dało to możliwość odejścia od poprowadzenia różnego rodzaju planów, schematów z układem sieci a paszporty poszczególnych urządzeń stały się archiwum dokumentów z którego korzysta się w wyjątkowych sytuacjach. Każdy z pracowników stał się cennym źródłem informacji – korzystające z systemu osoby starają się na bieżąco zgłaszać wszystkie zmiany.

4. Wdrożone funkcje systemu GIS

Rozszerzenie funkcjonalności wprowadzanego systemu to przede wszystkim powiązanie GIS z innymi systemami komputerowymi. Schemat takiego powiązania przedstawiono na rys.4. W rozpoczętej budowie wymiany informacji szczególną rolę odgrywa przygotowany interfejs do systemu bilingowego. Pozwolił on na powiązanie odbiorców energii elektrycznej z miejscem dostawy. W GIS wprowadzony został obiekt: sieciowy punkt przyłączenia którego wyróżnikiem jest unikalne ID (rys. 5.) Wartość ID wprowadzona jest obecnie w systemie bilingowym w opisie punktu poboru. Takie połączenie gwarantuje jednoznaczne określenie lokalizacji poboru energii elektrycznej, co będzie wykorzystane do lokalizacji odbiorców w terenie, określania rzeczywistego zużycia energii w danych węzłach sieci na potrzeby obliczeń technicznych, analiz obszarów bilansowania energii z wykorzystaniem topologii sieci. Rozwiązanie to wykorzystywane jest w wyżej zaprezentowanym wspomaganiu pracy zespołu pogotowia energetycznego.

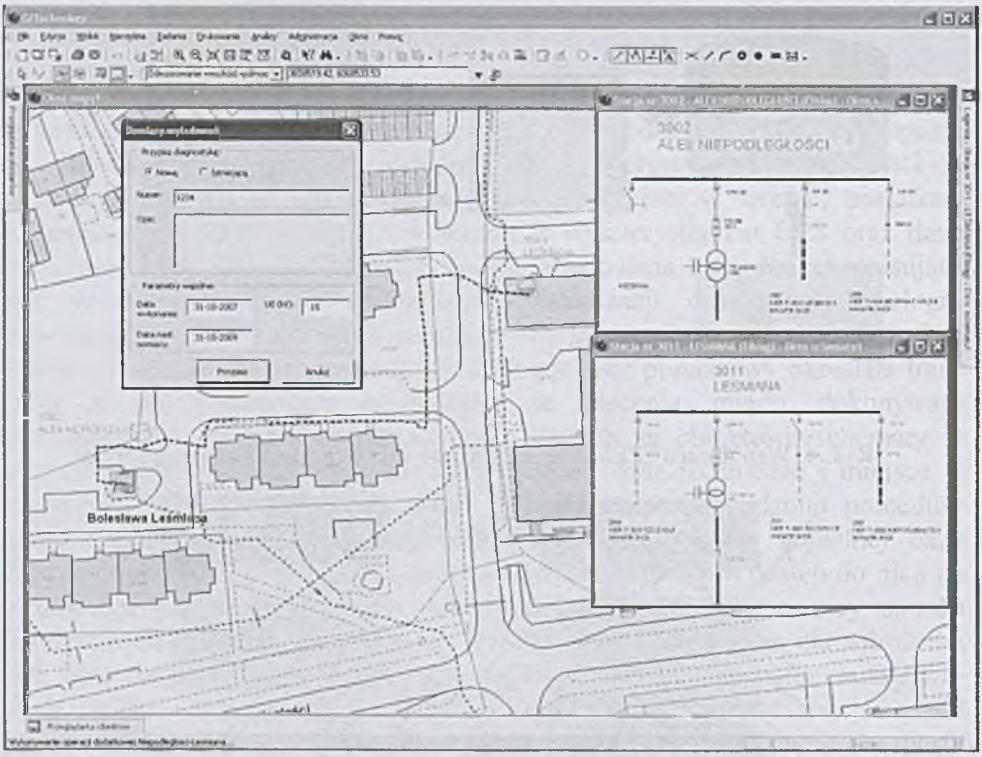


Rys. 4. Współpraca G/Electric z pozostałymi systemami komputerowymi



Rys. 5. Powiązanie G/Electric z bilingiem poprzez sieciowy punkt przyłączenia

Kolejną wprowadzoną funkcją systemu GIS była rozbudowa bazy danych o informacje dotyczące dokonanych pomiarów na elementach sieci. Jako przykład rozbudowanej funkcjonalności w tym zakresie są pomiary wyładowań niezupełnych w liniach kablowych średniego napięcia (SN) i śledzenie procesu starzenia się kabli. Przygotowana dodatkowa funkcja w GIS pozwala automatycznie wyszukać wszystkie elementy linii kablowej (głowice, mufy, odcinki kabla) i opisać je informacjami otrzymanymi z analizy wyników pomiaru. W prowadzonej diagnostyce istotnym elementem jest posiadanie informacji historycznych (o liniach unieczynnionych, przebudowanych) które wpływają na jakość diagnozy. Wymusiło to przebudowę GIS i wprowadzenie historii życia linii kablowych. Rozbudowana funkcjonalność analityczna G/Electric służy do analizy gromadzonych danych i opracowywania standardów oceny stanu technicznego linii kablowych.



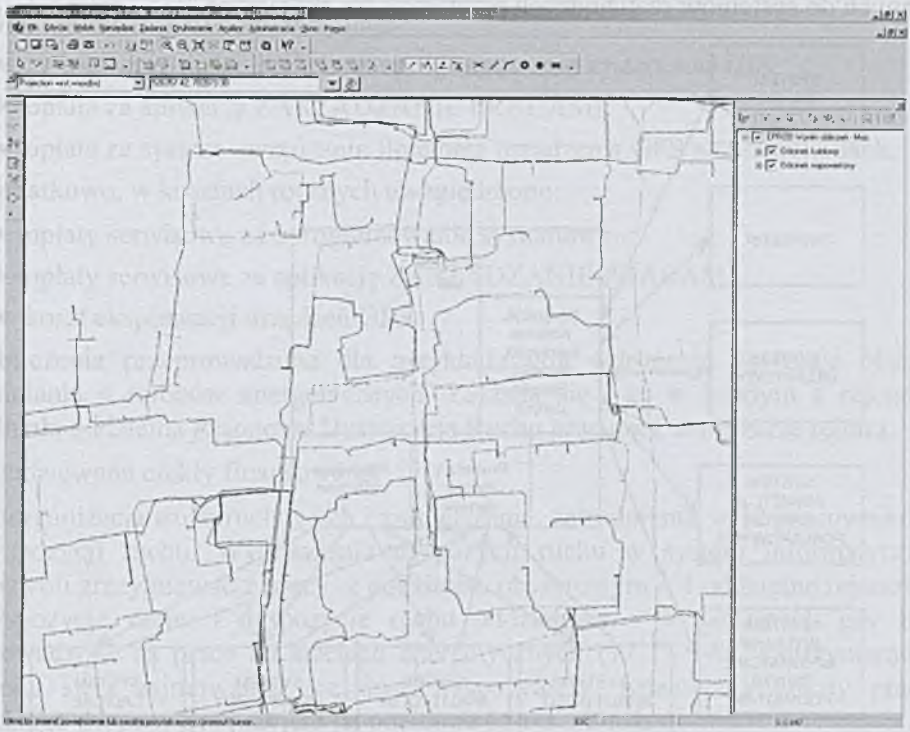
Rys. 6. Widok G/Electric z wybrana linią kablową do opisu wyładowań niezupełnych

Powyżej przedstawiono przykłady rozszerzające funkcjonalność GIS ponad system paszportyzacji sieci. W praktyce tych elementów pojawia się coraz więcej. Z pewnością można zaliczyć do nich:

1. Prowadzenie gospodarki transformatorami gdzie zmianę lokalizacji transformatora dokonuje się poprzez przesuniecie symbolu transformatora pomiędzy schematami stacji lub magazynem.
2. Wspomaganie bilansowania energii elektrycznej w obrębie stacji transformatorowych.
3. Wspomaganie rozliczania zajęcia pasa drogowego. GIS umożliwia uporządkowanie corocznych opłat za zajmowanie pasa przez urządzenia energetyczne.
4. Udostępnienie danych o wybranych fragmentach sieci biurom projektowym przygotowującym na zlecenie przedsiębiorstwa projekty budowlane związane z rozbudową sieci (wykorzystanie G/Mobilewiever).

Istotnym elementem systemu informatycznego wspomagającego procesy zarządzania w spółce dystrybucyjnej jest system analiz technicznych (obliczenia przepływów w elementach sieci, spadków napięć, strat, estymacja obciążeń, estymacja stanu sieci, prognozowanie obciążeń, obliczenia zwarciowe, obliczenia optymalizacyjne, obliczenia dotyczące bezpieczeństwa sieci i

użytkowników) [2,3]. System analiz technicznych powinien obejmować źródła energii odnawialnej (szczególnie farmy wiatrowe, które mogą przysporzyć dużo problemów na etapie analiz a także planowania rozwoju sieci) oraz uwzględniać generację rozproszoną.

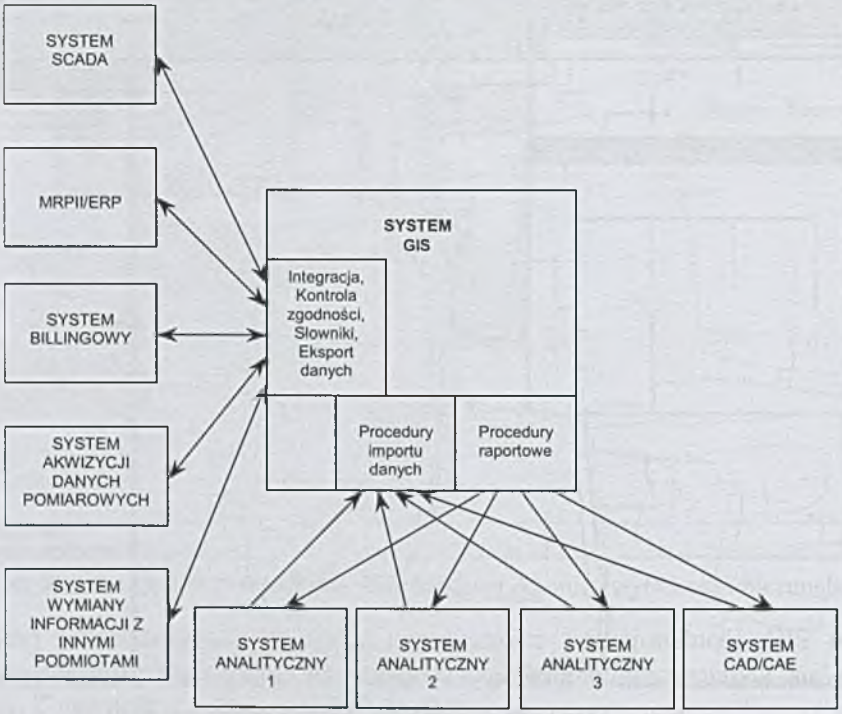


Rys.7. Mapa obciążenia sieci elektroenergetycznej

Na podstawie danych dotyczących sieci zawartych w systemie GIS, w systemie analiz technicznych tworzone jest odpowiednie odwzorowanie sieci a po wykonaniu obliczeń ich wyniki wprowadzane są do baz danych systemów GIS. Dzięki takiemu podejściu możliwa jest wizualizacja wyników analiz na mapie – także generowanie raportów, wykresów i zestawień realizowane byłoby w systemie GIS. Oznacza to konieczność opracowania dwukierunkowego interfejsu między systemem GIS a systemami analiz technicznych w celu osiągnięcia maksymalnych korzyści z łączenia takich systemów. Na rys. 7 podano przykład mapy obciążenia linii elektroenergetycznych – kolor linii zależy od stopnia obciążenia linii.

Wyniki otrzymywane z systemu analiz technicznych wykorzystane mogą być do wspomagania procesu minimalizacji strat sieciowych (wyznaczanie optymalnych punktów rozcięć), lokalizowania źródeł nieuzasadnionych strat mocy i energii elektrycznej, planowania rozwoju sieci czy analizy efektywności inwestycji.

Można założyć, że w przyszłości system GIS będzie stanowił centralny punkt wprowadzania i koordynacji danych oraz tworzenia raportów, jak pokazano na rys. 8 [1]. Dodatkowym zadaniem systemu GIS będzie integracja danych pochodzących z systemów analitycznych oraz przekazywanie części tych danych do systemów realizujących przetwarzanie operacyjne.



Rys. 8. Koncepcja współpracy między systemem GIS a innymi systemami informatycznymi

5. Ocena efektywności rozbudowy komputerowego systemu wspomagania działalności dystrybucyjnej

Ocenę efektywności przeprowadzono dla wdrożenia nowego modułu – ZARZĄDZANIE AWARIAMI (Outages Management System) wyposażonego w następujące funkcje:

- obsługę awarii (przyjęcie zgłoszenia, lokalizację zgłoszenia na mapie, zlecenie wykonania pracy, planowanie trasy dojazdu, rejestracja danych o awarii),
- planowanie i obsługa prac planowych,
- analityczne stanowisko dyspozytorskie (wspomaganie pracy dyspozytorów ruchu),
- stanowiska dla zespołów pracowników,

- moduł powiadamiania odbiorców o występujących przerwach w dostawie energii elektrycznej,
- stanowiska do analizy gromadzonych danych.

W oparciu o wstępne oferty przyjęto koszt inwestycyjny wdrożenia w wysokości 1400 tys. PLN. Koszt ten obejmuje:

- zwiększenie ilości licencji na oprogramowanie systemowe GIS,
- opłata za aplikację ZARZĄDZANIE PRACAMI,
- opłata za system zarządzania flotą oraz urządzenia GPS na samochodach.

Dodatkowo, w kosztach rocznych uwzględniono:

- opłaty serwisowe za oprogramowanie systemowe,
- opłaty serwisowe za aplikację ZARZĄDZANIE PRACAMI,
- koszt eksploatacji urządzeń GPS.

Obliczenia przeprowadzona dla przykładu gdy wdrożenie obejmuje obszar działania 4 rejonów energetycznych. Zakłada się , że w każdym z rejonów istniała oddzielna Rejonowa Dyspozycja Ruchu pracująca w obszarze rejonu.

Spodziewane efekty finansowe to:

Reorganizacja służb ruchowych i zmniejszenie zatrudnienia o 12 pracowników dyspozycji ruchu. Wyposażenie dyspozycji ruchu w system informatyczny pozwoli zrezygnować z pracy z podziałem obszarowym – 4 oddzielne rejonowe dyspozycje ruchu i dyspozycje ruchu oddziałową. W godzinach gdy nie prowadzone są prace na sieciach energetycznych ($17^{00} - 6^{00}$) utrzymaniem ruchu sieci zajmowaliby się dwaj dyspozytorzy. Szacowany koszty pracy jednego dyspozytora przyjęto na poziomie 65 tys. zł/ rok.

System komputerowy będzie zawierał w sobie także zarządzanie flotą w oparciu o GPS. Poprawa organizacji pracy w tym zakresie. Opierając się o dane z wdrożeń innych systemów zarządzania flotą, planuje się zmniejszyć wykonywane rocznie przejazdy samochodów o ok. 10%. Daje to oszczędności w Oddziale w wysokości 100 tys. zł/rok.

Dodatkowo założono (ze względów bezpieczeństwa projektu):

Wzrost wartości inwestycji o 10%.

Etapowanie reorganizacji służb ruchowych ze względu na sprzeciw załogi (związków zawodowych). Zakłada się, że w pierwszym roku z pracy w służbach ruchowych odejdzie sześciu dyspozytorów a w kolejnym roku kolejnych sześciu.

W związku z etapową centralizacją służb ruchowych wdrażanie systemu także odbędzie się w przeciągu 2 lat. Pierwszy okres to wdrożenie systemu w dwóch rejonach energetycznych i centrali Oddziału. Drugi etap to wdrożenie systemu w pozostałych dwóch rejonach.

Analizę finansową przeprowadzono dla okresu 3 lat. Wyniki analizy przedstawiono w tabelach 1 i 2 [4,5].

Tab. 1. Przepływy pieniężne (FCF –Free Cash Flow) w tys. zł

Lata	1	2	3	4
Przychody	490,0	880,0	880,0	880,0
Koszty (bez amortyzacji)	60,5	151,0	151,0	151,0
EBITDA	429,5	729,0	729,0	729,0
Amortyzacja	383,3	466,7	466,7	83,3
EBIT	46,2	262,3	262,3	645,7
Podatek (19%)	8,8	49,8	49,8	122,7
EAT (wynik po opodatkowaniu)	37,4	212,5	212,5	523,0
Amortyzacja	383,3	466,7	466,7	83,3
CF brutto	420,7	679,2	679,2	606,3

Tab. 2. Wyznaczenie wartości bieżącej projektu

Współczynnik dyskontowy	1	0,893	0,797	0,712	
CFC		420,7	679,2	679,2	PV
CFC dyskontowane		375,7	541,4	483,4	1400,5
Wydatki inwestycyjne	-1150	-250			
Wydatki inwestycyjne zdyskontowane	-1150,0	-223,2			-1373,2

Na podstawie podanych wyników wyznaczono:

NPV = 1400,5- 1373,2 = 27,3 tys. zł

IRR = 13,1 %

Inwestycja ta analizowana była dla okresu pełnej amortyzacji, jednak można się spodziewać że przychody finansowe będą generowane także przez następne lata. Dodatkowo należy spodziewać się w tej chwili trudnych do oszacowania finansowych efektów wynikających z poprawy organizacji pracy, efektywniejszego zarządzania majątkiem.

6. Podsumowanie

W ENERGA-OPERATOR SA Oddział w Elblągu wdrożenie GIS zakończyło się sukcesem. System ten stał się narzędziem wspomagającym wiele stanowisk pracy. Rozszerzona funkcjonalność ponad system paszportyzacji sieci daje nowe możliwości wspomagania zarządzania majątkiem sieciowym. Szczegółowa baza

danych wspomaga podejmowanie decyzji związanych z prowadzeniem eksploatacji i rozwojem sieci.

Na podstawie doświadczeń można wyróżnić trzy etapy we wprowadzaniu systemu GIS w spółce dystrybucyjnej:

1) System GIS jako system paszportyzacji – mniej więcej jest to etap wprowadzania danych, czyli tworzenia numerycznego odwzorowania sieci elektroenergetycznej
2) Rozbudowa systemu o proste narzędzia analityczne – np. analiza awarii, wspomaganie wydawania warunków przyłączenia (w sensie wspomagania procesu obiegu dokumentów), łączenie systemu GIS z systemem billingowym
3) Wprowadzanie systemów analitycznych bądź wiązanie istniejących systemów z systemem GIS.

Dodatkowo, w każdym z trzech etapów wdrażania systemu może pojawić się udostępniane danych i elementów procesów za pośrednictwem internetu. Określane jest to jako najwyższy poziom zasięgu systemu informatycznego – udostępnianie danych całemu społeczeństwu. Prowadzić to by mogło np. do bardzo dużej transparentności procesów decyzyjnych – w niektórych zagadnieniach może być to społecznie uzasadnione a także zdecydowanie wpływać na pozytywny wizerunek przedsiębiorstwa.

Wprowadzenie systemów GIS do zarządzania infrastrukturą sieciową, wyposażonych we wspomniane moduły analityczne będzie niewątpliwie miało istotny wpływ na wzrost efektywności (między innymi przez spodziewaną obniżkę kosztów eksploatacji) systemów energetycznych. Należy jednak zauważyć, że ocena nakładów powinna być przeprowadzana oddzielnie dla każdej spółki dystrybucyjnej wdrażającej systemy GIS, mając na uwadze iż koszty pozyskania danych mogą sięgnąć nawet do 80% kosztów wdrażania systemu. Ocena poniesionych nakładów i efektywności ich wykorzystania pozwoli także na planowanie rozwoju przedsiębiorstwa sieciowego (a ściślej planowanie rozwoju sieci elektroenergetycznej) uwzględniając uwarunkowania rynkowe.

Wykorzystanie systemów informatycznych do zarządzania wiedzą to np. system korzystający z informacji zgromadzonych w GIS, SCADA i systemie bilingowym a następnie udostępniający je dyspozytorom i elektromonterom oraz zapewniający odpowiednią komunikację. Z dokonanej analizy finansowej takiego przedsięwzięcia wynika, że inwestycja taka zwraca się w okresie 3 lat. Nie można tu zapomnieć o efektach niepoliczalnych, takich jak: odzyskanie wykształconego i doświadczonego personelu, skrócenie przerw w dostawie, zadowolenie klientów, większa pewność decyzji oparta o zgromadzona wiedzę.

LITERATURA

1. Baczyński D., Helt P.: System GIS jako hurtownia danych technicznych w spółce dystrybucyjnej. Olejniczak Z., Nowak J., Grabara J. – „Informatyka. Strategie i zarządzanie wiedzą", PTI, Katowice 2005, ISBN
2. Helt P.: Wdrażanie systemów obliczeń technicznych w spółkach dystrybucyjnych. Systemy informatyczne – Zastosowanie i wdrożenia 2002, Tom II, Wydawnictwo Naukowo – Techniczne, Warszawa – Szczyrk 2002
3. Helt P., Baczyński D.: Funkcje systemu GIS w zakładach energetycznych, IX Międzynarodowa Konferencja Naukowa APE'99 "Aktualne Problemy w Energetyce", Jurata 1999.
4. Shim J., K., Siegel J., G.: Dyrektor Finansowy. Oficyna Ekonomiczna, Kraków 2005
5. Sobczyk M.: Matematyka finansowa. Wydawnictwo „Placet", Warszawa 2003

Rozdział 18

Efektywność systemów informatycznych w branży meblarskiej

Dominika Biniasz
Politechnika Opolska
dominika.biniasz@data.pl

Streszczenie

W rozdziale przedstawiono rolę systemów informatycznych i ich zastosowanie w prowadzeniu działalności biznesowej branży meblarskiej. Ze względu na ochronę danych nazwy firm zostały zmienione. Na potrzeby rozważań wybrano badania na losowej grupie przedsiębiorstw, które wykorzystują technologie informatyczne w celach komunikowania się i realizacji zleceń. Podjęto próbę identyfikacji najczęściej wykorzystywanych narzędzi informatycznych w realizacji zadań biznesowych. Oszacowano efektywność ekonomiczną zastosowanych systemów informatycznych i pokazano ich przydatność w codziennej działalności badanych firm.

1. Wprowadzenie

Szybko zachodzące zmiany, uwarunkowane rozwojem nauki i techniki spowodowały, że znaczenie życiowe nowoczesnych rozwiązań w zakresie technologii informacyjno-komunikacyjnych jest ogromne, szczególnie teraz, gdy następuje okres technizacji życia. Znajomość rozwiązań ICT staje się coraz bardziej potrzebna w życiu codziennym, a tym bardziej w pracy zawodowej, czy zarządzania firmą.

Gwałtowne zmiany w przedsiębiorstwach spowodowane wprowadzaniem nowych rozwiązań technicznych, informatycznych i komunikacyjnych wymusiły na przedsiębiorcach konieczność ich zastosowania w praktyce. Zwiększająca się globalność konkurencji otworzyła nowe możliwości działania i współpracy, opartej często na partnerstwie w interesach. Zyskały na tym szczególnie mikro,

małe i średnie przedsiębiorstwa, które dzięki nowym technologiom mogły zaistnieć ponownie na rynku i konkurować nawet z firmami większymi od siebie.

Rozwój technologii informatycznej wymusza stosowanie zintegrowanych systemów zarządzania, które mają za zadanie wspomagać przedsiębiorstwo we wszystkich aspektach działalności na globalnym rynku konkurencji. Podmioty gospodarcze starannie lokują swoje pieniądze w rozwój organizacji oraz w rozwój kadry pracowniczej. Wspomaganie działalności przez odpowiednie systemy informatyczne znacznie ułatwia proces zarządzania organizacją.

2. Rola technologii informatycznych

Technologia to stosunkowo młody termin w nauce, a technologie informatyczne jeszcze młodszy. Choć to nowa nauka, bardzo prężnie się rozwija i diametralnie zmienia otaczającą nas rzeczywistość, nasz sposób działania, komunikowania się – właściwie ma wpływ na całokształt naszego życia. Rozwój społeczeństwa informacyjnego zauważony został w Unii Europejskiej w 1994 roku przez raport komisarza Unii do Spraw Przemysłu i Technologii Informacyjnych. Od tego czasu powstało wiele różnorodnych systemów informatycznych skierowanych zarówno do dużych, jak i tych najmniejszych podmiotów gospodarczych. Powstały projekty systemów informatycznych mające na celu rozwój społeczeństwa wiedzy oraz gospodarki opartej na wiedzy, która mogłaby konkurować w świecie z tak zaawansowanymi technologicznie gospodarkami, jak np. USA. Oczywiście wraz z rozwojem Unii Europejskiej powstawały również i w Polsce programy rozwoju informatycznego kraju.

Budowanie przewagi konkurencyjnej wymaga od przedsiębiorstwa nieustannego zwiększania dostępu do wiedzy i informacji. Działania takie z powodzeniem wspomagane mogą być poprzez stosowanie różnorodnych narzędzi informatycznych, które z jednej strony doskonale nadają się do walki konkurencyjnej, ale też z drugiej strony dają możliwość wewnętrznej i zewnętrznej integracji organizacji. Natomiast konieczność zapewnienia wsparcia technologicznego w przedsiębiorstwie staje się tym większa, im większa jest liczba zatrudnionych, ilość dokumentów, czy projektów.

Kluczowym momentem dla zarządzania wiedzą był rozwój technologii informatycznych, a przede wszystkim powstanie sieci komputerowych. Możliwym stało się wówczas przetwarzanie danych na większej liczbie komputerów, przez co zwiększyła się ich efektywność jako całości [4].

2.1. Systemy informatyczne – krótka charakterystyka

Istnieją różne, ale podobne definicje opisujące technologię informatyczną, czy systemy informatyczne, które zostały stworzone jako rozwiązania technologicznych problemów pojawiających się w dziedzinie informatyki.

Wielość tych problemów oraz ich zróżnicowanie implikuje wielość oraz różnorodność technologii, która ją opisuje. Ta ostatnia okoliczność utrudnia bezpośrednie porównywanie rozwiązań technologicznych, skłaniając do wprowadzenia odrębnych kategorii technologii informatycznych [por. 3].

Wprowadzono następujące kategorie technologii informatycznych [3]:

- sterowniki sprzętowe – oprogramowanie odpowiadające za dostęp do sprzętu,
- systemy operacyjne – oprogramowanie służące jako środowisko, w którym funkcjonuje system informatyczny,
- serwery aplikacji – głównie serwery oferujące różne usługi, w większości rozwiązań pokrywają zapotrzebowania logicznej warstwy przetwarzania,
- systemy zarządzania bazami danych – w istocie są serwerami oferującymi usługę dostępu do danych, lecz ze względu na ich rolę w systemach informatycznych zostały zakwalifikowane jako osobna kategoria technologii,
- platformy programowe dla aplikacji typu gruby klient – oprogramowanie odpowiadające za przetwarzanie po stronie grubego klienta,
- platforma programowe dla aplikacji typu cienki klient – oprogramowanie odpowiadające za przetwarzanie po stronie cienkiego klienta,
- platforma programowe dla aplikacji typu „smart" klient – oprogramowanie odpowiadające za przetwarzanie po stronie cienkiego klienta typu „smart",
- systemy terminalowe – oprogramowanie odpowiadające za przetwarzanie po stronie terminala. Większość systemów terminalowych posiada własne serwery, które zarządzają swoimi klientami (terminalami).

Wielorodność technologii informatycznych oraz rozproszenie geograficzne danych
i użytkowników korzystających z systemu stają się często kluczowymi problemami klientów – użytkowników w odpowiednim ich wykorzystaniu.

Rolę informatycznych systemów wspomagających zarządzanie wyjaśniają modele rozwoju technologii informatycznych. Jeden z nich, tak zwany model MIT'90 wychodzi z założenia, że środowisko informatyczne potrafi nie tylko wspomagać organizacje, ale również rozwija w znaczący sposób strategie konkurencyjności obiektu. Model ten opisuje podział na trzy typy organizacyjnej infrastruktury technologii informatycznej [1]:

- niezależny – gdzie system odgrywa tylko rolę wspomagania informatycznego, natomiast brak jest koncepcji strategicznej. Poniesione nakłady wchodzą w skład bieżących kosztów administracyjnych;
- reaktywny – pod uwagę bierze się rolę informatycznego wspomagania czego efektem jest wzrost świadomości potrzeby wykorzystania technologii informatycznych podczas wybierania i realizowania strategii działania,

natomiast pomija podczas jej kształtowania. Nakłady traktuje się, jako koszty działalności;

- współczesny – uwzględnia się długofalowy plan strategiczny i w związku z nim dokonuje się wszelkich zmian i modyfikacji. Tak więc plan strategiczny a nie reakcja na plan są przyczyna zmian. Nakłady traktuje się, jako inwestycje gospodarczą.

Według powyższego modelu MIT'90 powiązania długofalowych planów strategicznych z planem rozwoju informatycznych rozwiązań zależy od skutecznego wykorzystania technologii informatycznych w organizacji. Obecne tendencje wskazują, że większość organizacji tworzy infrastrukturę na podobieństwo pierwszych dwóch typów modeli. Jak wskazują badania w Polsce ciągle jednak najpopularniejszy jest typ niezależny. W ten sposób ryzyko wdrażania systemów wzrasta. Głównym problemem organizacji w Polsce podczas decydowania się na wprowadzenie systemu nie jest brak korzyści, ale koszty wdrożenia, niedojrzałość technologiczna oraz brak przygotowania partnerów przedsięwzięć informatycznych. Mimo tych obaw, światowe standardy systemów informatycznych sprawiają, że w obecnych czasach bariery te są coraz skuteczniej przełamywane [por. 1].

Różne rozwiązania informatyczne dedykowane są do różnych pod względem wielkości, branży, sektora i specyfiki działania firm, które dopasowując umiejętności swoich pracowników lub szkoląc ich, wdrażają nowe systemy wspomagające zarządzanie i usprawniając działalność firm.

Systemy informatyczne charakteryzują jakość i funkcjonalność oprogramowania, jego elastyczność, szybkość oraz koszty jego rozwoju, a także modyfikacji dokonywanej zarówno przez użytkownika, jak i zewnętrznych klientów Przyjmując, jako kryteria klasyfikacji cel systemu, stosowaną technologię przetwarzania, złożoność podejmowania decyzji, użyteczność informacji generowanej przez system, poziom podejmowania decyzji w hierarchii organizacyjnej, można wyróżnić szereg systemów informacyjnych zarządzania, do których zalicza się:

- systemy transakcyjne,
- systemy z bazą danych,
- systemy wspomagania decyzji,
- systemy informowania kierownictwa,
- systemy eksperckie [5].

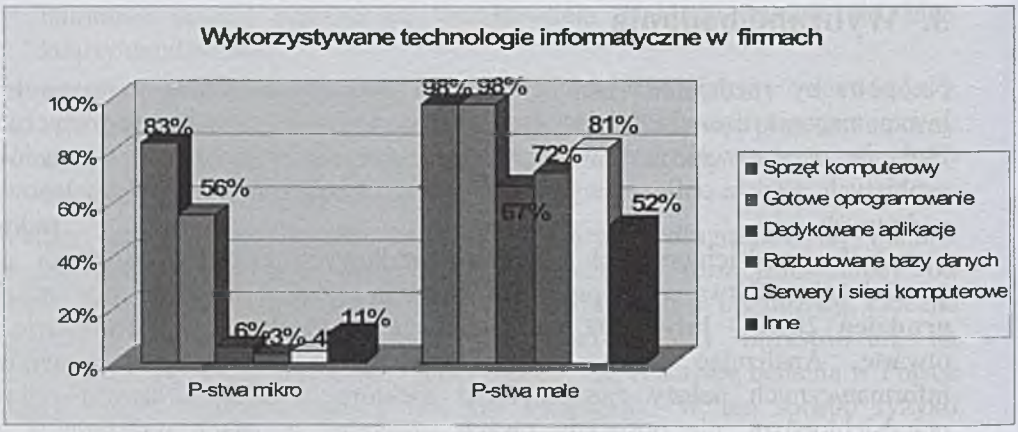
W zarządzaniu przedsiębiorstwem wszelkie systemy wspomagające działalność są postrzegane, jako konieczne i koszty związane z ich zakupem ponoszone na początku eksploatacji, w krótkim czasie procentują w zyskach firmy, co udowadniają poniższe badania.

3. Wybrane badania

Na potrzeby rozdziału wyselekcjonowano badania dotyczące zastosowania i wspomagania działalności biznesowej firm poprzez systemy informatyczne. Badania przeprowadzono w formie kwestionariuszy ankiet i wywiadów osobistych, a także obliczeń wybranych wskaźników ekonomicznych, w losowo wybranych mikro i małych przedsiębiorstwach usługowo-handlowych branży meblarskiej na terenie województwa wielkopolskiego. Badania przeprowadzono w miesiącach **grudzień 2008 – luty 2009 roku**. Ankieta zawierała pytania zamknięte i otwarte. Analizując badane firmy pod kątem posiadanych systemów informatycznych należy zaznaczyć, iż niektóre z nich, zwłaszcza mikro przedsiębiorstwa, nie posiadały owych systemów w latach poprzednich, a występowanie ich w małych przedsiębiorstwach szacowało się na około **25%**.

Skupiając się na klientach firm, poddano analizie liczbę nowych klientów, po zastosowaniu technologii informatycznych. Dzięki zastosowanej technice firmy pozyskały nowych klientów o **37%** większej w stosunku do lat ubiegłych w mikro przedsiębiorstwach, a o **54%** zwiększyła się liczba klientów w małych firmach, co świadczy o większym zainteresowaniu klientów firmami i ich oferowanym asortymentem oraz o polepszającym się systemie zarządzania relacjami z klientami. Miała na to wpływ również odpowiednia motywacja pracowników do zwiększenia zaangażowania w pozytywne relacje z klientami i ich szkolenia, mające na celu przygotowanie ich do korzystania z technologii informatycznych. Korzystanie z technologii informatycznej wpłynęło korzystnie na funkcjonowanie analizowanych przedsiębiorstw i ich wizerunek w stosunku do klientów.

Wśród 10 analizowanych przedsiębiorstw wszystkie zadeklarowały wykorzystywanie technologii informatycznej do prowadzenia działalności biznesowej, rysunek 1 pokazuje procentowe udziały firm w zastosowanych technologiach informatycznych.



Rys. 1. Wykorzystywane technologie informatyczne w badanych firmach

Widać, iż aż **83%** firm wielkości mikro wykorzystuje sprzęt komputerowy w codziennej działalności, w przypadku małych firm jest to **98%**. Firmy kupują gotowe oprogramowanie i w przypadku miko przedsiębiorstwa jest to **56%**, a w przypadku małych firm **98%**. Firmy korzystają z gotowych systemów informatycznych i to nie tylko administracyjnych, ponieważ coraz większego znaczenia nabierają systemy wspomagające działalność produkcyjną i usługowo-handlową w analizowanych firmach. Niestety na dedykowane aplikacje ma zapotrzebowanie zaledwie **6%** mikro przedsiębiorstw, ale już w małych widzimy, że jest to zwiększone do **67%**. Są to niewielkie programy pozwalające sprawniej obsługiwać klientów, czy kontrolować stany magazynowe i planować transport. Inaczej sprawa wygląda z rozbudowanymi bazami danych, tu wiodące są małe firmy, które w **72%** takie posiadają, natomiast w mikro firmach bazy danych pojawiają się tylko u **3%** badanych. Podobnie sprawa wygląda w przypadku korzystania z serwerów i sieci komputerowych, tu również małe firmy częściej, bo w **81%** wykorzystują je w swojej działalności, natomiast mikro przedsiębiorstwa tylko w **4%**. Pozostałe technologie, z których korzystają badane firmy meblarskie stanowią **11%** wśród mikro przedsiębiorstw i aż **52%** wśród małych firm, zalicza się do nich korzystanie
z Internetu, poczty elektronicznej, komunikatorów, itp.

Można zaobserwować, iż poddane badaniom firmy znają znaczenie technologii informatycznej i sprawnie wykorzystują jej zasoby do prowadzenia swojej działalności. Rozwijający się rynek informatyczny oferuje wiele nowych, ciekawych rozwiązań zarówno dla dużych przedsiębiorstw, jak i dla tych małych, czy rodzinnych firm, które coraz śmielej wkraczają na drogę globalnej konkurencji. Świadomość swoich możliwości mogły sprawdzić analizowane firmy w wynikach, jakie otrzymały po wprowadzeniu do swojej działalności systemów informatycznych, patrz rysunek 2. Badane przedsiębiorstwa zwiększyły szybkość obsługi klientów o **78%** w przypadku mikro

przedsiębiorstw i o **96%** w małych. Ich znaczenie wśród konkurencji wzrosło o **69%** dla mikro i o **85%** dla małych przedsiębiorstw, co poprawiło wizerunek firmy i wzmocniło jej markę. Również obsługa klienta znacząco się poprawiła i tak dla mikro firm o **87%**, a dla małych **96%**, co przejawia się również w lepszej wiedzy o potrzebach samych klientów. Wiedza firm o potrzebach klientów wzrosła aż o **65%** w mikro i o 93% w małych firmach. pozycji wśród konkurencji.

Rys. 2. Wyniki osiągane przy wykorzystaniu systemów informatycznych w badanych firmach

Otrzymane wyniki świadczą, iż analizowane firmy zaczęły coraz śmielej i częściej stosować technologie informatyczne w swojej działalności. W badaniach analizowanych mikro i małych przedsiębiorstw branży meblarskiej ustalono, iż nie wyobrażają one sobie prowadzenia biznesu bez informatyzacji. Należy wskazać pozytywne zjawisko, polegające na tym, iż wszystkie te firmy posiadają oraz wspomagają swoją działalność przy użyciu technologii informatycznej,
a w szczególności systemów informatycznych.

Chcąc dokładniej sprawdzić wiarygodność otrzymanych danych ankietowych dokonano również analizy wybranych wskaźników ekonomicznych scharakteryzowanych w poniższym rozdziale.

3.1. Próba oszacowania efektywności systemów informatycznych

Kolejnym elementem analizy wykorzystywania systemów informatycznych w działalności firm i wpływ na ich działalność jest obliczenie wskaźnika „*elastyczności dostaw*", „*gotowości do świadczenia dostaw*" oraz „*jakości dostaw*". Dane uzyskane do obliczeń wskaźników zostały podane przez badane firmy po analizie ich dokumentacji.

Badania zrealizowano na tych samych **10** losowo wybranych mikro i małych przedsiębiorstwach usługowo-handlowych branży meblarskiej na terenie województwa wielkopolskiego. Wyselekcjonowane firmy, dla których policzono kolejne wskaźniki opisane w tabelach. Badania realizowano w dwóch latach, w okresie na **koniec grudnia 2007 i na dzień 01.02.2009.** Tabela 1 zawiera dane procentowe otrzymane z obliczeń wskaźnika „elastyczności dostaw".

Tab. 1. Dane dotyczące wskaźnika elastyczności dostaw badanych firm. Opracowanie własne na podstawie danych otrzymanych z firm [por. 2]

WSKAŹNIK	FIRMA	2007	2009
ELASTYCZNOŚĆ DOSTAW $\frac{\textit{liczba zrealizowanych życzeń klientów}}{\textit{całkowita liczba zapotrzebowań}} * 100\%$	1	78,9%	**85%**
	2	75%	**86,7%**
	3	93,02%	**96,67%**
	4	82%	**89,5%**
	5	76,9%	**83%**
	6	94,3%	**96,2%**
	7	91%	**96,4%**
	8	87,2%	**91,6%**
	9	81,8%	**89,3%**
	10	90,8%	**94%**

Wskaźnik „elastyczności dostaw" świadczy o tym, czy firma jest w stanie zrealizować indywidualne zlecenie klienta, który ma specyficzne zapotrzebowanie. Jest otwarta na propozycje i sugestie klienta, zmierzające do realizacji zamówienia w taki sposób, aby był on jak najbardziej zadowolony i chciał dalej współpracować z daną firmą.

Analizowane w badaniach przedsiębiorstwa są gotowe do świadczenia dostaw w zależności od indywidualnych potrzeb danego klienta. Dzięki zbieraniu, przechowywaniu i analizowaniu wiedzy o kliencie stały się bardziej elastyczne i ukierunkowane na dążenie do wspólnych celów z klientem, który gwarantuje jego zadowolenie. Wzrost wskaźnika „elastyczności dostaw" pokazuje, w jakim stopniu firmy są gotowe realizować indywidualne życzenia klientów i je zaspokajać. Jak widać na wynikach umieszczonych w tabeli 1, wzrost analizowanego wskaźnika świadczy o tym, iż firmy zaczęły w większym stopniu zwracać uwagę na to, co dany klient chce od nich zakupić. W największym stopniu poprawiły swoje wyniki firmy, których wskaźnik elastyczności dostaw w roku 2007 szacował się na ok. 80%, gdyż ich dane **w roku 2009** są znacząco wyższe i klasyfikują się w okolicach **90%,** są to **Firmy 4, 8 i 9**. Największy skok wielkości wskaźnika odnotował **Firma 2**, która z 75% w 2007 roku, uzyskała **w roku 2009** aż **86,7%**. Bardzo dobrą pozycję zajęły w tych badaniach **Firmy 3, 7 i 10**, których wskaźnik w roku 2007 osiągał **ponad 90%**, a w roku **2009** jeszcze bardziej umocniły swoją pozycje

podwyższając ten wynik prawie do **100%**, co może świadczyć o tym, iż przedsiębiorstwa te realizują swoje dostawy zgodnie z życzeniami klientów i w oparciu o ich indywidualne preferencje.

Kolejnym etapem jest obliczenie wskaźnika „gotowości do świadczenia dostaw".

Analizowane wyniki wskaźnika „gotowości do świadczenia dostaw" ukazują, w jakim udziale procentowym dana firma była lub jest w stanie realizować dostawy z własnych zapasów magazynowych na przełomie jednego roku, patrz tabela 2.

Tab. 2. Dane dotyczące wskaźnika gotowości do świadczenia dostaw badanych firm. Opracowanie własne na podstawie danych otrzymanych z firm

WSKAŹNIK	FIRMA	2007	2009
GOTOWOŚĆ DO ŚWIADCZENIA DOSTAW $\frac{\textit{zapotrzebowanie realizowane w magazynie}}{\textit{całkowita liczba zapotrzebowań}} * 100\%$	1	86%	**92,1%**
	2	87,6%	**90,8%**
	3	95,3%	**98,3%**
	4	84,2%	**89%**
	5	76,8%	**80,2%**
	6	85%	**87%**
	7	94%	**97,5%**
	8	86,7%	**89,5%**
	9	82%	**84,3%**
	10	96,5%	**98%**

Z powyższej tabeli wynika, iż w porównaniu z rokiem 2007 firmy podniosły poziom realizacji dostaw świadczonych z ich magazynów. Należy jednak wspomnieć, iż każda z analizowanych firm posiada odmienny rodzaj świadczonych usług i różną wielkość firmy, co ma zasadniczy wpływ na otrzymane wyniki wskaźników. Branża meblarska jest bowiem bardzo specyficzną branżą, która ma różne sezony sprzedaży i bardzo różny asortyment. Widzimy, iż największą gotowość do świadczenia dostaw osiągnęły **Firmy 3,7 i 10**, które prężnie rozwijają się na rynku i w roku **2009** osiągnęły wynik bliski **100%**. Znalazły się jednak i takie przedsiębiorstwa, które do tej pory jeszcze nie radzą sobie dobrze z zapotrzebowaniem na ich towary, jak **Firma 5 i 9**. Nie posiadają własnych magazynów lub nie mają wystarczających mocy przerobowych, które umożliwiłyby im optymalne świadczenie dostaw z własnych magazynów, ich wynik przekracza w **2009** roku nieznacznie **80%**. Pomocne w realizacji dostaw stały się dla tych firm systemy informatyczne, dzięki którym mogą one realizować dostawy szybciej, dokładniej i zgodnie z indywidualnymi wymogami klienta. Badane przedsiębiorstwa ukazywały liczne dodatnie cechy systemów informatycznych działających w ich siedzibach, które

sprawiają, że zarządzanie, produkcja, sprzedaż i dostawa towarów do klienta stały się prężniejsze, łatwiejsze w wykonaniu i mniej pracochłonne.

Następny wskaźnik „jakości dostaw" odpowie nam na pytanie: ile procent ich realizacji, to reklamacje, które zgłaszają poszczególni klienci firmy?

Z danych zamieszczonych w tabeli 3 wynika, iż przedstawione w badaniach przedsiębiorstwa znacząco poprawiły swój poziom świadczonych usług i zmniejszyły średnią liczbę występujących reklamacji w ciągu roku. Jest to bardzo zadowalający wynik i świadczy o ogromnej kulturze i odpowiedzialności firm za ich wysyłki.

Tab. 3. Dane dotyczące wskaźnika jakości dostaw badanych firm. Opracowanie własne na podstawie danych otrzymanych z firm

WSKAŹNIK	FIRMA	2007	2009
JAKOŚĆ DOSTAW $\frac{\textit{liczba reklamacji}}{\textit{całkowita liczba zapotrzebowań}} * 100\%$	1	2%	**2,4%**
	2	2,7%	**2,3%**
	3	1,5%	**1%**
	4	2,3%	**1,9%**
	5	4%	**4,2%**
	6	3,2%	**2,9%**
	7	1,8%	**1,1%**
	8	2,6%	**2%**
	9	2,3%	**1,9%**
	10	1,2%	**0,5%**

Pozycja wśród konkurencji jest znacząca pod względem rozpoznawalności marki
i rozpiętości obszarów rynków zbytu. Osiągnięcie lepszego znaczenia tego wskaźnika procentuje zyskami dla firm i zdobyciem poważania wśród nawet największych konkurentów na rynku. Bardzo ważna w działalności firmy jest obsługa klientów
i wiedza o ich preferencjach. Uzyskane przez firmy meblarskie wysokie wskaźniki związane z klientami oznaczają profesjonalne podejście do klienta i chęć zaspokajania jego potrzeb w sposób perfekcyjny i szybki. Pozwalają na to zastosowane w firmach systemy informatyczne, dzięki którym mogą o każdej porze komunikować się z klientem, przekazywać najnowsze pomysły i wzory swojego asortymentu, projektować i podnosić jakość. Ma to odzwierciedlenie w zakupach klientów i zyskach firm, które dzięki nim generują. Najlepszą jakość dostaw osiągnęły **Firmy 3, 7 i 10**, których wskaźnik oscyluje wokół **1%**. Natomiast najgorszy wynik uzyskała **Firma 5**, która w roku **2009** podwyższyła liczbę reklamacji na swoje towary, uzyskując wartość wskaźnika jakości dostaw na poziomie **4,2%**. Powodem takiej sytuacji była zbyt częsta rotacja pracowników i czasowy kryzys na rynku.

Krótko podsumowując otrzymane dane z tabel, należy nadmienić, iż wybrane przedsiębiorstwa chciały podzielić się swoimi wynikami w przeświadczeniu o swoim wysokim poziomie. Możliwy był on do osiągnięcia dzięki zastosowaniu do prowadzenia swojej działalności systemów informatycznych i zmianie polityki relacji z klientami. Należy również dodać, iż analizowane firm skupiają się w województwie wielkopolskim, które nazywane jest w branży „zagłębiem meblowym", gdzie trudno rozpocząć prowadzenie interesu i gdzie trzeba bardzo się starać, aby go utrzymać i dobrze prosperować. Co widać udało się osiągnąć firmom poddanym powyższej analizie badawczej.

4. Podsumowanie

Reasumując otrzymane wyniki ankiet i wskaźników ekonomicznych należy zauważyć, iż branża meblarska jest specyficzną branżą i prowadzenie przez nią działalności przy wspomaganiu technologii informatycznej jest postrzegane wśród mniejszych przedsiębiorców, jako dodatkowe, lecz niekoniecznie potrzebne i przydatne, narzędzie pracy. Występuje to zjawisko głównie w firmach mikro i małych rodzinnych, które swoje siedziby mają w małych miasteczkach i wsiach. Inaczej sprawa wygląda w odniesieniu do małych firm położonych w większych miastach oraz średnich, te z kolei nie wyobrażają sobie istnienia i funkcjonowania bez możliwości korzystania z systemów informatycznych na tak konkurencyjnych rynkach.

Obecnie mamy do dyspozycji wiele różnych produktów informatycznych, które mają wspomagać działalność przedsiębiorstw. Dzisiejszy postęp technologii informatycznej ma swoje znaczące odbicie zarówno w produkcji sprzętu komputerowego, jak i tworzeniu oprogramowania do wspomagania firm. Wielorakość produktów technologii informatycznych podobnie, jak zróżnicowanie firm, wymaga zastosowania innych narzędzi i innych metod w odpowiednim zarządzaniu przedsiębiorstwem. Producenci prześcigają się w nowych narzędziach i ich przystępnością, zarówno pod względem ceny, jak i koniecznych umiejętności. Firmy muszą dokładnie analizować oferowane systemy, urządzenia oraz narzędzia wspomagające ich działalność, aby je odpowiednio i z oczekiwanym przez nie pozytywnym efektem wykorzystać.

LITERATURA

1. Adamczewski P.: Zintegrowane systemy informatyczne w praktyce. Wydawnictwo Mikom, Warszawa, 2003.
2. Biniasz D.: Wiedza o klientach firmy i jej wspomaganie w e-biznesie. [w:] Podstawy informatyczne w organizacji produkcji, pod red. Taranenko W. , Lubelskie Towarzystwo Naukowe, Lublin 2009, ss. 37-44.

3. Drozdowski E., Szpunar Z.: Dobór technologii w procesie realizacji bazodanowych systemów informatycznych. [w:] Rozwój zastosowań informatyki. Redakcja naukowa Kisielnicki J., Grabara J.K., Miłosz M., PTI, Katowice, 2006, ss. 179-206.
4. Kolbusz E., Olejniczak W., Szyjewski Z.: Inżynieria systemów informatycznych w e-gospodarce. Polskie Wydawnictwo Ekonomiczne, Warszawa 2005.
5. Pańkowska M.: Zarządzanie zasobami informatycznymi. Wydawnictwo Difin, Warszawa, 2001

Rozdział 19

Równoległe środowisko obliczeniowe jako narzędzie do tworzenia efektywnego portfela giełdowego

Agnieszka Ulfik
Politechnika Częstochowska
aulfik@gmail.com

Streszczenie

Analiza portfelowa to jedna z trzech podstawowych technik dostępnych dla inwestorów giełdowych. Powstała ona dzięki H. Markowitzowi, który swoją pierwszą pracę na ten temat opublikował w 1952 roku. Po przeszło pół wieku, stosowanie jego teorii wciąż jest sprawą skomplikowaną ze względu na dużą złożoność obliczeniową. Jedynym możliwym rozwiązaniem tego problemu jest zastosowanie systemów informatycznych, zdolnych do wyboru efektywnego portfela papierów wartościowych pomimo olbrzymiej ilości dopuszczalnych rozwiązań. Propozycją rozwiązania tego problemu są równoległe środowiska obliczeniowe – automaty komórkowe. Praca przedstawia modele funkcjonowania automatu komórkowego jako równoległego środowiska zdolnego do wyboru efektywnego portfela papierów wartościowych.

1. Portfel papierów wartościowych - Model Markowitza

Analiza portfelowa to jedna z trzech podstawowych technik, obok analizy technicznej i analizy fundamentalnej, stosowanych przez inwestorów giełdowych w celu optymalizacji swoich inwestycji. Główną zaletą analizy portfelowej jest dywersyfikacja kapitału powodująca zmniejszenie potencjalnego poziomu ryzyka.

W analizie portfelowej stworzonej przez H. Markowitza, na podstawie historycznych notowań spółek giełdowych, obliczana jest ich oczekiwana stopa zwrotu oraz odchylenie standardowe. Wielkości te interpretuje się jako spodziewany zysk z inwestycji oraz ryzyko jemu towarzyszące.

Na podstawie danych historycznych, wyznaczane są oczekiwana stopa zwrotu będącą zgodnie z tą teorią miarą poziomu przewidywanego zysku. Na podstawie stóp zwrotu wyznaczane są ich odchylenia standardowe będące miarą dyspersji i wyrażające ryzyko towarzyszące estymowanym zyskom.

Inwestorzy zwykle są zainteresowani walorami przynoszącymi duży dochód, związany z niskim poziomem ryzyka. Takie papiery wartościowe musiałyby posiadać wysoki poziom oczekiwanej stopy zwrotu przy jednoczesnym niskim stopniu dyspersji. Wartość stopy zwrotu w okresie t jest obliczana na podstawie następującego wzoru:

$$R_t = \frac{P_t - P_{t-1} + D_t}{P_{t-1}} \tag{1}$$

gdzie R_t – stopa zwrotu w okresie t,

P_t – cena waloru w okresie t,

P_{t-1} – cena waloru w okresie t-1,

D_t – dywidenda wypłacona w okresie t.

Dla każdego okresu t wyznaczana zostaje stopa zwrotu w związku z tym stopa zwrotu jest funkcją czasu. Wysokość zysku (lub straty) z inwestycji zależy od wielu czynników. W praktyce wartość oczekiwanej stopy zwrotu wyznacza się jako średnią arytmetyczną wszystkich zaobserwowanych stóp zwrotu. Prowadzi to do następującego wzoru na oczekiwaną stopę zwrotu z danego papieru wartościowego:

$$R = \frac{\sum_{t=1}^{N} R_t}{N} \tag{2}$$

gdzie: R – oczekiwana stopa zwrotu z danego papieru wartościowego,

R_t – empiryczna stopa zwrotu w okresie t,

N – liczba wszystkich analizowanych stóp zwrotu.

Tak określonemu poziomowi zysku zawsze towarzyszy ryzyko inwestycyjne. Pojęcie to jest niezwykle złożone. W praktycznych analizach giełdowych, ryzyko wyznacza się wykorzystując statystykę matematyczną. Wielkością interpretowaną jako ryzyko jest odchylenie standardowe będące pierwiastkiem kwadratowym z wariancji. Wyznacza się je na podstawie następującego wzoru:

$$S = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (R_i - R)^2} \tag{3}$$

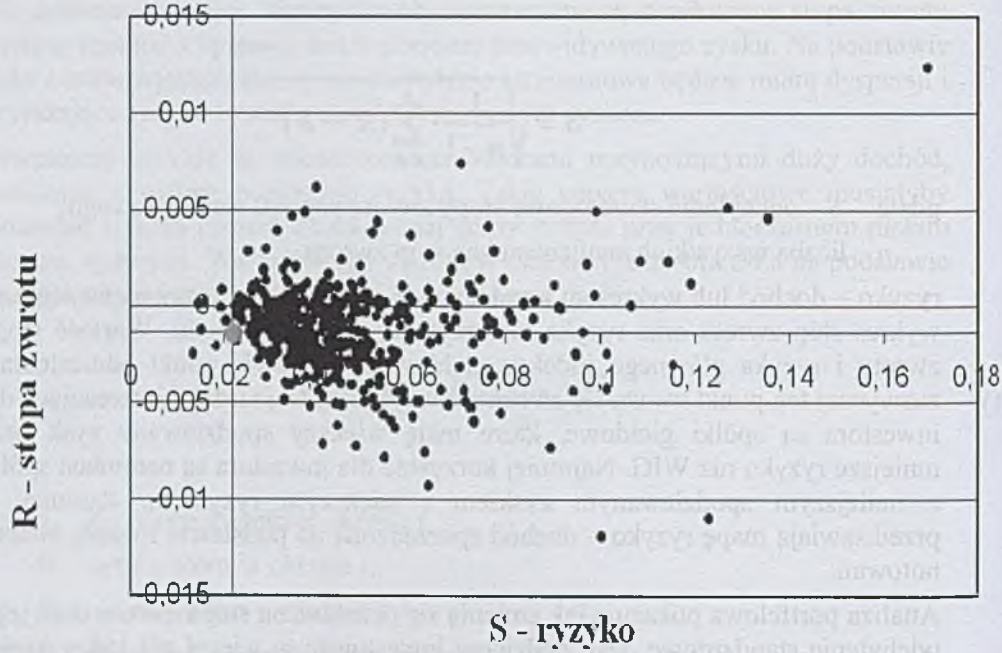
gdzie: S – odchylenie standardowe stopy zwrotu papieru wartościowego,

n – liczba wszystkich analizowanych stóp zwrotu.

ryzyko – dochód lub wykresem korelacyjnym. Powstaje on przez naniesienie na wykres stóp zwrotu oraz ryzyka towarzyszącego ich osiąganiu. Wartość stopy zwrotu i ryzyka głównego indeksu giełdowego stanowią punkt odniesienia i zazwyczaj ten punkt uważa się za centrum wykresu. Najbardziej interesujące dla inwestora są spółki giełdowe, które mają większy spodziewany zysk oraz mniejsze ryzyko niż WIG. Najmniej korzystne dla inwestora są natomiast spółki z mniejszym spodziewanym zyskiem i większym ryzykiem. Rysunek 1 przedstawiają mapę ryzyko – dochód sporządzoną na podstawie rocznej historii notowań.

Analiza portfelowa pokazuje jak zmienią się oczekiwana stopa zwrotu oraz jego odchylenie standardowe, jeśli będziemy inwestować w więcej niż jeden papier wartościowy, a także w jaki sposób dobrać do portfela inwestycyjnego jego składniki, aby zdywersyfikować ryzyko, czyli aby ryzyko portfela było mniejsze niż składników wchodzących w jego skład, przy jednoczesnym zachowaniu odpowiedniego poziomu zysku.

Dochód i ryzyko są głównymi kryteriami oceny rozpatrywanymi przez inwestora. Obie wielkości wyznaczane są dla wszystkich rozpatrywanych papierów wartościowych, a następnie umieszcza na wykresie potocznie nazywanym mapą.



Rys. 1. Mapa ryzyko – dochód dla większości spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie od 1-go czerwca 2008 roku do 1-go czerwca 2009 roku.

Tebele 1 i 2 przedstawiają procentową ilość spółek giełdowych w poszczególnych ćwiartkach mapy ryzyko dochód, przedstawionej na rysunku 1. Tabela 1 prezentuje te wyniki jeśli za centrum wykresu przyjmiemy indeks giełdowy WIG, a tabela 2 – jeśli indeks giełdowy WIG20.

Tab. 1. Procentowa ilość spółek giełdowych w poszczególnych ćwiartkach mapy ryzyko dochód, jeśli za centrum wykresu przyjmiemy indeks giełdowy WIG

IV – 1,6%	I – 52,8%
III – 0,5%	II – 45,1%

Tab. 2. Procentowa ilość spółek giełdowych w poszczególnych ćwiartkach mapy ryzyko dochód, jeśli za centrum wykresu przyjmiemy indeks giełdowy WIG20

IV – 5,4%	I – 46,6%
III – 1,6%	II – 46,4%

Spółki giełdowe najbardziej interesujące inwestorów, to spółki o niskim poziomie ryzyka oraz wysokim poziomie oczekiwanego zysku. Spółki te na mapie ryzyko-dochód znajdują się w IV ćwiartce i niestety nie jest ich dużo, bo

zaledwie 1,6% jeśli za punkt odniesienia przyjmiemy indeks giełdowy WIG, lub 5,4% jeśli za punkt odniesienia przyjmiemy indeks giełdowy WIG20.

W przypadku n składnikowego portfela papierów wartościowych, wartość stopy zwrotu z portfela R_p jest sumą stóp zwrotu poszczególnych walorów pomnożonych przez ich udziały w całości inwestycji. Wartość stopy zwrotu portfela n składnikowego wyznacza się na podstawie następującego wzoru:

$$R_p = \sum_{i=1}^{n} x_i \cdot R_i \tag{4}$$

gdzie $\sum_{i=1}^{n} x_i = 1$, $0 \le x_i \le 1$, dla $i = 1, 2, 3, \ldots, n$.

Odchylenie standardowe oczekiwanej stopy zwrotu S_p dla n składnikowego portfela papierów wartościowych będące miarą ryzyka, jest pierwiastkiem z wariancji. Można je wyznaczyć z jednego z dwóch równoważnych wzorów:

$$S_p = \sqrt{\sum_{i=1}^{n} x_i^2 \cdot S_i^2 + 2 \cdot \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_i \cdot x_j \cdot S_i \cdot S_j \cdot \rho_{ij}} \tag{5}$$

$$S_p = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} x_i \cdot x_j \cdot S_i \cdot S_j \cdot \rho_{ij}} \tag{6}$$

przy czym ρ_{ij} to współczynnik korelacji, wyznaczany na podstawie następującej formuły:

$$\rho_{ij} = \frac{\sum_{k=1}^{n} (R_{ik} - R_i) \cdot (R_{jk} - R_j)}{S_i \cdot S_j} \tag{7}$$

W przypadku n składnikowego portfela papierów wartościowych, wartość stopy zwrotu z portfela R_p oraz odchylenie standardowe oczekiwanej stopy zwrotu S_p będące miarą ryzyka, to dwa podstawowe parametry służące do porównywania różnych n składnikowych portfeli papierów wartościowych, używane w zarządzaniu portfelem papierów wartościowych.

2. Automaty komórkowe

Automaty komórkowe zostały stworzone w latach czterdziestych ubiegłego wieku, aby emulować procesy występujące w naturze, a jednocześnie tworzyć samoreplikujące się maszyny obliczeniowe. Za ich twórcę uważa się Johna von Neumanna. Swój początek miały one w naśladowaniu żywych organizmów . Wkrótce po ich odkryciu, okazały się one bardzo interesujące i przydatne w wielu dziedzinach nauki. Początkowo zainteresowali się nimi fizycy i zaczęli z powodzeniem stosować automaty komórkowe do symulacji złożonych zagadnień. Dziś automaty komórkowe są stosowane w matematyce, mechanice, ekonomii, grafice, socjologii, w symulacjach ruchu ulicznego i powietrznego, grach komputerowych, kryptografii i wielu innych zagadnieniach. Jednym z takich zagadnień może być symulowanie zachowania się rynku akcji i inwestorów, a także zagadnienie doboru portfela inwestycyjnego.

Automaty komórkowe to dyskretne modele używane głównie w fizyce, matematyce i teoriach obliczeniowych. Są to struktury takich samych elementarnych automatów komórkowych – nazywanych komórkami – ułożonych w siatkę. Zazwyczaj jest ona jedno, dwu lub trójwymiarową kratownicą, choć istnieją również inne sposoby ułożenia komórek przykładowo przypominające plaster miodu lub ułożonych jako sąsiadujące ze sobą trójkąty. Liczba wymiarów automatu komórkowego, może być większa niż trzy, choć w praktyce takie rozwiązania stosuje się rzadko ze względu na trudność w implementacji takiego modelu i niewielkie korzyści, w porównaniu z bardziej typowymi automatami komórkowymi .

W automacie komórkowym każda komórka komunikuje się z przylegającymi do niej sąsiadami. Wyróżnia się dwa podstawowe rodzaje sąsiedztw: von Neumanna i Moore'a. W pierwszym przypadku jako sąsiad rozumiana jest komórka przylegająca do danej całym bokiem, a w drugim również wierzchołkiem.

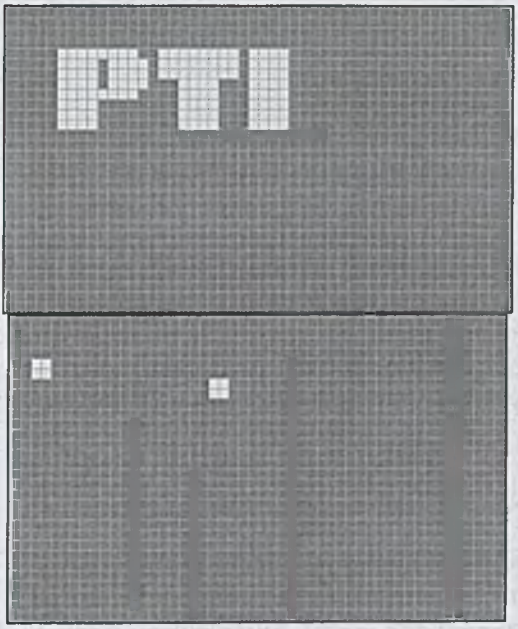
Każdy automat komórkowy oprócz swojej struktury oraz typu sąsiedztwa, musi mieć ustalone trzy następujące parametry:

- typ komórek budujących automat komórkowy, co oznacza ustalenie jakiego typu informacje muszą przechowywać elementarne automaty komórkowe,
- wartość początkową każdej komórki,
- funkcję przejścia, która jest algorytmem decydującym jaki będzie stan komórki w obecnej iteracji na podstawie wartości komórek sąsiednich w poprzedniej iteracji.

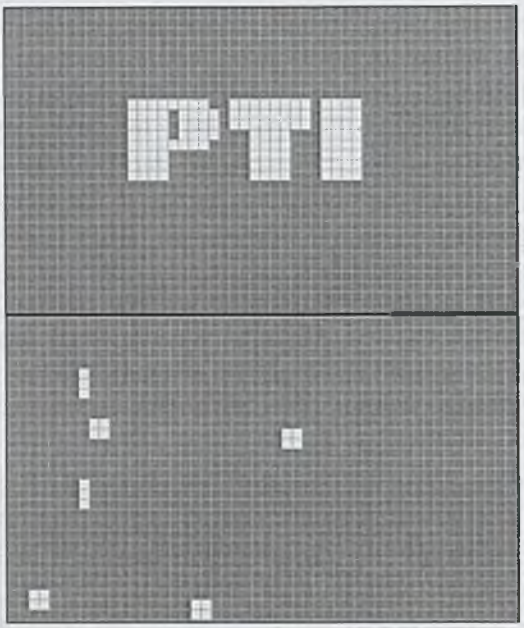
Najlepiej znanym przykładem zastosowania automatu komórkowego jest gra „Życie" stworzona przez Johna Conway'a. W grze tej automat komórkowy ma strukturę dwuwymiarowej siatki. Każda komórka otrzymuje stan początkowy: może być aktywna – żywa lub nieaktywna – martwa. W grze tej są ustanowione

reguły zachowania się każdej komórki, które mogą doprowadzić do ożywienia danej komórki lub też do jej obumarcia, w zależności od tego jakie wartości mają komórki sąsiednie. Ta gra symuluje środowisko naturalne w którym zwierzęta mogą się rodzić i umierać, kiedy nie mają wystarczającej ilości pożywienia.

Rysunki 2 i 3 przedstawiają przykładowe symulacje w grze Życie, gdzie stanem początkowym są żywe komórki ułożone w znak logo PTI. Na obydwu rysunkach, lewa część to stan początkowy a prawa to stan końcowy. Rysunki te różnią się między sobą miejscem ulokowania żywych komórek czyli stanem początkowym. Oba te stany początkowe prowadzą do różnych efektów. Na rysunku 2, komórki doszły do stabilnej konfiguracji, w której nic się nie zmieniało, a na rysunku 3, żywe komórki ciągle się przemieszczały w pewnej stałej sekwencji.



Rys. 2. Przykładowa symulacja gry „Życie"



Rys.3. Przykładowa symulacja gry „Życie"

Ponieważ automat komórkowy charakteryzuje się dużą mocą obliczeniową, celowe jest sprawdzenie, w jaki sposób środowisko automatów komórkowych może pomóc w doborze parametrów portfela inwestycyjnego. Takie próby były już znane i dowiodły, iż proces doboru portfela papierów wartościowych może się odbywać z użyciem środowiska automatów komórkowych.

3. Symulacje

W celu zbadania przydatności automatów komórkowych do konstruowania portfela papierów wartościowych, napisany został autorski program komputerowy. Program ten został zaimplementowany w środowisku Builder C++ firmy Borland. Aplikacja ta wczytuje dane giełdowe z wybranych przez użytkownika plików w formacie tekstowym. Z danych tych wybierane są notowania dotyczące wybranego przez użytkownika przedziału czasowego. Na podstawie tych notowań, wyznaczane są podstawowe charakterystyki wybranych papierów wartościowych (oczekiwana stopa zwrotu oraz jej odchylenie standardowe). Jako kolejne wielkości, wyznaczane są współczynniki korelacji pomiędzy wszystkimi wybranymi spółkami.

Wszystkie wyznaczone wielkości trafiają jako informacje wejściowe do środowiska automatów komórkowych. Na podstawie tych danych dokonywana jest symulacja. Celem przeprowadzania symulacji było znalezienie portfela charakteryzującego się możliwie największym zyskiem przy najmniejszym

ryzyku. Przykładowe symulacje przeprowadzone w środowisku automatów komórkowych miały na celu stworzenie portfeli dwuskładnikowych. Badany przedział czasowy dotyczył danych od 1-go czerwca 2008 roku do 1-go czerwca 2009 roku.

Automat komórkowy użyty w symulacjach miał 100 komórek w pionie i 100 w poziomie, czyli składał się z 10 000 komórek – elementarnych automatów komórkowych. Na początku każdej symulacji, każda komórka miała losowo wybierany skład portfela inwestycyjnego. Następnie komórki komunikując się ze sobą, wybierały portfel papierów wartościowych najbardziej interesujący inwestora przy zadanych warunkach np. maksymalizacji zysku czy minimalizacji ryzyka. Tabele 3-6 przedstawiają wyniki uzyskane w 10 symulacjach, a dodatkowo w tabeli 3 umieszczono charakterystyki indeksu giełdowego WIG.

Tab. 3. Charakterystyki dla WIGu oraz dwuelementowych portfeli papierów wartościowych

Numer realizacji	R	S
1	0,007536	0,057708
2	0,003120	0,030226
3	0,006291	0,078914
4	0,006223	0,050309
5	0,005748	0,044643
6	0,006123	0,056502
7	0,005911	0,062542
8	0,010551	0,056892
9	0,005317	0,043450
10	0,004811	0,047673
WIG	0,001188	0,013148

Tab. 4. Charakterystyki dla trójelementowych portfeli papierów wartościowych

Numer realizacji	R	S
1	0,006387	0,047457
2	0,005298	0,030660
3	0,004031	0,037111
4	0,004448	0,046314
5	0,004661	0,050315
6	0,004151	0,021038
7	0,004521	0,027484
8	0,006127	0,040485
9	0,007111	0,036084
10	0,006032	0,034586

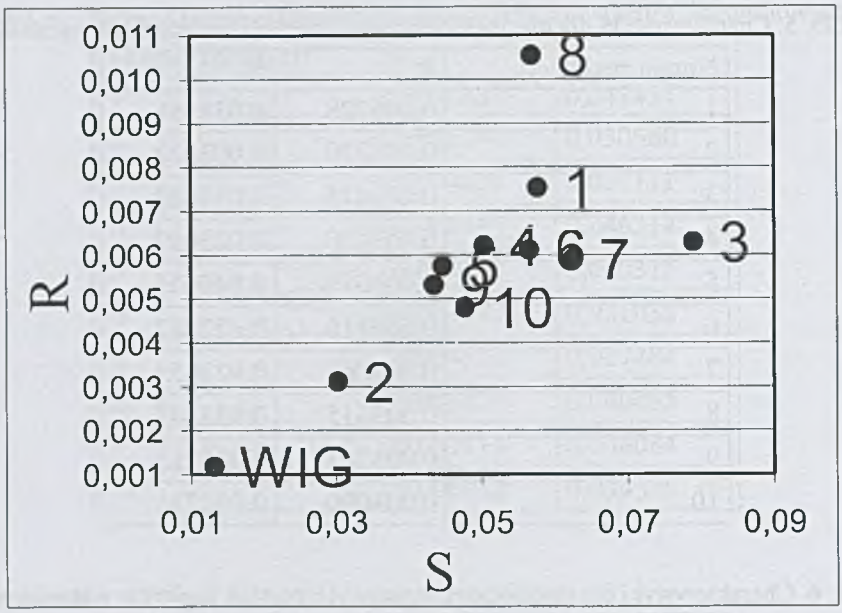
Tab. 5. Charakterystyki dla pięcioelementowych portfeli papierów wartościowych

Numer realizacji	R	S
1	0,005728	0,038461
2	0,003270	0,003270
3	0,003415	0,031080
4	0,003420	0,025089
5	0,006629	0,040605
6	0,005116	0,027682
7	0,004063	0,021059
8	0,003315	0,033345
9	0,003703	0,026313
10	0,004090	0,038234

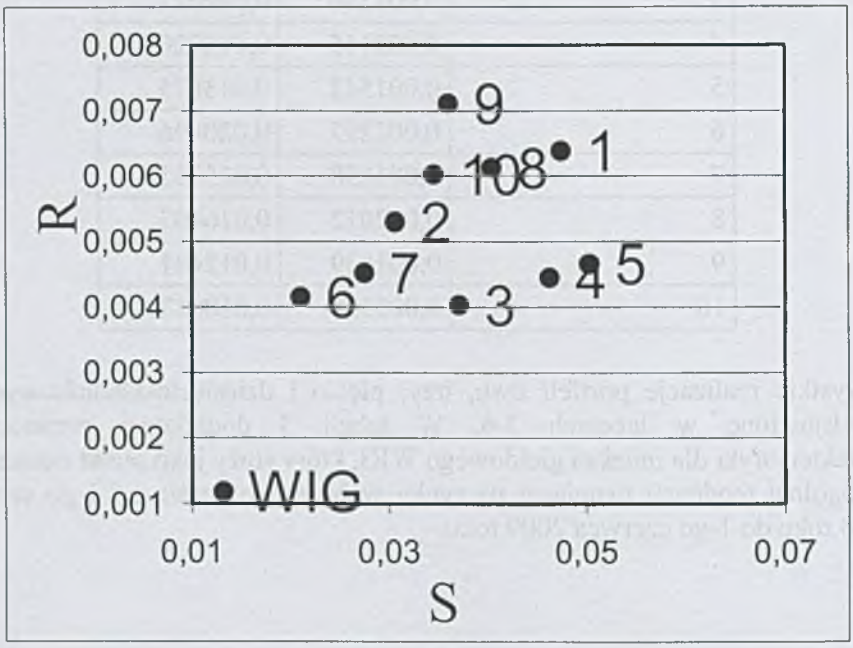
Tab. 6. Charakterystyki dla dziesięcioelementowych portfeli papierów wartościowych

Numer realizacji	R	S
1	0,001480	0,020040
2	0,003299	0,020797
3	0,001390	0,010627
4	0,002145	0,014026
5	0,001542	0,013875
6	0,002297	0,020496
7	0,001898	0,015732
8	0,002012	0,016467
9	0,001639	0,012481
10	0,002336	0,018017

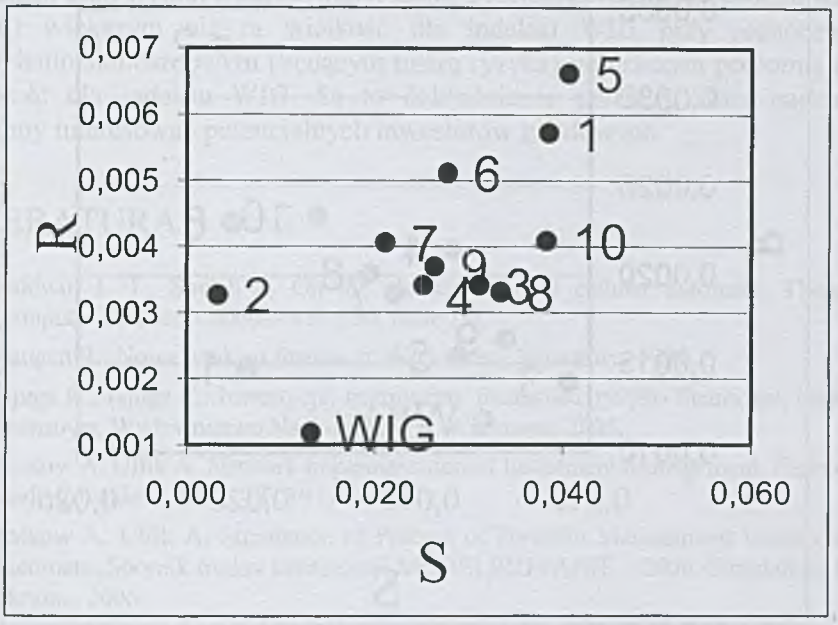
Wszystkie realizacje portfeli dwu, trzy, pięcio i dziesięcioskładnikowych są przedstawione w tabelach 3-6. W tabeli 3 dodatkowo umieszczono charakterystyki dla indeksu giełdowego WIG, który służy jako punkt odniesienia do ogólnej tendencji panującej na rynku w badanym czasie od 1-go czerwca 2008 roku do 1-go czerwca 2009 roku.



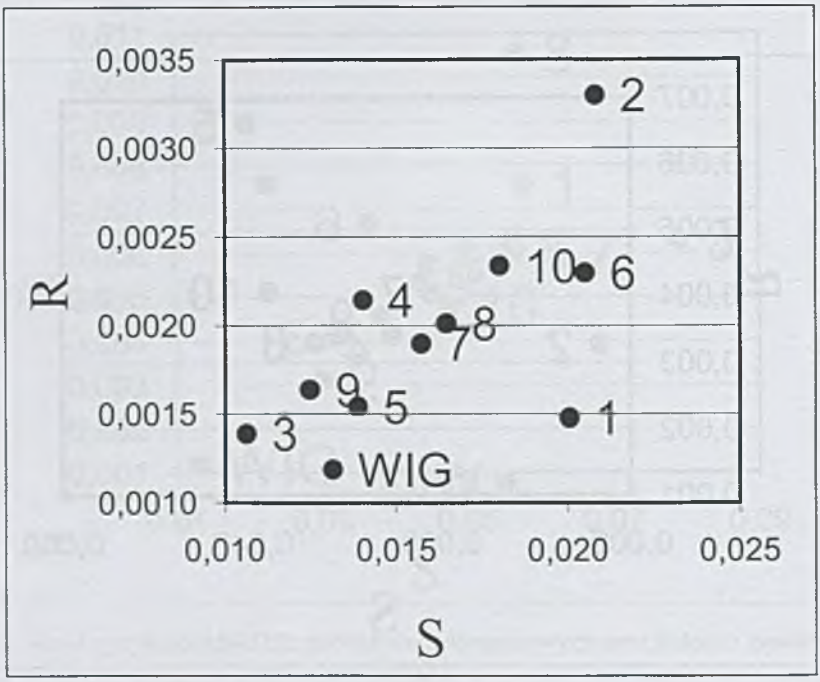
Rys. 4. Mapa ryzyko-dochód dla portfeli dwuelementowych oraz indeksu giełdowego WIG



Rys. 5. Mapa ryzyko-dochód dla portfeli trójelementowych oraz indeksu giełdowego WIG



Rys. 6. Mapa ryzyko-dochód dla portfeli pięcioelementowych oraz indeksu giełdowego WIG



Rys. 7. Mapa ryzyko-dochód dla portfeli dziesięcioelementowych oraz indeksu giełdowego WIG

Rysunki 4-7 przedstawiają portfele papierów wartościowych, odpowiednio dwu, trzy, pięcio i dziesięcioelementowe, przedstawione na mapach ryzyko-dochód. Na wszystkich tych rysunkach został również umieszczony indeks giełdowy WIG w celu porównania wylosowanych portfeli do ogólnej sytuacji na rynku giełdowym.

4. Podsumowanie

Przeprowadzone symulacje, potwierdziły przypuszczenia, że automat komórkowy potrafi tworzyć efektywne portfele papierów wartościowych, bazujące na klasycznym portfelu papierów wartościowych. Elementarne automaty komórkowe współpracując ze sobą, wymieniają się informacjami o wybranych portfelach papierów wartościowych i w ten sposób cały automat wybiera portfel papierów wartościowych o charakterystykach poszukiwanych przez inwestora, w zależności od przyjętej przez niego strategii inwestycyjnej.

Wieloelementowe portfele papierów wartościowych widoczne na mapach ryzyko-dochód, na rysunkach 4-7, w graficzny sposób prezentują wyniki przeprowadzonych symulacji w środowisku automatów komórkowych. Prawidłowością jest iż w miarę zwiększania ilości składników w portfelu, jego charakterystyki się stopniowo poprawiają, co jest zgodne z klasyczną teorią

portfelową Markowitza. W przypadku portfeli dziesięcioelementowych, część z symulacji dało wyniki o oczekiwanej stopie zwrotu (czyli mierze spodziewanego zysku) większym niż ta wielkość dla indeksu WIG przy jednoczesnym odchyleniu standardowym (będącym miarą ryzyka) na niższym poziomie niż ta wielkość dla indeksu WIG. Są to dokładnie te portfele, które najbardziej powinny interesować potencjalnych inwestorów giełdowych.

LITERATURA

1. Baldwin J. T., Shelah S. On the classifiability of cellular automata, Theoretical Computer Science.– 2000.– Vol. 230, Issue 1-2.
2. Haugen R.: Nowa nauka o finansach. WIG–Press, Warszawa, 1999.
3. Jajuga K., Jajuga T.: Inwestycje, instrumenty finansowe, ryzyko finansowe, inżynieria finansowa. Wydawnictwo Naukowe PWN, Warszawa, 2005.
4. Katkow A. Ulfik A. Network Implementation of Investment Management, Elektronnoe modelirovanie. – 2007. – T. 29 No 4.
5. Katkow A. Ulfik A. Simulation of Process of Portfolio Management Using Cellular Automata. Sbornik trudov konferencii MODELIROVANIE – 2006, Simulation, Kijev, Ukraine, 2006.
6. Kari Jarkko Theory of cellular automata: A survey// Theoretical Computer Science.– 2005.– Vol. 334, Issue 1– 3.
7. Markowitz H.: Portfolio selection, The Journal of Finance, 1952, Vol.7 No1.
8. Sharpe W. A Simplified Model For Portfolio Analysis, Management Science.– 1963.– Vol. 9, Issue2.
9. Tarczyński W.: Fundamentalny portfel papierów wartościowych. Polskie Wydawnictwo Ekonomiczne, Warszawa, 2002.
10. Ulfik A. The Simulation of Parallel Chaotic Processes in Sequential Environments, Elektronnoe modelirovanie. – 2005. – T. 27 No 2

Rozdział 20

Wykorzystanie narzędzi internetowych w systemie komunikacji z telewidzami

Roman Kmieciak
Politechnika Śląska
roman.kmieciak@polsl.pl

Streszczenie

Niniejszy rozdział podejmuje zagadnienie wykorzystania Internetu w procesie komunikacji z widzami w przedsiębiorstwie telewizyjnym. W analizie studium przypadku omówiono niektóre rozwiązania zastosowane w systemie informacyjnym brytyjskiego nadawcy publicznego. W badaniach ilościowych audytorium porównano aktywność użytkowników serwisów internetowych europejskich nadawców telewizyjnych. Na tej podstawie podjęto próbę określenie efektywności działań podejmowanych w ramach systemu zarządzania relacjami z telewidzami.

1. Wprowadzenie

Stacje telewizyjne rozpoczęły swą działalność w roku 1928. Początkowo przekaz telewizyjny realizowano za pomocą anten naziemnych. Ewolucja techniczna pozwoliła na rozpowszechnianie programu telewizyjnego drogą satelitarną i kablową [2]. Wraz z rozwojem Internetu dla przedsiębiorstw telewizyjnych pojawił się nowy kanał dystrybucji i nowy sposób dotarcia do telewidza.

W przedsiębiorstwie, aby właściwie realizować funkcje zarządzania, konieczne jest gromadzenie i przetwarzanie informacji o swoich klientach. Zasada ta dotyczy także, działających w warunkach rynkowej konkurencji, przedsiębiorstw telewizyjnych, dla których marketingowa strategia orientacji na klienta prowadzi do ciągłego poznawania, analizowania i zaspokajania potrzeb telewidzów. Dynamiczne i burzliwe otoczenie, nowe możliwości techniczne, wzrost konkurencyjności na rynku telewizyjnym oraz coraz bardziej

wymagające audytorium – wszystko to wymaga od nadawców telewizyjnych przedefiniowania sposobów budowania i utrzymywania kontaktów ze swoimi odbiorcami. Widzowie są na tyle ważni dla stacji telewizyjnych, iż można nawet stwierdzić, że „intensywność więzi z odbiorcami jest istotnym elementem wyróżniającym media wobec innych rodzajów działalności" [5].

Reakcje i sygnały od telewidzów są źródłem istotnych informacji, które mogą wpływać nie tylko na tematykę podejmowaną w audycjach telewizyjnych, ale także na decyzje podejmowane przez menedżerów w przedsiębiorstwie telewizyjnym.

Szczególnym i aktualnym wyzwaniem dla tradycyjnych nadawców telewizyjnych jest wykorzystanie Internetu. Globalna sieć komputerowa stwarza możliwość nie tylko alternatywnego sposobu rozpowszechniania produktów telewizyjnych i uzyskiwania dodatkowych dochodów [8, 11], ale także pozwala budować głębsze relacje i dłuższe interakcje z widzami [1, 9, 10].

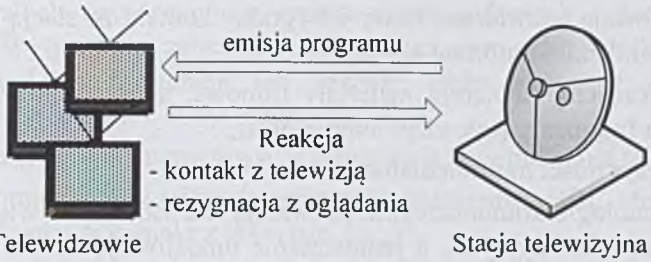
Niniejszy rozdział podejmuje zagadnienie wykorzystania stron internetowych jako części systemu komunikacji pomiędzy stacją telewizyjną a telewidzami. Pierwsza część pracy stanowi omówienie znaczenia systemów CRM w przedsiębiorstwie telewizyjnym. Część druga rozdziału stanowi analizę studium przypadku brytyjskiego nadawcy publicznego. Ostatnia część przedstawia wyniki badań użytkowników internetowych serwisów stacji telewizyjnych.

2. CRM, działalność telewizyjna i telewidzowie

Działalność marketingowa współczesnych przedsiębiorstw nie ogranicza się jedynie do reklamy i promocji, ale obejmuje ciągłe badanie potrzeb klientów, ich zaspokajanie i kreowanie nowych potrzeb. Aby analizować potrzeby i zachowania klientów, konieczne jest zarządzanie zgromadzoną informacją, która jest podstawą podejmowania decyzji w przedsiębiorstwie. Informacje i technologie informacyjne stały się kluczowymi czynnikami nowoczesnego marketingu [7], a efektywne zarządzanie jest możliwe tylko wówczas, kiedy mamy dostęp do informacji o organizacji i jej otoczeniu [3]. Sprawna komunikacja z klientami przyczynia się do rozbudowy bazy danych o klientach, umożliwię lepsze poznanie zachowań, potrzeb i preferencji klientów. Gromadzenie, przetwarzanie i analizowanie informacji od i o klientach pozwala na personalizację kontaktów, wzmacnia więzi miedzy przedsiębiorstwem a klientami oraz poprawia efektywność kampanii marketingowych.

Zasady nowoczesnego marketingu znajdują swoje odzwierciedlenie w zarządzaniu specyficznym przedsiębiorstwem, jakim jest przedsiębiorstwo telewizyjne.

W początkowych okresach działalności stacji telewizyjnych interakcja z telewidzami przybierała bardzo proste formy. Rysunek 1 przedstawia prosty schemat komunikacji między telewizją a telewidzem.



Rys. 1. Komunikacja między stacją telewizyjną a telewidzami

Stacja telewizyjna nadaje program telewizyjny, a telewidzowie kontaktują się z nią poprzez listy, telefon lub osobisty kontakt (wizyta w ośrodku telewizyjnym). Treścią kontaktu jest najczęściej:

- reakcja telewidzów na wyemitowaną audycję telewizyjną (pytania, komentarze, krytyka, pozytywne i negatywne opinie),
- propozycje podjęcia określonych tematów w audycjach lub prośba o interwencję (źródło tematów dla audycji np. informacyjnych, reportaży).

Upraszczając można stwierdzić, że głównym celem stacji telewizyjnych jest dążenie do jak największej oglądalności programu telewizyjnego, która przekłada się na wpływy reklamowe. Największą siłą telewidzów jest natomiast możliwość rezygnacji z oglądania telewizji lub przełączenia na innych kanał telewizyjny.

Zmiany zachodzące przez ostatnie lata bardzo znacząco wpłynęły na ekonomikę i zarządzanie w stacji telewizyjnej. Zmianie uległa struktura rynkowa telewizji, która z monopolu, poprzez oligopol zbliża się do konkurencji monopolistycznej [4]. Wynika to miedzy innymi z rozwoju Internetu, zwłaszcza szerokopasmowy, co skutkuje zmianami dokonującymi się w technologii dystrybucji. Bariery technologiczne (np. brak częstotliwości) odgrywają coraz mniejszą rolę. Nasila się konkurencja na polu jakości, różnorodności i atrakcyjności programów.

Nie bez znaczenia jest umiejętność budowania i utrzymywania przez stację telewizyjną poprawnych relacji z telewidzami. Siła i jakość tych relacji może być ważnym czynnikiem wpływającym na intensywność kontaktów telewidzów z telewizją, rozumianych tutaj także jako czas spędzony na oglądaniu audycji telewizyjnych.

Według Kotlera [6] jedną z podstawowych zasad, jakimi powinni się kierować współczesne przedsiębiorstwa, aby zapewnić sobie sukces jest opracowanie jasnej koncepcji do czego i w jaki sposób można wykorzystać Internet. Zasada ta dotyczy także szczególnego rodzaju przedsiębiorstw – stacji telewizyjnych.

Nadawcy telewizyjni współpracują z portalami internetowymi lub tworzą własne rozbudowane serwisy internetowe. Serwisy te:

- mają na celu promocję stacji i audycji telewizyjnych,

- umożliwiają telewidzom łatwy i wygodny kontakt ze stacją telewizyjną, jej pracownikami i autorami audycji,
- coraz częściej zawierają materiały filmowe, audycje telewizji tradycyjnej lub ich fragmenty (telewizja internetowa),
- zawierają treści multimedialne i interaktywne.

Nowe technologie komunikacyjne pozwalają telewidzom na większą interakcję z audycjami telewizyjnymi, a jednocześnie umożliwiają stacjom telewizyjnym uzyskanie dodatkowych dochodów z nowych źródeł. Takimi źródłami są na przykład:

- połączenia telefoniczne i wiadomości SMS wysyłane przez telewidzów jako reakcja na konkursy, gry i głosowania,
- opłaty za dostęp do materiałów filmowych umieszczonych na stronach internetowych lub w usłudze *video-on-demand* (*pay-per-view*).

Działania stacji telewizyjnych względem telewidzów składają się na strategię zarządzania relacjami z klientami (CRM), z tą różnicą, że klientami są telewidzowie. Wyzwanie dla stacji telewizyjnych jest takie zbudowanie relacji z telewidzami, aby po zakończeniu emisji w tradycyjnej telewizji, kontakt z audycją był kontynuowany. Możliwości takie stwarza serwis internetowy zawierający interaktywne treści powiązane tematycznie z wyemitowana audycją [9]. Celem jest utrzymanie widza, nagradzanie lojalnych widzów i wzrost zainteresowania audycją poprzez wykorzystanie działań promocyjnych.

Dla stacji telewizyjnych kontakt z widzami drogą poczty elektronicznej i SMS jest okazją do zbudowania i rozwijania bazy danych widzów (adresów e-mail i numerów telefonów), przyczynia się do lepszego poznania widzów i w konsekwencji efektywniejszego dotarcia do grup docelowych zarówno przez autorów audycji jak i reklamodawców.

3. Realizacja komunikacji z telewidzami na przykładzie strony internetowej BBC

BBC (*British Broadcasting Corporation*) jest brytyjskim publicznym nadawcą radiowo-telewizyjnym. Początki działalności sięgają roku 1922. Obecnie w skład korporacji wchodzi m.in. 10 sieci radiowych i 8 kanałów telewizyjnych. Krajowa działalność BBC jest finansowana przede wszystkim z abonamentu telewizyjnego (*licence fee*), który w 2009 roku wynosi 142,50 funtów rocznie dla każdego gospodarstwa domowego. Portal BBC [12] został wybrany do analizy pod względem stosowanego systemu komunikacji z telewidzami z uwagi na:

- renomę telewizji – BBC jest największym nadawcą publicznym na świecie i uznawanym powszechnie za jednego z najlepszych. Nadawca znany jest ze swojego zobowiązania do zapewnienia wysokich standardów

dziennikarskich i etycznych w emitowanych audycjach i świadczonych usługach. Wytyczne są zebrane w 225 stronicowym przewodniku *Editorial Guidelines* [13]. Dokument ten zawiera także wytyczne w zakresie kontaktów i interakcji z audytorium.

- rozbudowany serwis internetowy, w tym serwis komunikacji z telewidzami.

W obszarze komunikacji z telewidzami serwis internetowy BBC oferuje szereg narzędzi ułatwiających kontakt z telewizją.

Ponieważ każdego dnia do stacji BBC napływają znaczne ilości listów elektronicznych, zachęca się widzów, którzy chcieliby zadać pytanie, aby w pierwszej kolejności zapoznali się z działem FAQ (*Frequently Asked Questions*). Witryna FAQ zawiera odpowiedzi na pytania, które już wcześniej były kierowane do BBC przez telewidzów, radiosłuchaczy i gości portalu. Strona pozwala w łatwy sposób przeszukiwać zadane pytania wraz z możliwością ograniczenia wyników przeszukiwań do pewnych obszarów, np. informacji o programie telewizyjnym. Zapytania użytkowników są na bieżąco analizowane i na ich podstawie są dodawane nowe lub uzupełniane istniejące odpowiedzi.

Zgłaszanie skarg (ang. *complaints*) przez telewidzów odbywa się poprzez:

- kontakt telefoniczny,
- tradycyjną pocztę kierowana do działu *BBC Complaints* lub bezpośrednio do osób odpowiedzialnych za konkretna audycje lub przedsięwzięcie,
- stronę internetową.

Na szczególną uwagę zasługuje proces przesyłania skarg z wykorzystaniem narzędzi na stronie internetowej (rys. 2). Procedura formułowania skargi obejmuje pięć kroków:

W pierwszym etapie należy wybrać rodzaj reakcji – informacji zwrotnej, jaka telewidz planuje przesłać. Wybór dokonany w tym kroku determinuje kroki kolejne. Do wyboru jest kilka opcji:

- zadanie pytania,
- przesłanie komentarza (własnego punktu widzenia),
- przesłanie słów uznania (pochwały),
- zgłoszenie propozycji (przesłanie sugestii),
- zgłoszenie usterek technicznych,
- złożenie skargi.

W przypadku składania skargi, w kroku drugim należy wybrać jedna z opcji, która najlepiej opisuje usługę, której dotyczy skarga. Wyboru można dokonać spośród: audycji telewizyjnych, audycji radiowych, stron internetowych, telewizji interaktywnej (*BBC Red Button*), teletekstu (*Ceefax*) lub ogólnie BBC.

W trzecim kroku należy wypełnić formularz, którego postać zależy od wybranej w drugim kroku opcji. Jeśli skarga dotyczy programu telewizyjnego,

w formularzu należy: wskazać kanał telewizyjny, nazwę audycji i datę emisji, wpisać treść skargi i wskazać, czy osoba składająca skargę życzy sobie otrzymać odpowiedź na tę skargę.

Jeżeli oczekujemy odpowiedzi ze strony nadawcy, należy podać swoje dane, m. in.: imię, nazwisko, adres e-mail, kraj. Jeśli nie oczekujemy odpowiedzi, dane te są opcjonalne, ale ich podanie pomaga w dokładnym zarejestrowaniu i przekazaniu skargi do kierownictwa. W obu przypadkach należy wskazać, czy osoba składająca skargę ma poniżej 13 lat. Jeśli tak, wymagana jest zgoda rodziców (opiekunów).

Ostatni etap stanowi podsumowanie skargi i umożliwia jej zatwierdzenie (przesłanie).

Rys. 2. Schemat procesu przesyłania skarg telewidzów

Skargi powinny być zgłaszane w ciągu 30 dni od emisji audycji lub wydarzenia. Odpowiedź następuje zwykle w ciągu 10 dni, ale jest to uzależnione od charakteru skargi. Na stronie internetowej zamieszczone są oficjalne odpowiedzi BBC w sprawach, które poruszyły szczególnie widzów lub były przedmiotem wielu skarg. Odpowiedzi te zwykle są dostępne przez 6 miesięcy.

Na swoich stronach internetowych BBC publikuje także:

- comiesięczne podsumowanie głównych problemów pojawiających się w skargach. W czerwcu 2009 roku BBC odpowiedziała na ponad 15000 skarg,
- kwartalne raporty o skargach skierowanych do *Editorial Complaints Unit* – zespołu zajmującego się istotnymi skargami dotyczącymi naruszenia standardów BBC,
- comiesięczne raporty o rozpatrzonych skargach przez *Editorial Standards Committee* – komisję odwoławczą przy *BBC Trust* (rodzaj komitetu nadzorującego BBC).

Interakcje między telewizją a telewidzami przybierają także formę komentarzy. Chcą wyrazić swoja opinię o audycji, jej treści lub poruszanym temacie, użytkownik serwisu ma do dyspozycji:

- *BBC Message Boards* i *Points of View* – internetowe fora dyskusyjne,
- interaktywny formularz, który przesyłane wiadomości pozwala podzielić na: komentarze, zapytanie lub słowa uznania. Formularz wymaga podania szczegółów audycji, której dotyczy komentarz oraz danych kontaktowych użytkownika. Przesyłanie tych danych odbywa się z wykorzystaniem protokołu szyfrowania TLS.

Przedstawione powyżej rozwiązania w obszarze kontaktów z telewidzami są jedynie częścią rozbudowanego systemu.

4. Efektywność internetowego serwisu nadawcy telewizyjnego

Efektywność systemów CRM najczęściej mierzy się za pomocą: wzrostu poziomu satysfakcji klienta, liczby nowych transakcji, wskaźników retencji lub stopnia redukcji kosztów. Ze względu na specyfikę zarządzania relacjami z telewidzami, istotnym miernikiem efektywności systemu wydaje się być poziom satysfakcji. Interesującym zagadnieniem jest sposób pomiaru satysfakcji telewidzów korzystających z serwisu internetowego. Satysfakcję tę można mierzyć z wykorzystaniem badań jakościowych (np. badanie opinii) i/lub ilościowych. Obiektywnym, ilościowym wskaźnikiem satysfakcji użytkowników serwisu jest średni czas spędzony w serwisie. Zgodnie z tym podejściem, im wyższy stopień zadowolenie (czerpanej satysfakcji) z serwisu internetowego, tym dłuższy czas spędzony w serwisie.

Tab. 1. Wyniki pomiaru aktywności użytkowników serwisów internetowych publicznych nadawców telewizyjnych[1]

L.p.	Nadawca	Kraj	Strona www	Średni czas spędzony w serwisie [min]	Średnia liczba stron / użytkownika	Ranking w kraju
1.	ORT	Austria	www.ort.au	7	5,69	5
2.	BBC	W. Brytania	www.bbc.co.uk	6,7	4,85	7
3.	DR	Dania	www.dr.dk	5,4	4,94	10
4.	RAI	Włochy	www.rai.it	5,4	3,57	77
5.	YLE	Finalandia	http://yle.fi	5,4	5,84	16
6.	TVE	Hiszpania	www.rtve.es	4,9	3,68	78
7.	CT	Czechy	ceskatelevize.cz	4,8	5,56	67
8.	NRK	Norwegia	www.nrk.no/	4,8	3,5	19
9.	RTP	Portugalia	ww1.rtp.pt	4,6	3,25	61
10.	TVP	Polska	www.tvp.pl	4,2	4,15	136
11.	SVT	Szwecja	http://svt.se	4	4,25	25
12.	RTE	Irlandia	www.rte.ie	3,9	3,83	13
13.	ERT	Grecja	www.ert.gr	3,4	3,19	122
14.	TVR	Rumunia	www.tvr.ro	3,4	3,88	649
15.	ZDF	Niemcy	www.zdf.de	3,3	3,36	131
16.	STV	Słowacja	www.stv.sk	3,2	4,13	171
17.	ARD	Niemcy	www.ard.de	3,1	2,95	142
18.	MTV	Węgry	www.hirado.hu	2,5	1,94	251
Średnia				**4,44**	**4,03**	**110**

W tabeli 1 przedstawiono średnie czasy spędzone przez użytkowników w serwisach wybranych[2] największych europejskich publicznych stacji telewizyjnych, w tym analizowanego serwisu BBC.

Ponadto wyniki uzupełniono o średnią liczbę stron serwisu, jaką użytkownik wyświetlił podczas jednego dnia oraz pozycję w rankingu stron internetowych w danym kraju, wyznaczoną na podstawie liczby użytkowników, którzy odwiedzili tę stronę oraz liczby stron wyświetlonych przez tych użytkowników. Średnie zostały obliczone na podstawie danych trzymiesięcznych (od maja do lipca 2009 r.).

Z tabeli 1 wynika, że oprócz austriackiej ORT, średnio najwięcej czasu użytkownicy spędzają na stronach brytyjskiego nadawcy BBC. Poza tym serwis

[1] opracowanie własne z wykorzystaniem [14]

[2] Nie uwzględniono np. publicznych telewizji, których stacje telewizyjne posiadają oddzielny serwis internetowy (z różnymi domenami), np. telewizja francuska.

BBC cieszy się dużą popularnością mierzoną pozycją w krajowym rankingu stron internetowych. Można zatem zaryzykować stwierdzenie, że BBC bardzo dobrze w porównaniu z badanymi nadawcami telewizyjnymi wypełnia zadania w ramach zarządzania relacjami z telewidzami z wykorzystaniem narzędzi internetowych. Nie można tutaj jednak wyciągnąć wniosku o efektywności tych działań (rozumianej jako stosunek efektów do nakładów), nie znając nakładów poniesionych na utworzenie i utrzymanie poszczególnych serwisów.

5. Podsumowanie

W pracy skoncentrowano się na wskazaniu internetowych narzędzi interakcji z audytorium, jednak dla tworzenia kompleksowego systemu zarządzania relacjami nie są one wystarczające. Nie można bowiem zapominać o innych formach i narzędziach budowy i utrzymywania poprawnych stosunków z telewidzami, takimi jak np.: urządzenia mobilne, infolinie, telewizja interaktywna, nagrywanie audycji z udziałem publiczności, organizacja przedsięwzięć artystycznych, konkursy, głosowania, gry i inne techniki interaktywne.

Oprócz stosowania środków technicznych, konieczne jest opracowanie polityki komunikacji z telewidzami, która określi np. zasady gromadzenia i ochrony danych osobowych telewidzów, postępowanie w przypadku skarg, pytań i propozycji programowych. Istotnym zagadnieniem są badania ilościowe i jakościowe audytorium.

W przeprowadzonych badaniach case study przedstawiono niektóre rozwiązań wdrożone na łamach serwisu internetowego przez BBC. BBC jest nadawcą publicznym, co wpływa w pewnym stopniu na podejście do telewidza. Już wstępna analiza wykazała, że Internet nie jest postrzegany przez brytyjskiego nadawcę jako potencjalne źródło dochodów, ale przede wszystkim jako ważny kanał komunikacji z odbiorcami i dystrybucji produkcji radiowo-telewizyjnych. Z tego też powodu do pomiaru efektów zarządzania relacjami z telewidzami zaproponowano mierniki poziomu satysfakcji, a nie wskaźniki związane z transakcjami, dochodami i kosztami.

LITERATURA

1. Chan-Olmsted S. M., Ha L. S.: Internet Business Models for Broadcasters: How Television Stations Perceive and Integrate the Internet. Journal of Broadcasting & Electronic Media 2003, 47:4, s. 597-616.
2. Karwowska-Lamparska A.: Rozwój radiofonii i telewizji. Przegląd telekomunikacyjny 2003, nr 1, s. 31-37.
3. Kisielnicki J., Sroka H.: Systemy informacyjne biznesu: informatyka dla zarządzania. Placet, Warszawa 2005.

4. Kowalski T. Jung B.: Media na rynku. Wprowadzenie do ekonomiki mediów. Wydawnictwa Akademickie i Profesjonalne, Warszawa 2006.
5. Kowalski T.: Między twórczością a biznesem. Wprowadzenie do zarządzania w mediach i rozrywce. Wydawnictwa Akademickie i Profesjonalne, Warszawa 2008.
6. Kotler Ph.: Kotler o marketingu. Jak kreować i opanować rynek. Profesjonalna Szkoła Biznesu, Kraków 1999.
7. Nowicki A. (red.): System informacyjny marketingu przedsiębiorstw: modelowanie. Polskie Wydawnictwo Ekonomiczne, Warszawa 2005.
8. Picard R.: Changing Business Models of Online Providers. International Journal on Media Management 2000, t. 2, nr 2, s. 60-68.
9. Price I.: Using the Web to build lasting relationship. New Media Age, 6 maja 2004.
10. Rosser M.: TV channels bed downs viewers' relationship. Precision Marketing, 17 stycznia 2003.
11. Waterman D.: The Economics of Internet TV: New Niches vs Mass Audience. info 2001, t. 3, nr 3, s. 215-229.
12. Oficjalny serwis internetowy BBC. http://www.bbc.co.uk, 8.05.2009.
13. Editorial Guidelines – oficjalny dokument BBC. www.bbc.co.uk/guidelines/editorialguidelines/edguide/, 9.05.2009
14. http://www.alexa.com/, 1.08.2009.

Rozdział 21

Efektywność nauczania informatyki w szkołach ponadgimnazjalnych przez pryzmat wyników egzaminów państwowych

Sławomir Iskierka, Janusz Krzemiński, Zbigniew Weżgowiec
Politechnika Częstochowska
wezgow@el.pcz.czest.pl

Streszczenie

W pracy podjęto próbę analizy efektywności nauczania informatyki przez pryzmat wyników osiąganych przez uczniów szkół ponadgimnazjalnych na egzaminach maturalnych i dyplomowych. Przeanalizowano zależności pomiędzy stopniem nasycenia placówek edukacyjnych w sprzęt teleinformatyczny w poszczególnych województwach a wynikami uzyskiwanymi na egzaminach państwowych. Zwrócono uwagę na duże zróżnicowanie istniejące pomiędzy poszczególnymi regionami kraju związane z szerokopasmowym dostępem szkół do Internetu. Wskazano na wciąż niedostateczne przygotowanie uczniów do świadomego i bezpiecznego korzystania z zasobów Internetu jak również bardzo pobieżną znajomość problemów związanych z ochroną własności intelektualnej. Podjęto próbę zdiagnozowania przyczyn wpływających na aktualny stan nauczania informatyki w szkołach ponadgimnazjalnych.

1. Wstęp

Zastosowanie technologii informatycznych i teleinformatycznych we współczesnej dydaktyce obejmuje coraz większe jej obszary, wkraczając praktycznie do wszystkich przedmiotów nauczania. W związku z tym wzrasta rola i znaczenie informatyki jako przedmiotu realizowanego w szkołach ponadgimnazjalnych. Wobec tych faktów, zasadnym jest przeanalizowanie wpływu dostępności sprzętu komputerowego na poziom wiedzy uczniów szkół

ponadgimnazjalnych z informatyki. Dobrym wskaźnikiem tego wpływu mogą być wyniki uzyskiwane przez uczniów na egzaminie maturalnym. Przeanalizowanie tych wyników ciągu kilku ostatnich lat i skorelowanie ich z rozwojem infrastruktury informatycznej może być przyczynkiem do dyskusji nad skutecznością, celowością i ewentualnie opracowywaniem nowych form nauczania informatyki. Problem staje się tym bardziej istotny, jeżeli przeanalizuje się wyniki rekrutacji na kierunek informatyka na wyższe uczelnie [7, 8]. W roku akademickim 2007/2008 na ogólną liczbę 565 765 studentów przyjętych do szkół publicznych i niepublicznych na informatykę zgłosiło się tylko 18 890 osób, co stanowi zaledwie 3,34% ogólnej liczby przyjętych na studia. Analogiczna sytuacja powtórzyła się w roku akademickim 2008/2009, w którym na ogólną liczbę przyjętych studentów wynoszącą 591 096 na informatykę przyjęto 19 488, co stanowi 3,30% ogólnej liczby przyjętych. Brak wykształconych informatyków może w przyszłości spowodować osłabienie tempa wzrostu gospodarki, w której technologie informatyczne i teleinformatyczne będą odgrywały coraz większą rolę.

2. Infrastruktura informatyczna szkół ponadgimnazjalnych

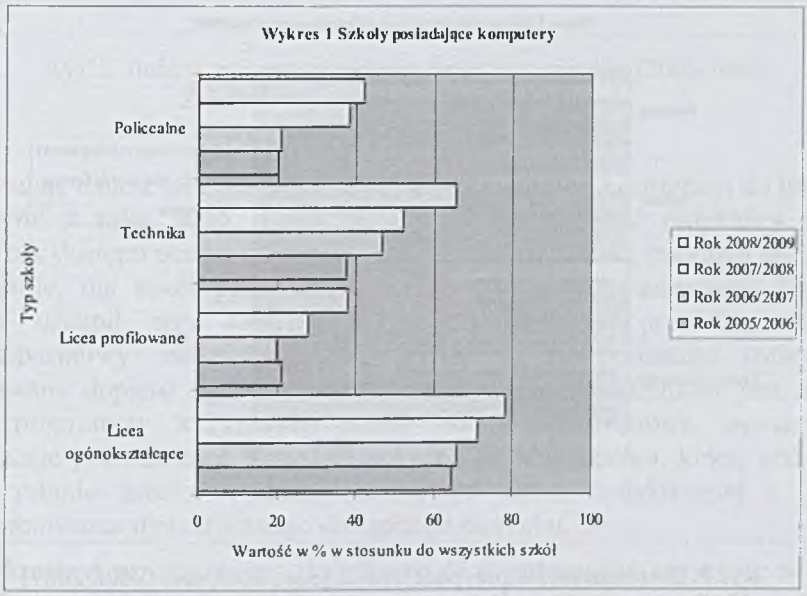
Wyposażenie szkół w sprzęt komputerowy jest bardzo zróżnicowane i zależy od typu szkoły i jej lokalizacji. Dobrym źródłem informacji na ten temat jest opracowanie Edukacja Informatyczna 2002. Zawarte tam dane są bardzo istotne albowiem obrazują stan wyposażenia polskiego szkolnictwa w sprzęt komputerowy u progu wprowadzanej reformy oświatowej. Dane te umożliwiają ocenę obecnej sytuacji i obserwację dynamiki zmian w wyposażaniu szkół w sprzęt komputerowy. Do najważniejszych wniosków sformułowanych w tym opracowaniu można zaliczyć:

- od czasu zmian ustrojowych w Polsce tj. od 1989 roku liczba komputerów w szkołach wzrosła prawie dziesięciokrotnie, co należy uznać za fakt bardzo pozytywny,
- średnio na jeden komputer przypadało wówczas 30 uczniów, a na komputer z dostępem do Internetu 38 uczniów. Uwzględniając tylko szkoły podstawowe i średnie porównano te dane dotyczące Polski i Unii Europejskiej. I tak, ilość uczniów przypadających na jeden komputer wynosiła wówczas w Polsce 33 (na komputer z dostępem do Internetu – 52), a w Unii Europejskiej 11 (na komputer z dostępem do Internetu – 24),

Interesującym jest zestawić te dane z danymi zawartymi w Małym Roczniku Statystycznym GUS za lata 2007, 2008 i 2009, dostępnymi na stronie internetowej GUS [5], dotyczącymi wyposażenia szkół w sprzęt komputerowy w latach 2005/2006 – 2008/2009. Na wykresie 1 przedstawiono, jaki procent szkół ponadgimnazjalnych, jest wyposażona w komputery.

Analizując ten wykres można stwierdzić dużą rozpiętość występującą pomiędzy poszczególnymi typami szkół. Najliczniej wyposażone w sprzęt komputerowy są licea ogólnokształcące i technika.. Wśród których odpowiednio (dane dotyczą roku szkolnego 2008/2009) 78,1%, i 65,6% jest wyposażona w komputery. Najmniejszy odsetek szkół wyposażonych w komputery to szkoły policealne i licea profilowane. Odsetek ten jest praktycznie stały dla tej grupy szkół i wynosi około 40%. Istotnym jest jednak fakt, że w ciągu ostatnich dwóch lat w tej grupie szkół nastąpił prawie dwukrotny wzrost tego wskaźnika. Należy uznać to za bardzo pozytywne działanie. Optymizmem napawa również fakt, że we wszystkich typach szkół, mino istniejących dysproporcji, procent szkół wyposażonych w komputery systematycznie wzrasta.

Technika plasują się na drugim miejscu z 65,6% udziałem. Mają natomiast największy procentowy przyrost wśród szkół, które zostały wyposażone w komputery w 2009 roku wynoszący ponad 13 %.

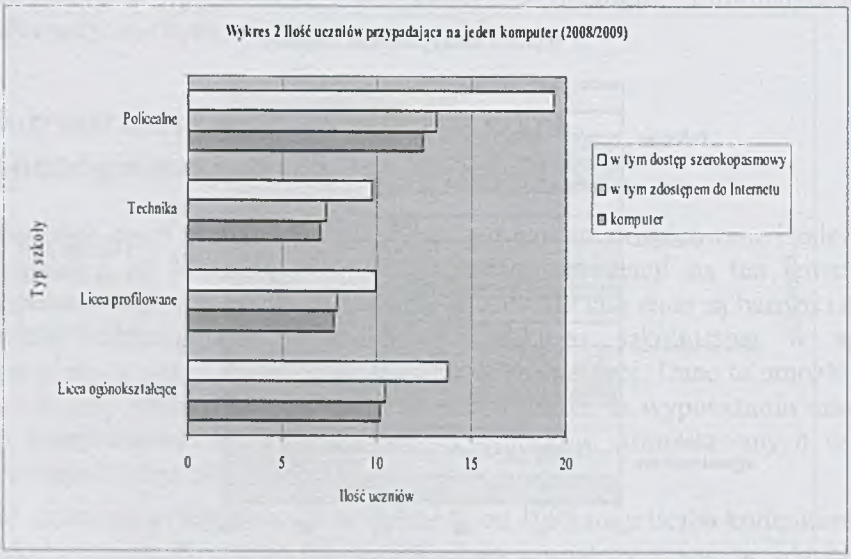


Rys.1. Szkoły posiadające komputery
Źródło: opracowanie własne na podstawie
Mały Rocznik Statystyczny GUS (lata 2007 - 2009)

Istotniejszym parametrem decydującym o możliwości wykorzystania komputerów w dydaktyce i zwiększającym komfort pracy tak ucznia jak i nauczyciela jest wskaźnik mówiący o ilości uczniów przypadających na jeden komputer. Parametr ten przedstawiono na wykresie 2.

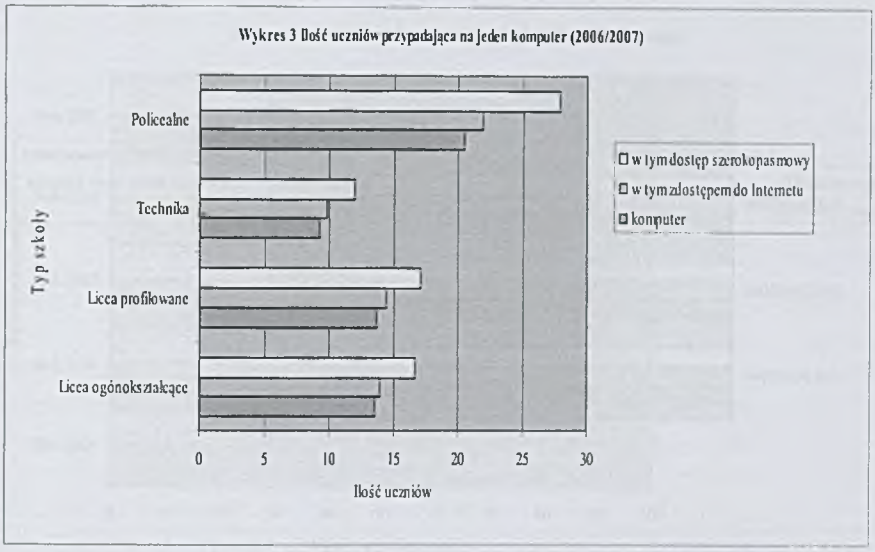
Najkorzystniejsza sytuacja występuje w grupie szkół technicznych. Wszystkie trzy parametry: ilość uczniów przypadająca na jeden komputer, ilość uczniów przydająca na jeden komputer z dostępem do Internetu i ilość uczniów przypadająca na jeden komputer z szerokopasmowym dostępem do Internetu są

najwyższe wśród wszystkich szkół. Przy czym ilość uczniów przydających na jeden komputer i na komputer dostępem do Internetu oscyluje wokół liczby 7. Jedynie wskaźnik związany z dostępem do szerokopasmowego Internetu zwiększa się do około 10. Podobne wskaźniki osiągnięto dla liceów profilowanych. Gorsza sytuacja występuje w liceach ogólnokształcących. Wskaźnik mówiący o ilości uczniów przypadających na jeden komputer i na komputer z dostępem do Internetu wynosi tutaj około 10, a wskaźnik związany dostępem szerokopasmowym 14. Najniekorzystniejsza sytuacja występuje w szkołach policealnych, gdzie wszystkie parametry przekraczają wartość 10, by dla dostępu szerokopasmowego osiągnąć wartość 18. Charakterystyczne jest również to dla pozostałych typów szkół, w których dostęp szerokopasmowy do Internetu jest też najbardziej utrudniony. Świadczy to o tym, że dostęp do najnowocześniejszych technologii teleinformatycznych, na przykład wideokonferencji, jest ciągle poza zasięgiem większości szkół w Polsce.



Rys.2. Liczba uczniów przypadająca na jeden komputer (2008/2009)
Źródło: opracowanie własne na podstawie
Mały Rocznik Statystyczny GUS 2009

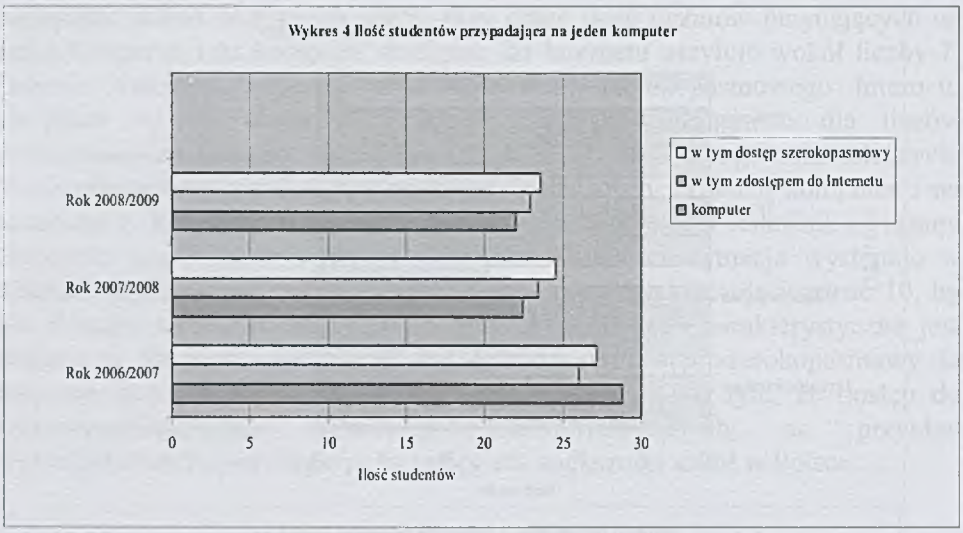
Dla porównania na wykresie 3 przedstawiono te same dane z przed dwóch lat. Widać znaczny postęp w dostępie uczniów tak do komputerów jak i do Internetu. Niemniej sytuacja, kiedy każdy uczeń będzie miał personalizowany dla swoich potrzeb komputer wydaje się być jeszcze odległa.



Rys. 3. Ilość uczniów przypadająca na jeden komputer (2006/2007)
Źródło: opracowanie własne na podstawie
Mały Rocznik Statystyczny GUS 2007

Porównując dane z lat 2006/2007 i 2008/2009 związane z dostępem do Internetu z danymi z roku 2000, można z satysfakcją zauważyć radykalny postęp. Wskaźnik dostępu ucznia do komputera, w polskiej szkole, poprawił się prawie trzykrotnie, dla szkół policealnych i pięciokrotnie dla techników. Podobny wzrost nastąpił przy dostępie ucznia do Internetu, przy czym dostęp szerokopasmowy może być, ze względu na nowoczesność technologii, realizowany dopiero od kilku lat. Poprawienie tych wskaźników jest zasługą wielu programów wdrażanych przez władze samorządowe, oświatowe i organizacje pozarządowe. Powstała pokaźna baza sprzętowa, której efektywne wykorzystanie zależy w dużej mierze od kadry dydaktycznej i jakości oprogramowania dydaktycznego dostępnego na rynku.

Na wykresie 4 przedstawiono dodatkowo dane pokazujące nasycenie w sprzęt komputerowy szkół wyższych, przy czym za wskaźnik przyjęto ilość studentów przypadająca na jeden komputer.



Rys. 4. Liczba studentów przypadająca na jeden komputer
Źródło: opracowanie własne na podstawie
Mały Rocznik Statystyczny GUS (lata 2007 – 2009)

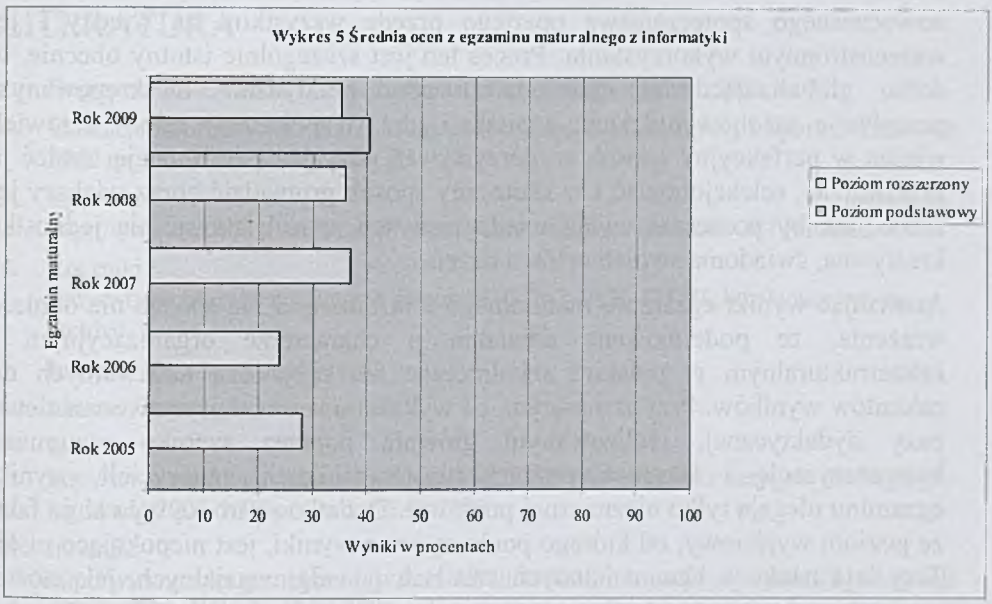
Parametr ten w sposób uśredniony pokazuje stan bazy komputerowej szkół wyższych. Zdając sobie sprawę z umownego charakteru tego wskaźnika, podano go jedynie jako pewne przybliżenie istniejącego stanu komputeryzacji szkół umożliwiające ocenę nasycenia szkół wyższych w sprzęt komputerowy

Nie może jednak pozostać niezauważony fakt, że osiągnęliśmy w Polsce wskaźniki, które w krajach Unii Europejskiej były praktycznie normą już prawie dziesięć lat temu. A pod względem dostępności do Internetu, znajdujemy się, według raportu przygotowanego między innymi przez Światowe Forum Ekonomiczne na 69 miejscu wśród 134 krajów świata [1].

3. Przegląd wyników egzaminu maturalnego

Przeanalizowanie wyników egzaminu maturalnego z informatyki z ostatnich lat ma na celu zaobserwowanie występujących zmian dotyczących poziomu opanowania przez absolwentów szkół ponadgimnazjalnych, wiedzy i umiejętności dotyczących tego przedmiotu.

Interesującym jest również zagadnienie jak wyposażenie szkół w sprzęt komputerowy wpływa na poziom wyników uzyskiwanych przez uczniów na egzaminie maturalnym z tego przedmiotu. Do analizy przyjęto wyniki z egzaminów maturalnych z lat 2005 – 2008, dostępnych na stronach internetowych Centralnej Komisji Egzaminacyjnej w Warszawie [3]. Średnie ocen z informatyki przedstawiono wykresie 5.



Rys. 5. Średnia ocen z egzaminu maturalnego z informatyki. Źródło: opracowanie własne na podstawie informacji udostępnianych przez Centralną Komisję Egzaminacyjną

Średnie wyniki z informatyki są niezadawalające. W latach 2005 i 2006 średnia nie przekroczyła 30%, a od roku 2007 waha się w okolicach 36%, co należy uznać za wynik również niekorzystny. Niepokojący jest również fakt, że informatykę na maturze wybiera bardzo nieliczna grupa uczniów. W analizowanym okresie odsetek uczniów wybierających informatykę na egzaminie maturalnym zmniejszył się kolejno z 1,5% w roku 2005 do 0,8% w 2006 i 0,49% w 2007 i 2009 roku.

4. Infrastruktura teleinformatyczna a wyniki egzaminu maturalnego z informatyki

Wyposażenie polskich szkół w sprzęt komputerowy jak wykazano w pierwszej części niniejszego opracowania uległo radykalnej poprawie. W ramach różnych programów zainstalowano w szkołach tysiące sieci komputerowych z dostępem do Internetu, w tym z dostępem szerokopasmowym, który wniósł nową jakość do dydaktyki w tym dydaktyki informatyki, stwarzając olbrzymie szanse (jak i zagrożenia) związane z praktycznie nieograniczonym dostępem do informacji. Wielu nauczycieli w ostatnich latach podnosiło swoje kwalifikacje zawodowe związane z wykorzystaniem technologii informacyjnych i teleinformatycznych na różnego rodzaju seminariach, kursach i studiach podyplomowych. Wszystkie te działania były i są podejmowane w celu poniesienia poziomu wykształcenia młodzieży, tak by mogła ona w przyszłości aktywnie włączyć się w rozwój

nowoczesnego społeczeństwa opartego przede wszystkim na wiedzy i jej wszechstronnym wykorzystaniu. Proces ten jest szczególnie istotny obecnie, w dobie globalizującej się gospodarki, przy praktycznie nieskrępowanym przepływie zasobów ludzkich, kapitału i idei. Współczesny młody człowiek winien w perfekcyjny sposób wykorzystywać pozyskaną informację, umieć ją przetwarzać, selekcjonować i w skuteczny sposób gromadzić coraz większy jej zasób, tak by poszerzać swoją wiedzę i w ten sposób stawać się jednostką kreatywną, świadomą swoich celów i dążeń.

Analizując wyniki egzaminu maturalnego z informatyki nie sposób nie odnieść wrażenia, że podejmowane działania o charakterze organizacyjnym i infrastrukturalnym w polskim szkolnictwie nie przynoszą adekwatnych do nakładów wyników. Przy znaczącym, co wykazano wcześniej, unowocześnieniu bazy dydaktycznej, realizowanym głównie poprzez szeroko rozumianą komputeryzację i istotnym podniesieniu kwalifikacji nauczycieli, wyniki egzaminu ulegają tylko nieznacznej poprawie. Dodatkową troskę wywołuje fakt, że poziom wyjściowy, od którego porównywano wyniki, jest niepokojąco niski. Trzy lata nauki w liceum i innych szkołach ponadgimnazjalnych, jak można przypuszczać, nie są w stanie znacząco podnieść poziomu wiedzy i umiejętności większości uczniów. Dodatkowo analizując wyniki egzaminów maturalnych w różnych szkołach ponadgimnazjalnych trudno nie odnieść wrażenia, że część młodzieży ze względu na swoje predyspozycje intelektualne nie jest w stanie przyswoić sobie wymaganego zakresu materiału obowiązującego na egzaminie maturalnym z informatyki. Niepokojącym jest również fakt, tak małego zainteresowania młodzieży informatyką, co może wydawać się sprzeczne z obiegowym poglądem o wysokich umiejętnościach młodzieży dotyczących wykorzystania nowoczesnych technik teleinformatycznych. Daje się również zauważyć olbrzymie dysproporcje występujące w ilości uczniów wybierających informatykę na egzaminie maturalnym w zależność od miejsca zamieszkania. Uczniowie pochodzący ze szkół wiejskich stanowili na przykład w 2005 roku zaledwie 1% wszystkich uczniów zdających na maturze informatykę. Nadzieją natomiast może napawać obligatoryjne wprowadzenie egzaminu maturalnego z matematyki, który być może wymusi na młodzieży bardziej systematyczną naukę powiązaną z logicznym analizowaniem przerabianego materiału

5. Podsumowanie

Autorzy zdają sobie sprawę z wielu uproszczeń przy prezentowaniu niniejszych zagadnień. Głównym celem tych rozważań jest jednak przede wszystkim potrzeba podzielnia się swoimi niepokojami i przemyśleniami dotyczącymi stanu polskiej oświaty, a w szczególności dydaktyką informatyki na poszczególnych etapach kształcenia, perspektywami jej rozwoju i grożącymi jej ewentualnie zaburzeniami

LITERATURA

1. Boguszewicz T, Polska internetową pustynią, „Rzeczpospolita" 2008, nr 72
2. Edukacja informatyczna 2002, Ministerstwo Edukacji Narodowej i Sportu, Warszawa 2002 na stronie www.men.waw.pl/.
3. Egzamin maturalny: http://www.cke.edu.pl/index.php?option=content&task=view&id=247 &Itemid=147
4. Roczniki statystyczne:http://www.stat.gov.pl/gus/5840_737_PLK_HTML.htm?action=show_archive
5. Internet: http://www.stat.gus.pl/
6. Internet: http://www.men.gov.pl/
7. Internet: http://www.nauka.gov.pl/mn/_gAllery/38/27/38276/wyniki_rekrutacji_na_studia_w_roku_akad._2007-2008.pdf, z dn.15.09.2009 r.
8. Internet: http://www.nauka.gov.pl/mn/_gAllery/49/50/49501/Informacja_o_wynikach_rekrutacji_2008-2009.pdf, z dn.15.09.2009 r.

Część 3.
Zastosowania systemów informatycznych

Rozdział 22

Zastosowanie reguł do wspomagania procesu analizy danych

Anna Zygmunt, Jarosław Koźlak, Piotr Domider, Wojciech Wójcik
Akademia Górniczo-Hutnicza
{azygmunt, kozlak}@agh.edu.pl, pdomider@gmail.com, wwojcik@student.agh.edu.pl

Streszczenie

W rozdziale opisany zostanie sposób analizy danych bilingowych przy pomocy metody analizy sieci społecznych. Obliczone miary sieci społecznej wykorzystane zostają do wyliczenia ról pełnionych przez poszczególne osoby w organizacji. W opracowanym systemie wspomaganie tego procesu, jak również rozwiązanie problemu poprawy jakości danych wejściowych próbuje się rozwiązać przy pomocy dodatkowej warstwy w postaci zdefiniowanych reguł systemu ekspertowego, wyrażającego wiedzę eksperta zarówno dotyczącą postaci danych wejściowych jak i struktury analizowanych organizacji.

1. Wprowadzenie

Regułowe systemy ekspertowe swą popularność zawdzięczają głównie prostocie ich konstrukcji i jasności procesu wnioskowania, w którym biorą udział. Klasyczne systemy ekspertowe traktowane były jako niezależne systemy z bazą wypełnioną wiedzą uzyskaną od eksperta i zapisaną w postaci reguł. Użytkownik takiego systemu – przeprowadzając z nim konsultacje (odpowiadając najczęściej na szereg pytań) – uzyskiwał diagnozę.

Takie podejście powodowało szereg problemów (związanych głównie z akwizycją wiedzy) i doprowadziło do spadku zainteresowania systemami ekspertowymi. Obecnie można obserwować wzrost zainteresowania tematyką, a

to głównie za sprawą tzw. *Business Rules* [4], w którym baza reguł stanowi dodatkową warstwę logiczną dla systemu bazodanowego.

W naszych pracach wykorzystaliśmy reguły do wspomagania procesu wstępnego przygotowania danych bilingowych, które następnie będą wykorzystywane do budowy sieci społecznych. Analizując pewne miary tak wygenerowanej sieci (głównie centralność) próbujemy poszczególnym osobnikom przypisać role, które mogą pełnić w organizacji. Poprawę dokładności takiej klasyfikacji próbowaliśmy uzyskać wprowadzając – przy pomocy reguł – pewną wiedzę dziedzinową odnośnie funkcjonowania danej organizacji.

2. Sieci społeczne

Interakcje zachodzące w grupach społecznych są coraz bardziej zróżnicowane (np. spotkania, rozmowy telefoniczne, wymiana poczty elektronicznej, czy komunikacja przy pomocy komunikatorów internetowych), a szybki rozwój technologii komputerowych ułatwia ich obserwowanie i analizowanie.

Zależności takie można przedstawić przy pomocy sieci zależności, zwanych Sieciami Społecznymi (ang. *Social Network*) [3, 7]. Role węzłów pełnią w nich osoby, a powiązania reprezentują zachodzące między nimi interakcje określonego typu. Dziedziną zajmującą się badaniem występujących w takich sieciach zależności jest Analiza Sieci Społecznych (ang. *Social Network Analysis – SNA)*, w ramach której wykrywane są i interpretowane społeczne wzorce zależności między osobami.

Metody SNA są szeroko stosowane w wielu dyscyplinach i przy analizie różnorodnych zjawisk. Wykorzystanie metod sieciowych jest szczególnie przydatne w badaniach funkcjonowania organizacji, w analizie treści publikowanych na blogach, do analizy relacji między instytucjami i firmami, systemów finansowych i bankowych, badań współautorstwa i cytowań w literaturze naukowej.

Istotą analiz SNA jest koncentrowanie się nie na indywidualnych cechach jednostek, ale na własnościach całego układu, na który składają się oddziałujące ze sobą elementy. Całość staje się tutaj czymś więcej niż sumą jednostek i w związku z tym istotne stają się relacje pomiędzy osobnikami należącymi do danej społeczności.

2.1. Metody analizy sieci społecznych

Analiza sieciowa ujawnia strukturę społeczną i pomaga śledzić drogi, którymi informacje mogą się rozprzestrzeniać oraz wyjaśniać, dlaczego pewne struktury społeczne pozwalają na szybkie ich rozprzestrzenianie się, podczas gdy inne mogą zawierać obszary, do których trudno jest dotrzeć [5]. Jedną z pierwszych metod zaproponowanych przez analityków sieci społecznych była ocena

centralności położenia jednostek w ich sieciach społecznych, które to położenie wskazuje na względną ważność. Do tej pory opracowano w teorii grafów i analizie sieci wiele sposobów wyliczania centralności węzłów w sieci. Miary te można podzielić na cztery podstawowe grupy [3, 2]:

1. centralność wierzchołków według stopnia (ang. *degree centrality*),
2. centralność oparta o wektor własny (ang. *eigenvector centrality*),
3. centralność pomiędzy (ang. *betweenness centrality*),
4. centralność wierzchołków według bliskości (ang. *closeness centrality*).

I tak: centralność wierzchołków według stopnia mierzy się ilością bezpośrednich powiązań jakie dany węzeł posiada. Węzeł, którego centralność wierzchołków według stopnia jest równy zero, jest izolowany. Z kolei węzeł z najwyższą wartością tego rodzaju centralności traktowany jest w sieci jako najbardziej aktywny. Centralność wierzchołków według stopnia jest najprostszym pomiarem centralności i często wysoce efektywnym w ocenie wpływu czy ważności danego węzła: w wielu sytuacjach osoby z większą liczbą powiązań mają większą moc. Niekoniecznie jednak zawsze im więcej połączeń, tym lepiej, gdyż ważne jest dokąd te połączenia prowadzą i w jaki sposób węzły te są połączone dalej. Spostrzeżenie to znalazło odzwierciedlenie w drugiej kategorii centralności: według wektora własnego. Okazuje się, że połączenia z ludźmi, którzy są ważni czy wpływowi powodują, że osoba sama staje się bardziej wpływowa, niż gdyby była połączona z mnie wpływowymi. Centralność według wektora własnego zależy zarówno do liczby jak i jakości powiązań: węzeł z mniejszą liczbą ale wysokiej jakości połączeń może mieć większą wartość centralności według wektora własnego niż inny, z dużo większą liczbą, ale przeciętnych połączeń. Miara ta traktowana jest jako wyznacznik popularności. Wariantem centralności według wektora własnego jest Ranking Page'a (ang. *Page Rank*) wykorzystywany przez przeglądarkę Google do nadawania indeksowanym stronom internetowym określonej wartości liczbowej oznaczających ich jakość. Innymi miarami oceny ważności węzłów z tej kategorii są stopień koncentracji (ang. *hubness*) i stopień autorytetu (ang. *authoritativeness*).

Dwie następne kategorie miar centralności opierają się na koncepcji ścieżek grafu. Centralność wierzchołków według bliskości wyraża jak blisko pozostałych węzłów w sieci znajduje się dany węzeł: wysoka wartość oznacza, że ma on najszybszy dostęp do wszystkich pozostałych wierzchołków w sieci, gdyż ma najkrótsze ścieżki do wszystkich innych. Jest to bardzo dobra pozycja do monitorowania przepływu informacji w całej sieci. Przykładami miar z tej kategorii może być centralność środka ciężkości (ang. *Bary Center*) albo centralność Markov'a (ang. *Markov Centrality*).

Ostatnią grupą miar jest centralność pomiędzy, która wskazuje jak bardzo dany węzeł pośredniczy między pozostałymi (ważnymi) węzłami w sieci. Jest to liczba dróg pomiędzy każdymi dwoma węzłami w grafie przechodząca przez

dany węzeł, podzielona przez liczbę wszystkich dróg pomiędzy każdą parą węzłów. Centralność ta mierzy stopień, w jakim dany węzeł może funkcjonować jako pośrednik (broker) między węzłami w sieci.

3. Role w organizacjach

Na podstawie informacji zawartych w bilingach telefonicznych można zbudować sieć społeczną, w której węzłami są numery abonentów, natomiast relacje pomiędzy nimi określają wykonane połączenia. Mamy tu do czynienia z relacjami skierowanymi (osoba A dzwoni do osoby B) i wartościowanymi ilością wykonanych połączeń. Na podstawie informacji dotyczących numery rozmówcy i jego partnera wyliczane są podstawowe miary centralności omówione w poprzednim rozdziale.

Wyliczone w ten sposób parametry zostaną wykorzystane do wyliczania ról, jakie pełnią dane osobniki w sieci. John Arquilla i David Ronfeldt w [1] zaproponowali kilka ról, jakie mogą odgrywać przestępcy w organizacji. Są to:

- **Organizatorzy** - osoby stanowiące rdzeń całej organizacji, czasem małe grupy sterujące działaniem całej sieci, które determinują skalę i kierunek działania.
- **Izolatorzy** - ich rolą jest izolowanie rdzenia organizacji (organizatorów) przed infiltracją i zapobieganie narażaniu ich na zagrożenie. Jednostki te są również odpowiedzialne za transmisję informacji i zarządzeń od rdzenia do brzegów sieci. Poza tym kontrolują, czy informacja przepływająca z peryferii sieci nie była w stanie w żaden sposób narazić rdzenia na wyłączenie go z sieci.
- **Komunikatorzy** - odpowiadają za przepływ informacji pomiędzy dwoma węzłami w całej sieci. Są oni odpowiedzialni za transmisję rozporządzeń z rdzenia i powracających do niego odpowiedzi.
- **Strażnicy** - koncentrują się na bezpieczeństwie sieci i minimalizują podatność sieci na zewnętrzne ataki lub infiltrację. Kontrolują kto jest rekrutowany do sieci oraz oceniają jego lojalność w stosunku do sieci. Wymuszają również lojalność wszystkich członków grupy oraz ich rodzin. Zapobiegają również dezercji osobników z sieci, a gdy taka sytuacja nastąpi, minimalizują ryzyko z tym związane. Starają się naprawić powstałe w sieci uszkodzenia i nieprawidłowe działanie.
- **Rozszerzający** - zajmują się poszerzeniem sieci poprzez rekrutowanie nowych członków oraz poprzez łączenie innych sieci. Zachęcają do współpracy osobniki z innych sieci (np. prawników, policjantów, polityków). Kiedy im się to udaje, sieć zyskuje nowe źródła informacji - wpływa to na jej lepsze działanie. Ich cele to głównie wpływowe osoby, które są w stanie wprowadzić do sieci wysoki stopień ochrony.

- **Monitorujący** - są odpowiedzialni za wydajność sieci. Raportują słabości sieci
  i jej problemy do organizatorów, którzy dzięki temu mogą podjąć kroki naprawcze. Są odpowiedzialni za poprawę działania sieci, nieraz sami nadzorują korekty w sieci. Dbają o to, aby sieć była zdolna zareagować na nowe okoliczności oraz aby sieć była elastyczna.
- **Łącznicy** - osoby, które zostały zwerbowane do sieci, ale mimo wszystko kontynuują swoje działanie w innych sieciach (legalnych instytucjach rządowych, politycznych, finansowych). Takie osobniki również mogą wprowadzać dodatkową informację i ochronę.

Powyższe role dotyczą jedynie osobników silnie ustrukturalizowanych, czyli takich, którzy istnieją w sieci dłuższy okres. Pełnią role wymagające od nich zaufania ze strony pozostałych osobników. W tej charakterystyce brakuje jednak pozostałej części osobników, które pojawiają się w sieci sporadycznie lub nie pełnią żadnej ważnej lub zaufanej funkcji. Zatem oprócz wspomnianych powyżej ról, zaproponowano uwzględnienie jeszcze czterech innych ról.

Są to:

- **Żołnierze** - osoby nie pełniące w zasadzie żadnej ważnej funkcji w sieci. Wykonują one jedynie rozkazy innych osobników w sieci, stojących wyżej w hierarchii. Ich zwierzchnikami mogą być komunikatorzy, strażnicy, monitorujący, łącznicy, a nawet izolatorzy. Jest to dominująca rola w sieci kryminalnej. Tych osób zwykle jest najwięcej. Mogą się one zajmować podstawowymi, najmniej zaawansowanymi działaniami grupy. W zależności od charakteru grupy mogą to być kradzieże samochodów, rozprowadzanie narkotyków, wyłudzanie pieniędzy, zwalczanie członków innych grup przestępczych rywalizujących z ich macierzystą strukturą.
- **Rekruci** - są to nowe osoby w sieci. Mogą jeszcze nie być uznani przez sieć jako jej członkowie, mimo wszystko zaczynają już działać w jej ramach. Nowa osoba może być rekrutowana na różne pozycje, nie tylko te najniższe (żołnierz). Może się nawet okazać, że nowo rekrutowana osoba stanie się strażnikiem. Najbardziej charakterystyczną cechą rekruta jest jego krótka obecność w sieci.
- **Postronni** - są to osoby, które pomimo faktu przynależności do sieci społecznej nie należą do sieci kryminalnej. Takie osoby to np. rodzina przestępców - jego żona nie musi działać w grupie przestępczej, a mimo tego należy do sieci (z racji częstych kontaktów z jednym z osobników sieci). Postronny z czasem może stać się aktywnym członkiem grupy przestępczej, jednak równie dobrze może nie mieć w ogóle świadomości jej istnienia.
- **Przypadkowi** - osoby, z którymi sieć komunikuje się bardzo rzadko. Są to osoby nie należące do sieci. Może się również zdarzyć, że sieć często komunikuje się

z przypadkowym, a mimo wszystko osobnik nie utożsamia się z siecią. Może to być numer pizzerii, czy też np. numer audiotele.

W pierwszym podejściu do problemu [6] założono, że jest to standardowy zestaw ról występujący w każdej organizacji przestępczej. Rolom tym przypisano pewne stereotypy działania wyrażone w zakresach parametrów SNA. Przyjęto np. że monitorujący, który jest odpowiedzialny za wydajność sieci i poprawę jej działania, a o słabościach sieci raportuje organizatorom powinien mieć niską wartość centralności według bliskości, gdyż raczej monitoruje obrzeża sieci. Z kolei izolator, którego zadaniem jest ochraniać rdzeń organizacji powinien mieć dość wysoką wartość centralności pomiędzy.

Przeprowadzając szereg eksperymentów z analizowaniem różnych organizacji, okazało się [8] że zdefiniowany, opisany powyżej, zestaw 11 podstawowych ról i ich charakterystyka wyrażona zakresami parametrów SNA jest zbyt ogólny. Każdy rodzaj organizacji ma swój specyficzny zestaw ról, różniący się zarówno licznością, jak i charakterystyką.

4. Wstępne przetwarzanie danych za pomocą reguł

Bilingi w oparciu, o które budowane są sieci społeczne dostarczane są przez kilku operatorów sieci telefonicznych i różnią się między sobą znacząco, między innymi formatami oraz liczbą pól. Dostarczyciele usług telefonii komórkowej niestety nie dopracowali się standardowego formatu bilingów. Dodatkowo, operatorzy zmieniają własne formaty bilingów, poprzez np. wprowadzanie dodatkowych do nich informacji. Struktura bilingów jest więc dynamiczna i należałoby wprowadzić pewne mechanizmy wstępnego przetwarzania danych wejściowych (ich czyszczenie) uwzględniający tę dynamikę i umożliwiające łatwe w obsłudze reagowanie na wprowadzane przez operatorów zmiany.

Weryfikacja poprawności prowadzona była w dwóch etapach: częściowo w czasie wczytywania bilingów do bazy, a następnie - po ich wczytaniu, gdy dane bilingowe były już zapisane w odpowiednich tabelach bazy danych.

Procedury wykonywane w ramach pierwszego etapu są bardziej uniwersalne i zdefiniowane zostały w odpowiednich klasach aplikacji. Dane są sprawdzane pod kątem poprawności dat, poprawności numerów telefonicznych (znaki nienumeryczne w numerach telefonów), długości i kierunku rozmów (czy wchodzące, czy wychodzące). Błędne rekordy są odrzucane przed zapisaniem do bazy, a informacje o tym wpisywane do odpowiednich rejestrów.

Na tym etapie usuniętych zostaje większość błędnych danych. Pozostaje jeszcze problem np. nadmiarowych wpisów, do których można zaliczyć:

- występowanie numerów będących postfiksami innych numerów (np. poprawny numer telefonu oraz identyczny numer poprzedzony cyframi 78),

- występowanie identycznych numerów (wraz z pozostałymi danymi) znajdujących się w różnych wierszach,
- występowanie powielonych rozmów: różnice w czasie rzędu 2 sekund albo powielone rozmowy ze zmienionym kierunkiem.

Nadmiarowości te usunięte zostają w drugim etap walidacji, który jest dużo bardziej elastyczny, bo przeprowadzony przy pomocy reguł systemu ekspertowego Drools[1].

Reguła taka jest wyrażeniem postaci "*when* przesłanka *then* konkluzja". W przesłance określamy warunki, które muszą zostać spełnione, żeby wykonana została akcja opisana w konkluzji reguły.

Baza reguł stanowi niejako zewnętrzną, niezależną warstwę w stosunku do systemu. Dane pobierane są z bazy danych przy pomocy Spring JDBC[2], po czym następuje inicjalizacja systemu Drools i załadowanie pliku z regułami. W wyniku działania reguł, dla każdego wiersza danych odnotowane zostaje czy wiersz ten należy zmienić lub usunąć, następnie dane po modyfikacji zostają ponownie przesłane do bazy danych.

Oprócz zauważonych nadmiarowości, wykorzystano informację, którą można traktować jako pewnego rodzaju wiedzę dziedzinową, dotyczącą sposobu dzielenia przez jednego z operatorów długich rozmów na kilka krótszych. W bilingach pewnego operatora jedna rozmowa pojawiała się jako kilka o określonych czasach trwania (po 900 sekund).

Aktualnie zaimplementowano następujące reguły:

1. regułę znajdującą pary bilingów z tymi samymi rozmówcami i o podobnej dacie (w przypadku gdy czas trwania rozmowy dla co najmniej jednego bilingu nie jest podany),
2. regułę znajdującą pary bilingów z tymi samymi rozmówcami i o podobnej dacie (w przypadku gdy czas trwania rozmowy jest podany w obu bilingach),
3. regułę wykrywającą numery będące postfiksami innych numerów,
4. regułę znajdującą identyczne numery wśród rozmówców,
5. regułę łączącą rozdzielone rozmowy w jedną całość.

Reguły 1., 2. i 5. automatycznie usuwają znalezione nieprawidłowości z bazy. W wyniku działania reguł 1. i 2., przy usuwaniu powielonej rozmowy w bazie danych zawsze pozostawiany zostaje wpis z wcześniejszą z dat rozmów oraz dłuższym z czasów jej trwania (może on zostać w tym celu zmodyfikowany, by uwzględnić dane z obu powielonych wpisów). Reguły te korzystają z

[1] http://www.jboss.org/drools/.
[2] http://static.springframework.org/spring/docs/2.0.x/reference/jdbc.html

konfigurowalnego parametru, który określa dopuszczalną tolerancję między datami rozpoczęcia rozmów.

Reguła 5. może łączyć niekoniecznie tylko rozmowy dotyczące zaobserwowanych przypadków, gdzie czas jednego fragmentu rozmowy wynosi wskazane przez analityka 900 s. Dowolna para rozmów, gdzie jedna zaczyna się dokładnie po drugiej, a kolejność rozmówców jest ustalona i zachowana zostaje scalona w jedną rozmowę.

Reguła 3. i 4. nie modyfikują zasadniczych danych w bazie, ale wskazują potencjalnie podejrzane wpisy, które mogą świadczyć o błędnych lub sprzecznych danych.

W ten sposób możemy wprowadzać do systemu dowolną wiedzę dziedzinową.

5. Analiza interakcji pomiędzy rolami

Jak już zostało to opisane w rozdziale 3., w opracowanym systemie jest możliwość definiowania specyficznego zestawu ról dla różnych rodzajów organizacji.

W celu poprawy jakości przeprowadzanych analiz zdecydowano się dodatkowo rozbudować system o fragmenty umożliwiające bezpośrednie włączanie posiadanej przez analityka wiedzy dziedzinowej do struktury wygenerowanego przez system grafu. Wiedza ta - w aktualnej wersji systemu - dotyczy specyficznych dla konkretnej grupy informacji o strukturze połączeń między osobami pełniącymi określone role.

Struktura ta jest różna i charakterystyczna dla różnych typów organizacji przestępczych.

Włączenie takiej wiedzy do przeprowadzanych w systemie analiz powinno wpłynąć korzystnie na jakość dopasowania ról do niezidentyfikowanych jeszcze numerów oraz doradzać, które numery należy obserwować bardziej szczegółowo, gdyż mogą one potencjalnie pełnić ważne role w badanej organizacji.

Jednym z możliwych rozwiązań jest przeprowadzenie - dla każdego wierzchołka w grafie - analizy jego bezpośredniego sąsiedztwa pod kątem przyznanych przez system ról. Na podstawie charakterystyk bezpośrednich interakcji danego rozmówcy z innymi rozmówcami, można wnioskować do jakiej roli (lub ról) w danym rodzaju organizacji można jego zachowanie dopasować. Metoda ta zakłada poprawę dopasowania ról, po wyznaczeniu ich podstawowym algorytmem systemu. Do zdefiniowania sposobu zmian ról na podstawie informacji o sąsiedztwie wierzchołków wykorzystano system regułowy Drools. Reguły takie są automatycznie uruchamiane i niezależne od modułów systemu realizujących przetwarzanie danych bilingowych.

Docelowo analityk będzie mógł - dzięki rozwiązaniu opartego na zastosowaniu reguł - wpisywać znane mu z innych źródeł zależności (dotyczące nie tylko interakcji między rolami).

5.1. Opis działania systemu z możliwością definiowania reguł

W podstawowym systemie analiza danych polegała na przeprowadzeniu analizy statystycznej oraz ewentualnym uwzględnieniu wiedz dziedzinowej w sposób manualny, tzn. należało ręcznie zaznaczyć zmianę ról oraz poszukiwać powiązania między nimi). Często ogólna struktura organizacji przestępczej znana jest w całości lub znane są podstawowe zależności występujące między rolami. Przykładowo dla niektórych organizacji można, z dużą dozą prawdopodobieństwa założyć, iż organizator będzie komunikował się przede wszystkim z innymi organizatorami, a komunikacja z innymi rolami będzie odbywała się za pośrednictwem roli izolatora (czyli organizator będzie połączony przede wszystkim z pozostałymi organizatorami oraz izolatorami, a powiązania z pozostałymi rolami będą znacząco słabsze).

Taką wiedzę można prosto zapisać za pomocą reguł, które będą automatycznie uruchamiane po wybraniu typu organizacji i ustawieniu odpowiednich progów tolerancji określających stopień dopasowania do wzorców zachowań.

Podsumowując, działanie systemu będzie wyglądało następująco:

- wczytanie danych do systemu,
- wstępne przygotowanie danych z uwzględnieniem reguł czyszczenia opisanych w rozdziale 4.,
- przeprowadzenie analizy statystycznej, wyznaczenie parametrów oraz przydzielenie ról do węzłów,
- ręczne wskazanie ról dla wierzchołków, pełniących już zidentyfikowane funkcje w badanej organizacji, co spowoduje, że takie wskazania nie będą modyfikowane przez system regułowy,
- opracowanie reguł uwzględniających zależności pomiędzy rolami oraz uruchomienie systemu regałowego,
- analiza wyników, ewentualne wprowadzenie zmian (modyfikacja samych reguł, zmiana progów tolerancji, etc). Na tym etapie możliwe są również zmiany ręczne dla przypadków nie ujętych w regułach (analogicznie jak w punkcie 4.),
- wyciągnięcie wniosków do dalszego działania (ewentualne rozszerzenie zakresu bilingowanych numerów lub podjęcie działań).

5.2. Koncepcja reguł uwzględniających zależności między rolami

W systemie jest możliwość definiowania dwóch schematów reguł: *podstawowego* i *rozszerzonego*. Dla każdego wierzchołka w analizowanym grafie przeprowadzone zostaje odrębne wnioskowanie na podstawie tych schematów.

5.2.1. Podstawowy schemat reguł

Jest to przypadek, w którym uwzględnia się jedynie wiedzę o fakcie sąsiedztwa dwóch ról bez uwzględniania siły tego sąsiedztwa. Można np. wprowadzić do systemu wiedzę mówiącą o tym, że rola Organizatora nie komunikuje się bezpośrednio z Żołnierzami, a czyni to za pomocą Izolatorów oddzielających go od mniej istotnych ról w organizacji. Taka wiedza przekłada się na regułę: jeśli w danych istnieje węzeł, który komunikuje się z Izolatorami, to należy zwiększyć przydział punktów dopasowania do roli Organizatora dla tego węzła. Na potrzeby analizy efektywniej jest analizować całe grupy sąsiadów o wyznaczonych rolach (np. grupa Izolatorów), a nie indywidualne interakcje badanego węzła z sąsiednimi węzłami, gdyż dana reguła, dotycząca konkretnych ról sąsiednich, może być uruchomiona tylko raz (każda uwzględniona w regule grupa sąsiadów, jako całość, występuje w analizowanych danych albo nie) i nie jest czuła na liczność grup.

5.2.2. Rozszerzony schemat reguł

W tym przypadku, poza faktem występowania grup sąsiadów pełniących określone role, uwzględnia się również moc ich połączenia z każdą grupą (tworzonej na podstawie ilości nawiązanych rozmów oraz ich długości) oraz liczności grup.

Analityk musi posiadać bardziej szczegółową wiedzę na temat charakteru powiązań między rolami w badanej organizacji.

Taka wiedza pozwala wzbogacić podstawową postać reguły o dodatkowe warunki – grupa nie tylko musi wystąpić w danych, ale musi być odpowiednio liczna (co jest rozumiane jako procentowy udział liczności grupy w stosunku do liczby wszystkich sąsiadów badanego wierzchołka), a komunikacja odpowiednio intensywna (co jest rozumiane jako procentowy stosunek wagi połączenia z daną grupą do sumy wag wszystkich grup sąsiadów danego wierzchołka).

Dodatkowo, można tutaj rozbić analizę na dwa niezależne przypadki: analiza połączeń wychodzących oraz analiza połączeń przychodzących. Takie podejście może być uzasadnione, jeśli komunikacja w organizacji ma silnie asymetryczny charakter.

Możliwa jest np. sytuacja, gdy rola Organizatora posiada powiązania z kilkoma różnymi grupami. Z jedną grupą powiązanie jest jednorazowe lub rzadkie,

a z pozostałymi - regularne. Może to odpowiadać sytuacji, w której grupa rzadko komunikująca się z Organizatorem pełni rolę Przypadkowego (np. pizza, taxi), a grupa komunikująca się z Organizatorem częściej – rolę Izolatora. Dodatkowo komunikacja z rolą Przypadkowy powinna mieć charakter jednostronny i zawierać głównie połączenia wychodzące od Organizatora do tej grupy.

Na podstawie takiego opisu można ułożyć już bardziej rozbudowaną regułę dla roli Organizator: jeśli wierzchołek posiada dwie grupy sąsiadów: jedną z rolą Izolatora (> 50% liczność, > 50% komunikacja, w obie strony) oraz drugą Przypadkowy (liczność 10%, komunikacja 20% dla połączeń wychodzących), to należy zwiększyć przydział punktów dopasowania do roli Organizator dla tego węzła.

5.2.3. Budowa reguł sąsiedztwa

Wiedza dziedzinowa odnośnie struktury połączeń między wierzchołkami została pozyskana poprzez analizę ilości połączeń wchodzących i wychodzących dla każdej roli z innymi rolami oraz porównanie tych wartości z ogólnymi (odrębnie dla każdej sprawy). Do konstrukcji reguł przydatne były następujące wyliczenia:

1. ilość połączeń pomiędzy rolami dla zagregowanych danych,
2. procent ze wszystkich połączeń inicjowanych przez daną rolę dla zagregowanych danych,
3. procent ze wszystkich połączeń w sieci dla zagregowanych danych,
4. ilość połączeń pomiędzy rolami dla nie zagregowanych danych,
5. procent ze wszystkich połączeń inicjowanych przez daną rolę dla nie zagregowanych danych,
6. procent ze wszystkich połączeń w sieci dla nie zagregowanych danych.
7. Przykład takiego zestawienia pokazuje tab. 1.

Tab. 1. Liczba połączeń Organizatora z innymi rolami dla wybranej sprawy

Rola inicjująca połączenie	Rola odbierająca połączenie	1	2	3	4	5	6
Organizator	Komunikator	2	1%	0%	17	0%	0%
Organizator	Strażnik	2	1%	0%	24	0%	0%
Organizator	Izolator	5	2%	0%	434	6%	0%
Organizator	Żołnierz	143	65%	5%	632	82%	5%
Organizator	Postronny	67	30%	2%	783	10%	1%
Organizator	Organizator	2	1%	0%	94	1%	0%

Przy opracowywaniu reguł przyjęto założenie, że w celu uwzględnienia wiedzy o sąsiedztwie należy dla każdego wierzchołka zbadać jego najbliższe otoczenie

– jego bezpośrednich sąsiadów – i sprawdzić jakie role maksymalne zostały dla nich przydzielone w wyniku analizy SNA (lub BD).

Następnie wśród sąsiadów wyszukiwane są takie ich grupy (zestawy ról) jakie odpowiadają charakterystyce komunikacji pewnej określonej roli (na podstawie wyliczeń zestawionych w Tab. . Za każdą sytuację, pasującą do wzorca, zwiększany jest przydział punktowy badanego wierzchołka do tej roli. Takie postępowanie jest powtarzane niezależnie dla każdej z analizowanych ról.

Jako rezultat, każda z analizowanych ról generuje przynajmniej jedną regułę, która bada dopasowanie do niej wierzchołków w grafie. Założono, że każde uruchomienie reguły będzie zwiększać przydział punktowy badanego wierzchołka dla odpowiadającej jej roli o pewną konfigurowalną wartość.

By uwzględnić siłę powiązań między badanym wierzchołkiem a grupami jego sąsiadów, części warunkowe reguł powinny zostać rozbudowane o dodatkowe warunki.

I tak, każda grupa sąsiadów powinna posiadać dwa dodatkowe parametry wyznaczane na podstawie struktury grafu: liczność grupy (*size*) oraz sumaryczną wagę krawędzi łączących daną grupę z wierzchołkiem analizowanym (*weight*).

Podobnie, analizowany wierzchołek powinien posiadać dodatkowe parametry określające sumaryczną liczbę przychodzących krawędzi od wszystkich grup (*inTotal*) i wychodzących krawędzi do wszystkich grup (*outTotal*) oraz sumaryczne wagi krawędzi przychodzących (*inWeightsTotal*) i wychodzących (out*WeightsTotal*).

Na podstawie wzajemnej relacji zestawów parametrów z analizowanego wierzchołka
i danej grupy sąsiadów można w sposób heurystyczny, szacować siłę powiązania pomiędzy nimi w następujących sposób:

- *size/inTotal* (lub *size/outTotal*) - określa jaki procent rozmówców węzła należało do danej grupy (parametr ten odpowiada kolumnie drugiej) ,
- *weight/inWeightsTotal* (lub *weight/outWeightsTotal*) - określa jaki procent rozmów trafiało do danej grupy (parametr ten odpowiada kolumnie piątej tabeli).

Na podstawie przeprowadzonych analiz zestawionych w tab. 1, można określić jaka powinna być wartość obu tych wskaźników, tak aby badany przypadek uznać za pasujących do wzorca danej roli.

Dodatkowo, aby warunek dopasowania mógł zostać spełniony, należy ustawić pewną tolerancję rozbieżności między zaobserwowanymi i wzorcowymi wartościami parametrów. Progi tolerancji są niezależnie konfigurowalne przez analityka
i określają:

- *sdelta* - próg tolerancji liczności grup (zakres [0..1]), określa o ile wartość *size/inTotal* (*size/outTotal*) może różnić się od wzorca,

- *wdelta* - próg tolerancji siły powiązań (zakres [0..1]), określa o ile wartość *weight/ inWeightsTotal* (*weight/outWeightsTotal*) może różnić się od wzorca.

5.2.4. Sposób uruchamiania reguł

Przeprowadzone eksperymenty wykazały, że jednokrotne uruchomienie reguł dla całego grafu może być niewystarczające i należy uwzględnić możliwość uruchamiania reguł w kilku iteracjach. Takie podejście umożliwia wykorzystanie nowego przydziału ról, opracowanego w poprzedniej iteracji, jako podstawy do wykonania następnej. Ilość iteracji winna być możliwa do zdefiniowania przez analityka – zbyt duża prowadzić może do zbytniego zwiększenia przydziału punktów dla dopasowanych przez reguły ról, w wyniku wielokrotnego uruchamiania tych samych reguł. Zbyt mała z kolei, może sprawić, że zmiany dokonane przez reguły będą niezauważalne, a przydział ról w grafie w ogóle nie ulegnie zmianie (zależy to również silnie od ilości punktów dodawanych dla roli w wyniku pojedynczego uruchomienia reguły).

6. Podsumowanie

W rozdziale przedstawiono możliwości włączania wiedzy dziedzinowej do systemu znajdującego role w organizacjach przestępczych przy pomocy analizy sieci społecznych. Wiedza ta została zdefiniowana przy pomocy reguł systemu ekspertowego i wspomaga analityka na etapie przygotowania danych oraz poprawia jakość przyporządkowania osób do pełnionych w organizacji ról. Przeprowadzone eksperymenty potwierdziły słuszność takiego podejścia w obydwu przypadkach. Reguły definiuje się w prosty sposób i są elastyczne: analityk może na bieżąco dokładać nową wiedzę. Pozwalają na wyrażenie wiedzy, którą często trudno byłoby zapisać w tradycyjny sposób.

LITERATURA

1. Arquilla J., Ronfeldt D.: Networks and Netwars: The Future of Terror, Crime and Militancy. RAND Corporation, 2002.
2. Bonacich P.: Factoring and Weighting Approaches to Status Scores and Clique Identification. Journal of Mathematical Sociology, 1972, 2.
3. Carrington P., Scott J., Wasserman S.: Models and Methods in Social Networks Analysis, Cambridge University Press, 2005.
4. Halle von B.: Business Rules Applied, Building Better Systems Using the Business Rules. Wiley Computer Publishing, New York, 2002.
5. Nooy de W., Mrvar A., Batagelj V.:Exploratory Social Network Analysis with Pajek. Cambridge University Press, 2005.

6. Piekaj W., Skorek G., Zygmunt A., Koźlak J.: Środowisko do identyfikowania wzorców zachowań w oparciu o podejście sieci społecznych. Technologie Przetwarzania Danych: II Krajowa Konferencja Naukowa, Poznań, 2007.
7. Scott J.: Social Network Analysis: A Handbook. Sage Publication, 2009.
8. Zygmunt A., Koźlak J.: Zastosowanie podejścia sieci społecznych do wspomagania prowadzenia analizy kryminalnej dotyczącej danych billingowych. Praktyczne elementy zwalczania przestępczości zorganizowanej i terroryzmu: nowoczesne technologie i praca operacyjna. red. Paprzycki L., Rau Z., Wolters Kluwer Polska. 2009.

Rozdział 23

Przeprowadzanie oceny skoringowej obiektów za pomocą modeli eksploracji danych Data Mining

Mirosława Lasek, Marek Pęczkowski
Uniwersytet Warszawski, Wydział Nauk Ekonomicznych
mlasek@wne.uw.edu.pl, mpeczkowski@wne.uw.edu.pl

Streszczenie

W rozdziale opisano wykorzystanie programu Enterprise Miner firmy SAS na potrzeby prognozowania za pomocą modeli eksploracji danych Data Mining, takich jak modele regresji logistycznej, drzew decyzyjnych i sieci neuronowych. Przedstawiono możliwość wykorzystania modeli do prognozowania poza Enterprise Miner w Base SAS na podstawie kodu programowego (tzw. kodu skoringowego), wygenerowanego w Enterprise Miner, jak i przeprowadzanie prognozowania w Enterprise Miner z użyciem węzła Score.

1. Wstęp

W książkach, artykułach i referatach konferencyjnych dotyczących eksploracji danych Data Mining można znaleźć obszerne i szczegółowe, a nawet wręcz drobiazgowe opisy procesów budowania, a także architektur gotowych modeli prognozowania, takich jak modele regresji logistycznej, drzew decyzyjnych, sieci neuronowych. Również dużo miejsca zajmują opisy uzyskiwanych wyników – prognoz wraz z ich interpretacjami. Na tym tle, zauważalny jest fakt niewielkiej uwagi poświęcanej samemu sposobowi otrzymywania prognoz [3]. Stąd zagadnieniu, w jaki sposób wykorzystać model, aby uzyskać prognozę, z pewnością interesującemu dla wielu osób, zainteresowanych praktycznym stosowaniem modeli eksploracji danych, postanowiliśmy poświęcić niniejszy rozdział. Wiele osób zwracało się do nas z zapytaniem o konkretne polecenia dla przeprowadzenia prognozy. Z tego powodu, opracowanie w niektórych miejscach przyjmuje charakter instrukcji pokazującej, w jaki sposób

przeprowadzić prognozę na podstawie zgromadzonych danych i zbudowanego modelu prognostycznego.

W pracy wykorzystujemy program Enterprise Miner firmy SAS [1, 4]. Zakładamy, że czytelnik zna podstawy obsługi zarówno starszej, jak i nowszych wersji tego programu oraz istotę tworzenia modeli prognostycznych takich jak modele regresji, drzew decyzyjnych, sieci neuronowych.

2. Zasady tworzenia kodu skoringowego służącego sporządzaniu prognoz

Skoring jest to proces oceny, w wyniku którego informacje dotyczące ocenianego obiektu są przekształcane na ciąg liczb, które po zsumowaniu stanowią miarę oceny obiektu. Ocena ta może być podstawą podejmowania decyzji dotyczących analizowanych obiektów [5].

Przykładem skoringu jest system oceny zdolności kredytowej (ryzyka kredytowego) potencjalnych kredytobiorców ułatwiający podejmowanie decyzji o przyznaniu kredytu przez pracowników banku.

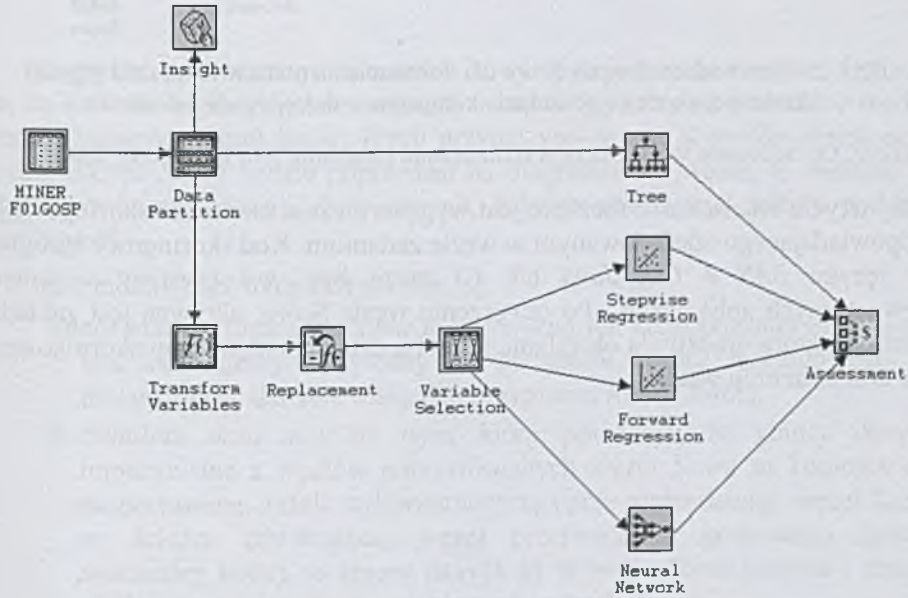
W ogólności ocena skoringowa ma na celu przewidywanie zachowania się badanego obiektu w przyszłości, co pozwala skutecznie prognozować np. przyszłe wyniki działalności firmy, zachowania się konsumenta albo sukcesy lub niepowodzenia w edukacji, sporcie.

Ocena skoringowa jest stosowana do nowo napływających danych, np. nowych klientów banku, nowych konsumentów, czy nowych uczniów.

Przed nadaniem oceny skoringowej należy opracować modele prognostyczne na podstawie danych z przeszłości, dla których znamy cechy i badane zachowanie się obiektów, a następnie należy wybrać najlepszy model, który posłuży do skoringu nowych obiektów.

Procesy tworzenia prognoz objaśnimy, aby osiągnąć większą jasność opisu, ilustrując nasze rozważania przykładem. Rozważmy przykładowy diagram zbudowany w *Enterprise Miner*, na którym umieścimy cztery modele prognostyczne (rys. 1). Są to dwa modele regresji: regresja krokowa (*Stepwise Regression*) i regresja „w przód" (*Forward Regression*), model drzewa decyzyjnego (*Tree*) oraz model sieci neuronowej (*Neural Network*). Na diagramie widoczne są także węzły, które wykorzystywaliśmy budując modele. Są to węzły: *Input Data Source* (pozwalający przypisać źródłowy zbiór danych, na diagramie *MINER.F01GOSP*), *Data Partition* (służący do podzielenia zbioru na podzbiory: treningowy, walidacyjny i testowy), *Insight* (pozwalający wyświetlić charakterystyki danych), *Transform Variables* (pozwalający utworzyć nowe zmienne na podstawie istniejących w zbiorze danych), *Replacement* (pozwalający rekodować dane i imputować wartości brakujacych danych), *Variable Selection* (służący do wyboru zmiennych biorących udział

w tworzeniu modelu prognostycznego). Zamieszczony jest także węzeł *Assessment*, umożliwiający nam porównywanie modeli i ich ocenę pod względem jakości dopasowania do danych. W wykorzystywanym przez nas przykładzie, używamy zbioru danych o nazwie *MINER.F01GOSP*, który składa się z obserwacji dotyczących gospodarstw domowych badanych przez GUS [2]. GUS prowadzi od wielu lat reprezentatywne badania budżetów polskich gospodarstw domowych na wylosowanej próbie, metodą rotacji miesięcznej (tzn. jedno gospodarstwo domowe jest badane przez jeden miesiąc). Badania te służą m.in. do ustalania wskaźników kosztów utrzymania oraz analizom naukowym. Rejestrowane są dane demograficzno-społeczne o gospodarstwie i jego członkach, o warunkach mieszkaniowych i wyposażeniu w dobra trwałe oraz o przychodach i wydatkach w badanym okresie. Jako zmienną objaśnianą (której wartości będziemy prognozować) w naszych modelach przyjęliśmy zmienną *posiadanie komputera z dostępem do Internetu* (o nazwie *U13*; 1 – symbolizuje posiadanie komputera, 0 - brak). Pozostałe zmienne są zmiennymi objaśniającymi. Zbiór liczy 31901 obserwacji, charakteryzowanych pod względem 448 zmiennych.

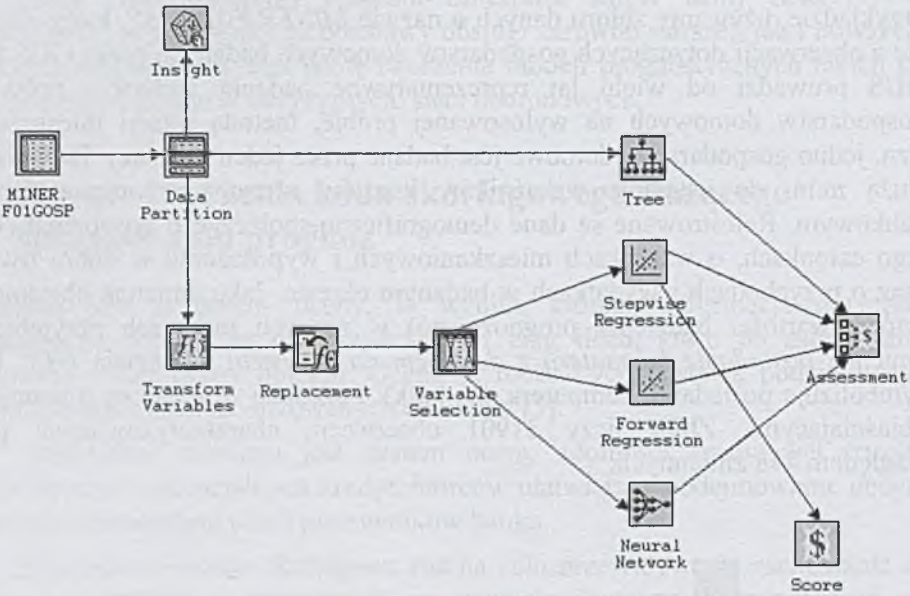


Rys. 1. Diagram analizy danych tworzenia i wykorzystania modeli prognozowania

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Przeprowadźmy ocenę skoringową wykorzystując jeden z utworzonych modeli. Niech to będzie przykładowo model regresji krokowej (*stepwise regression model*). Do przeprowadzenia oceny skoringowej służy węzeł *Score*, który należy

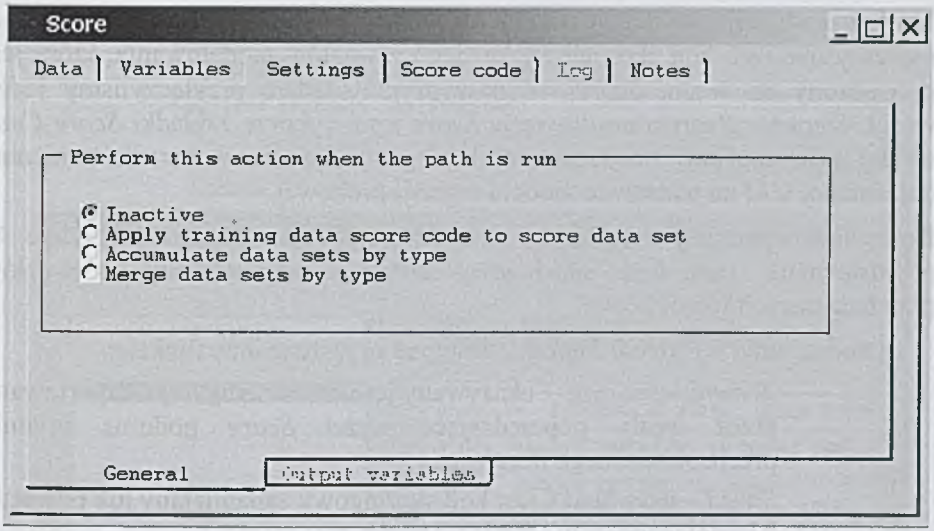
umieścić na diagramie, łącząc z nim węzeł *Stepwise Regression*. Diagram będzie wyglądał następująco (rys. 2).



Rys. 2. Wprowadzenie węzła *Score* dla dokonania na podstawie modelu regresji krokowej skoringu posiadania komputera z dostępem do Internetu

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Przy użyciu węzła *Score* możliwe jest wygenerowanie tzw. kodu skoringowego, odpowiadającego zdefiniowanym w węźle zadaniom. Kod skoringowy (program w języku *SAS 4 GL*, *Java* lub *C*) może być wykorzystany w innych zewnętrznych aplikacjach. Po otworzeniu węzła *Score*, aktywna jest zakładka *Settings*, która umożliwia określenie sposobu działania algorytmu skoringowego po uruchomieniu węzła.



Rys. 3. Okno *Score* z aktywną górną zakładką *Settings* i dolną zakładką *General* – ustalanie sposobu działania węzła *Score*

Źródło: *SAS Enterprise Miner*

Domyślnie, jak ilustruje to rysunek 3, wybrana jest opcja *Inactive*. Oznacza to, że kod skoringowy będzie tworzony i zapamiętywany na podstawie modeli poprzedzających węzeł *Score*. Jeżeli przyłączymy węzeł *Score* do określonego węzła tak, że *Score* będzie poprzedzał na diagramie ten węzeł, to zostanie do niego przekazany ostatni zbiór danych, którego obserwacje zostały poddane skoringowi.

Inne możliwości, to opcje (por. rys. 3):

Apply training data score code to score data set, która pozwala zastosować kod skoringowy otrzymany na podstawie zbioru treningowego do innego zbioru danych, mającego przypisaną rolę *SCORE*;

Accumulate data sets by type, która powoduje, że zbiory danych importowane z węzłów poprzedzających węzeł *Score* są kopiowane i eksportowane. Jeżeli wykorzystamy tę opcję umieszczając węzeł *Score* na ścieżce zawierającej węzeł przetwarzania grupowego (*group processing node*), to zbiory danych są w węźle *Score* łączone i zbiory wyjściowe z węzła *Score* są zbiorami po konkatenacji;

Merge data sets by type, po wyborze której, zbiory danych importowane z węzłów poprzedzających węzeł *Score* są łączone. Możemy zastosować tę opcję, aby połączyć w jeden zbiór, zbiory danych treningowych pochodzących z wielu węzłów modelowania w celu porównania wartości prognoz otrzymanych z różnych modeli.

Jeżeli pozostawimy wybraną opcję *Inactive*, to po wyborze zakładki *Score Code* możemy obejrzeć kod skoringowy każdego z węzłów modelowania, który jest przyłączony do węzła *Score*. W naszym przykładzie przyłączyliśmy jeden węzeł: *Stepwise Regression* do węzła *Score* i po wyborze zakładki *Score Code* węzła *Score* możemy obejrzeć kod skoringowy dla oceny wartości zmiennej objaśnianej *U13* na podstawie modelu regresji krokowej.

Domyślnie wybrana jest funkcja *Current Imports* (rys. 4), która powoduje, że przedstawiany jest kod skoringowy ostatnio importowany z węzłów poprzedzających węzeł *Score*.

Oprócz funkcji *Current Imports*, dostępne są jeszcze inne funkcje:

- *Accumulated runs* – ukazywany jest kod skoringowy eksportowany przez węzły poprzedzające węzeł *Score* podczas ostatnio przeprowadzonego trenowania,
- *Saved* – ukazywany jest kod skoringowy zapamiętany lub powstały z łączenia zbiorów,
- *All* – ukazywane są wszystkie utworzone kody skoringowe zarządzane przez węzeł Score.

Jeżeli kod skoringowy jest generowany z wielu modeli, to aby obejrzeć kod skoringowy wybranego modelu, musimy podwójnie kliknąć na nazwie modelu w oknie położonym w lewej dolnej części okna zakładki *Score code* (rys. 4). Kod skoringowy wyświetlany jest po prawej stronie. Kod skoringowy jest programem *SAS*, który możemy uruchomić korzystając z oprogramowania *Base SAS*.

Jeżeli zmienimy założenia modelu w węźle modelowania i uruchomimy przetwarzanie z węzła *Score*, to kod skoringowy powiązany z tym węzłem zostanie zaktualizowany.

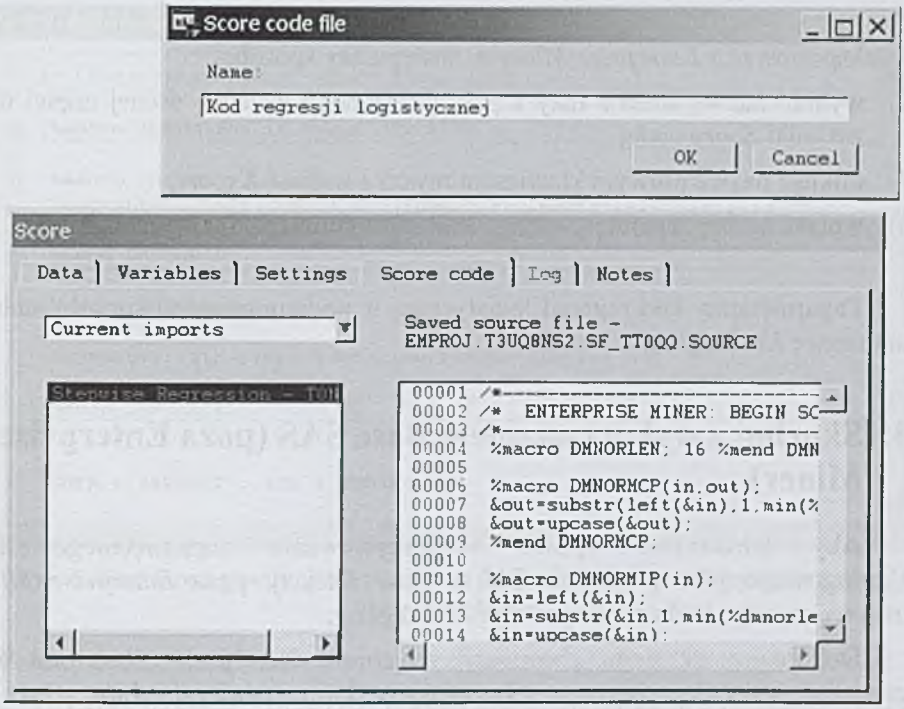
Jeżeli utworzyliśmy kod skoringowy, który uznamy za finalny i planujemy go stosować bez dalszych zmian, to możemy go zapamiętać w oddzielnym pliku dla potrzeb dalszego powtarzalnego wykorzystywania.

Aby zapamiętać kod skoringowy, należy wykonać następujące czynności:

1) wybrać z listy wyświetlanej w lewym dolnym polu zakładki *Score Code* nazwę modelu, którego kod skoringowy chcemy zapamiętać (w naszym przypadku jest to jeden model regresji logistycznej krokowej);

2) kliknąć wybrany model prawym klawiszem myszy i wybrać opcję *Save*;

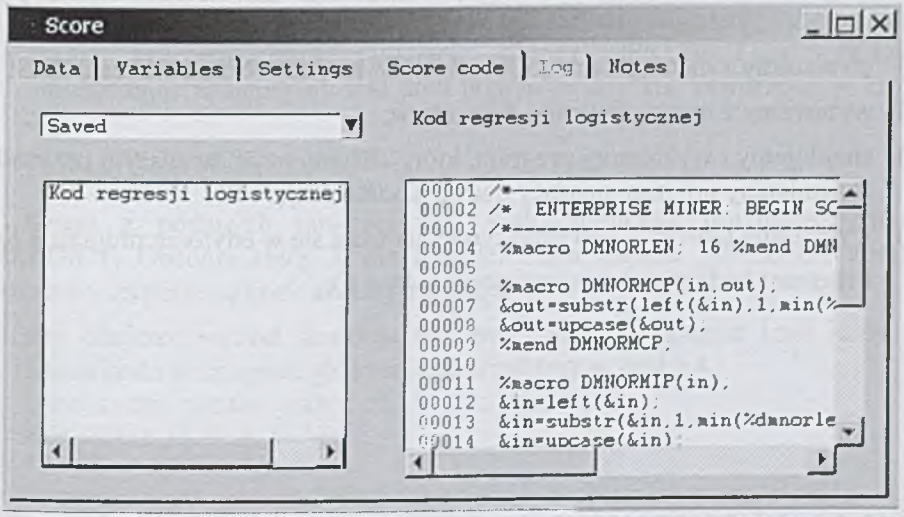
3) wprowadzić nazwę dla zapamiętywanego kodu w oknie dialogowym; w naszym przypadku wprowadźmy nazwę: *Kod regresji logistycznej* i wybrać przycisk *OK*.



Rys. 4. Zapisywanie kodu skoringowego

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Nazwa *Current Imports* zmieni się na *Saved*, sygnalizując, że kod został zapisany w *Enterprise Miner* (rys. 5).



Rys. 5. Zapisany kod skoringowy (dla naszego przykładu: *Kod regresji logistycznej*)

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Aby wykorzystać zapamiętany kod w *SAS*, poza *Enterprise Miner*, trzeba go wyeksportować z *Enterprise Miner* w następujący sposób:

1) wybrać nazwę kodu z listy z pola położonego w lewej dolnej części okna zakładki *Score code*;
2) kliknąć nazwę prawym klawiszem myszy i wybrać *Export*;
3) wpisać nazwę zapamiętywanego kodu (programu *SAS*) i wybrać *Zapisz*.

Zapamiętajmy kod regresji logistycznej w podany powyżej sposób, nadając mu nazwę *KodRegLog*.

3. Skoring z wykorzystaniem Base SAS (poza Enterprise Miner)

Aby zilustrować sposób wykorzystywania zapamiętanego kodu skoringowego jako programu *SAS* w *Base SAS*, tj. poza *Enterprise Miner*, użyjmy naszego kodu skoringowego *KodRegLog*.

Dokonajmy skoringu obserwacji ze zbioru *Nowygosp*. Zbiór ten musi zawierać wszystkie zmienne objaśniające, które zawierał zbiór użyty do utworzenia modelu i wygenerowania kodu skoringowego. Nie musi on zawierać (i na ogół nie zawiera) wartości zmiennej objaśnianej. W naszym przypadku zbiór zawiera obserwowane wartości zmiennej objaśnianej, dzięki czemu będziemy mogli porównać uzyskane oceny (prognozy) z rzeczywistymi wartościami zmiennej.

Aby dokonać oceny (skoringu) obserwacji ze zbioru wykonujemy następujące czynności:

1) wybieramy z menu głównego ***Window Editor***, aby rozpocząć sesję SAS;
2) wybieramy z menu głównego ***File Open***;
3) znajdujemy i wybieramy program, który chcemy użyć: w naszym przypadku zapamiętany program noszący nazwę *KodRegLog*;
4) wybieramy *Otwórz* – kod skoringowy ukazuje się w edytorze programu *SAS*.

Fragment kodu przedstawiono poniżej (rys. 6).

KodRegLog.sas

```
/*-----------------------------------------------------------------*/
/*   ENTERPRISE MINER: BEGIN SCORE CODE                            */
/*-----------------------------------------------------------------*/
%macro DMNORLEN; 16 %mend DMNORLEN;

%macro DMNORMCP(in,out);
&out=substr(left(&in),1,min(%dmnorlen,length(left(&in))));
&out=upcase(&out);
%mend DMNORMCP;

%macro DMNORMIP(in);
&in=left(&in);
&in=substr(&in,1,min(%dmnorlen,length(&in)));
&in=upcase(&in);
%mend DMNORMIP;

DATA &_PREDICT ; SET &_SCORE ;
```

Rys. 6. Kod skoringowy zapamiętany jako program *SAS* do wykorzystania poza *Enterprise Miner*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Zbiór danych _PREDICT zawiera prognozowane wartości. Zbiór danych reprezentowany przez _SCORE jest zbiorem danych, dla którego chcemy przeprowadzić ocenę (prognozę, skoring).

Ponieważ powyżej wymienione zbiory danych są wykorzystywane w macro (są poprzedzone znakiem &), zbiory danych muszą zostać zainicjowane.

5) przeprowadzamy skoring zbioru NOWYGOSP z biblioteki MINER. Aby to zrobić przeprowadzamy inicjację zbioru _PREDICT i _SCORE umieszczając następujący kod jako poprzedzający kod otworzony w kroku 4.:

```
%let _SCORE=MINER.NOWYGOSP;
%let _PREDICT=X;
```

Druga z podanych powyżej linii kodu zainicjuje utworzenie zbioru _PREDICT. Obecnie zbiór X nie istnieje. Zbiór danych _PREDICT zostaje utworzony za pomocą kodu skoringowego.

6) aby obejrzeć wyniki skoringu musimy dodać następujące linie kodu na końcu kodu skoringowego, który otworzyliśmy w kroku 4.:

```
PROC PRINT DATA=&_PREDICT;
VAR U13 P_U131 P_U130;
Run;
```

Przedstawiony powyżej kod spowoduje wypisanie wartości U13 oraz P_U131 (prognozowane prawdopodobieństwo U13=1) i P_U130 (prognozowane prawdopodobieństwo U13=0).

Możemy posortować obserwacje według wartości wybranej zmiennej, np. P_U131, P_U130. Aby to zrobić, musimy dołączyć procedurę sortowania przed podaną powyżej procedurą PROC PRINT. Dołączmy przykładowo procedurę sortującą obserwacje malejąco według wartości P_U131. Wówczas jako pierwsze zostaną przedstawione obserwacje – gospodarstwa domowe, które według przewidywań na podstawie modelu są gospodarstwami domowymi posiadającymi komputer z dostępem do Internetu.

Procedura sortująca będzie wyglądać następująco:

```
PROC SORT DATA=&_PREDICT;
BY DESCENDING U131;
RUN;
```

7) uruchamiamy przetwarzanie programu wybierając z menu głównego **Run Submit** lub ikonę *Submit* z paska narzędziowego. Możemy obejrzeć wyniki i porównać rzeczywistą wartość zmiennej U13 z uzyskanymi prognozami. Fragment wyników jest podany poniżej na rysunku 7.

Obs	U13	P_U131	P_U130
1	1	0.62522	0.37478
2	1	0.17892	0.82108
3	1	0.29347	0.70653
4	1	0.16354	0.83646
5	1	0.58604	0.41396
6	1	0.74083	0.25917
7	1	0.40415	0.59585
8	1	0.03468	0.96532
9	1	0.08773	0.91227
10	1	0.44378	0.55622
11	1	0.12047	0.87953
12	1	0.13067	0.86933
13	1	0.47690	0.52310
14	1	0.26477	0.73523
15	1	0.05132	0.94868
16	1	0.55626	0.44374
17	1	0.80659	0.19341
18	1	0.55313	0.44687
19	1	0.31788	0.68212
20	1	0.64301	0.35699
21	1	0.30533	0.69467
22	1	0.04368	0.95632
23	1	0.78994	0.21006
24	1	0.59726	0.40274
25	1	0.57758	0.42242
26	1	0.12539	0.87461
27	1	0.50376	0.49624
28	1	0.17117	0.82883
29	1	0.05591	0.94409
30	1	0.05314	0.94686
31	1	0.49175	0.50825
32	1	0.18201	0.81799
33	1	0.21457	0.78543
34	1	0.05413	0.94587
35	1	0.18352	0.81648
36	1	0.43249	0.56751
37	1	0.13622	0.86378
38	1	0.07566	0.92434
39	1	0.07703	0.92297
40	1	0.01281	0.98719
41	1	0.61042	0.38958
42	1	0.20167	0.79833
43	1	0.76353	0.23647
44	1	0.14777	0.85223
45	1	0.34041	0.65959
46	1	0.08869	0.91131
47	1	0.01754	0.98246
48	1	0.42700	0.57300
49	1	0.66089	0.33911
50	1	0.54166	0.45834
51	1	0.12008	0.87992
52	1	0.06031	0.93969
53	1	0.76301	0.23699
54	1	0.11148	0.88852

Obs	U13	P_U131	P_U130
31807	0	0.001011	0.99899
31808	0	0.008554	0.99145
31809	0	0.039160	0.96084
31810	0	0.020794	0.97921
31811	0	0.003212	0.99679
31812	0	0.026600	0.97340
31813	0	0.001467	0.99853
31814	0	0.006035	0.99396
31815	1	0.012360	0.98764
31816	0	0.003279	0.99672
31817	0	0.009087	0.99091
31818	0	0.003815	0.99619
31819	0	0.008371	0.99163
31820	0	0.014467	0.98553
31821	0	0.003036	0.99696
31822	0	0.017060	0.98294
31823	0	0.020124	0.97988
31824	0	0.001482	0.99852
31825	0	0.002430	0.99757
31826	0	0.019991	0.98001
31827	0	0.012582	0.98742
31828	0	0.000868	0.99913
31829	0	0.003497	0.99650
31830	0	0.001851	0.99815
31831	0	0.004503	0.99550
31832	0	0.003904	0.99610
31833	0	0.001431	0.99857
31834	0	0.007312	0.99269
31835	0	0.006518	0.99348
31836	0	0.028532	0.97147
31837	0	0.019221	0.98078
31838	0	0.001223	0.99878
31839	0	0.002325	0.99767
31840	0	0.004732	0.99527
31841	0	0.007952	0.99205
31842	0	0.079892	0.92011
31843	0	0.004717	0.99528
31844	0	0.001885	0.99811
31845	0	0.005238	0.99476
31846	0	0.002792	0.99721
31847	0	0.010618	0.98938

Rys. 7. Wyniki uzyskane po uruchomieniu kodu skoringowego zapamiętanego jako program *SAS* poza *Enterprise Miner*

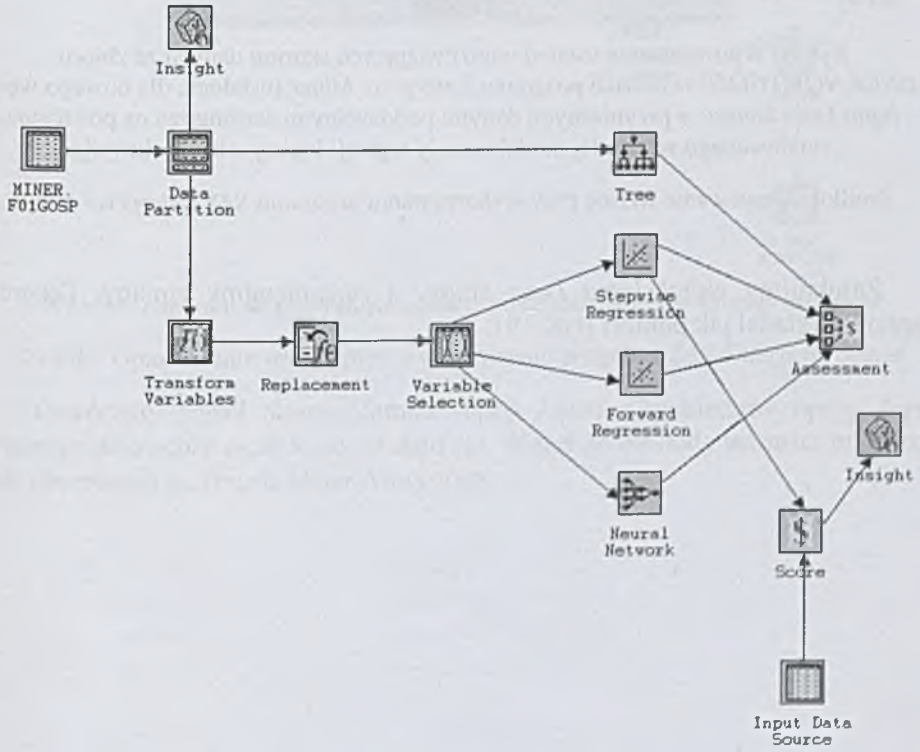
Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

4. Przeprowadzanie oceny skoringowej w Enterprise Miner

Przeprowadźmy skoring obserwacji z tego samego zbioru *Miner.Nowygosp*, jak w przypadku dokonywania skoringu w *Base SAS*.

Aby przeprowadzić skoring w *Enterprise Miner* należy dodać do diagramu, nowy węzeł *Input Data Source* umożliwiający dostęp do obserwacji ze zbioru *Miner.Nowygosp* i węzeł ten połączyć z węzłem *Score*. Dodajmy też węzeł *Insight* i połączmy węzeł *Score* z węzłem *Insight*.

Nasz diagram będzie teraz wyglądał tak jak ilustruje to poniższy rysunek (rys. 8):

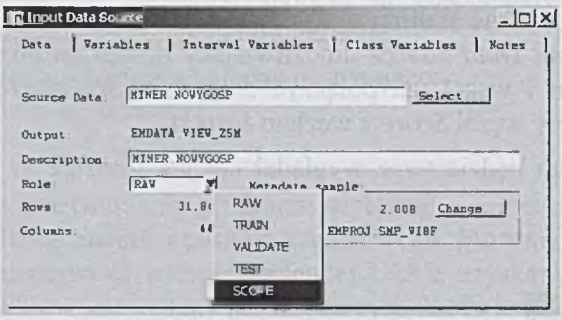


Rys. 8. Budowa diagramu do przeprowadzenia oceny skoringowej w *Enterprise Miner*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Przypiszmy zbiór danych *Miner.Nowygosp* z obserwacjami – gospodarstwami domowymi, dla których chcemy przewidzieć, czy mają

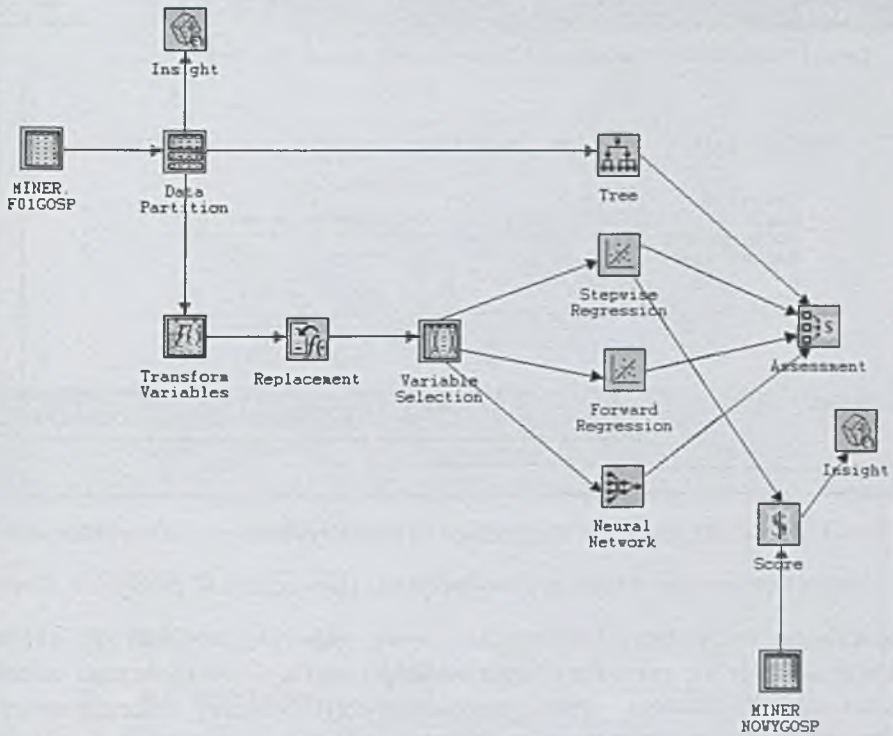
komputer z dostępem do Internetu, czy też nie, do wprowadzonego węzła *Input Data Source*. Zbiór danych, dla którego będzie przeprowadzany skoring musi mieć określoną rolę *Score*. Dlatego – jak zilustrowano na rysunku 9 zmieniamy rolę zbioru danych z *RAW* na *SCORE*.



Rys. 9. Wprowadzanie ustaleń umożliwiających skoring danych ze zbioru *MINER.NOWYGOSP* w ramach programu *Enterprise Miner* (ustalenia dla nowego węzła *Input Data Source* z przypisanymi danymi poddawanymi skoringowi na podstawie zbudowanego wcześniej modelu oceny regresji logistycznej krokowej)

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

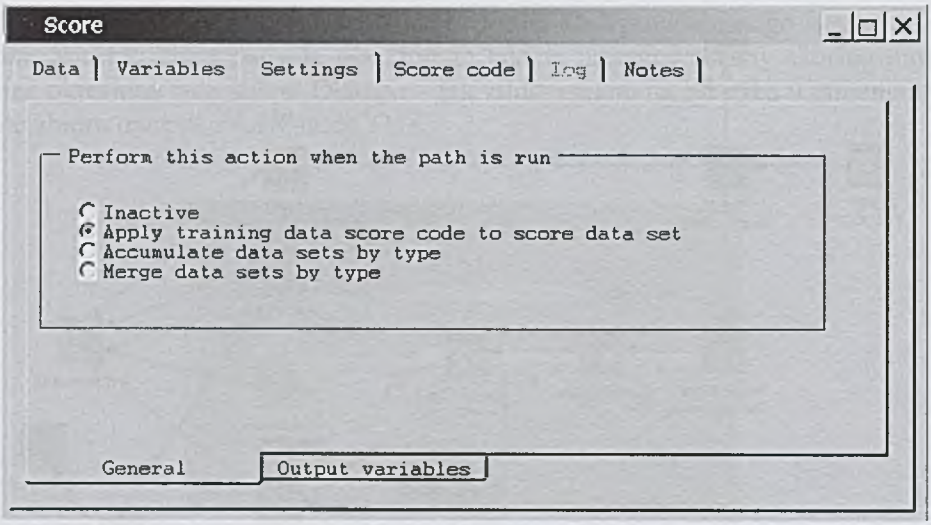
Zamknijmy węzeł *Input Data Source* i zapamiętajmy zmiany. Diagram będzie wyglądał jak poniżej (rys. 10):



Rys. 10. Diagram do przeprowadzenia oceny skoringowej w *Enterprise Miner*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Otwórzmy węzeł *Score*. Zamiast opcji *Inactive* wybierzmy opcję *Apply training data score code to score data set*. Węzeł *Score* doda wartości predykcji dla obserwacji ze zbioru *Miner.Nowygosp*.



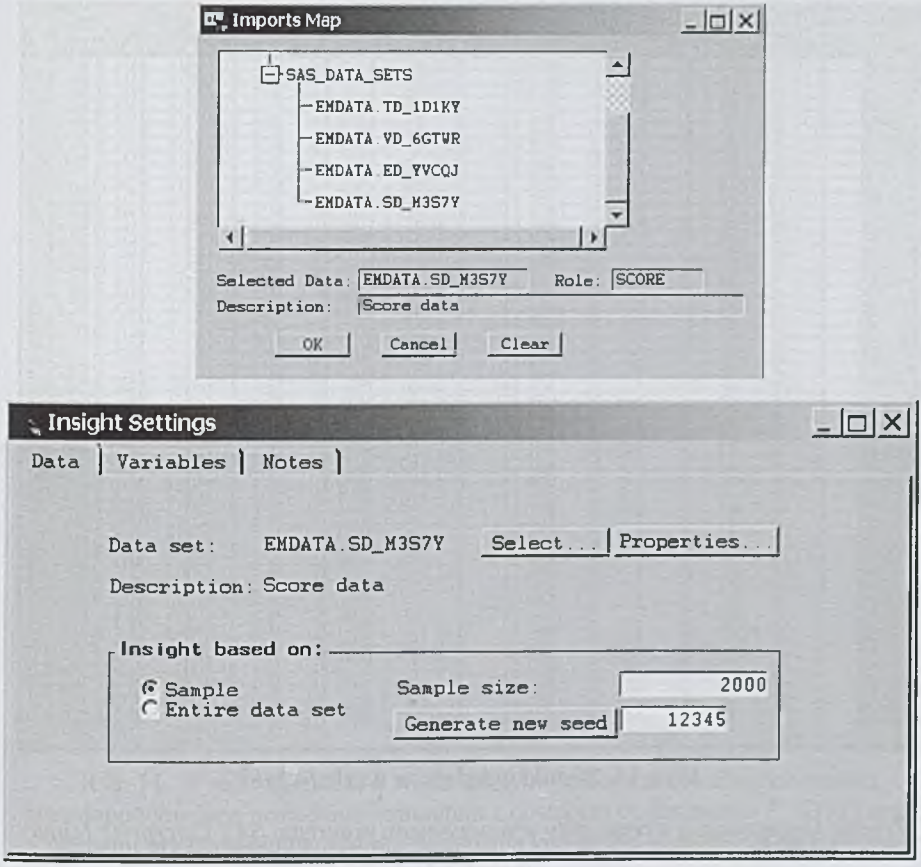
Rys. 11. Ustalenia *Settings* w węźle *Score* na potrzeby skoringu w *Enterprise Miner*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Po wyborze opcji *Apply training data score code to score data set*, aktywna (dostępna) staje się zakładka *Output variables* węzła *Score* (położona w dolnej części zakładki *Settings* – por. rysunek powyżej). Możemy wskazać zmienne, których nie chcemy umieszczać w zbiorze wyjściowym po dokonaniu skoringu. W naszym przykładzie pozostawmy wszystkie zmienne. Zamknijmy węzeł *Score*, zapamiętując wprowadzone zmiany. Pod nazwą węzła *Score* pojawi się napis *[Apply]*, wskazujący na wybrany przez nas sposób wykorzystania węzła.

Otwórzmy węzeł *Insight*. Wybierzmy przycisk *Select* z zakładki *Data*, aby wybrać zbiór danych, który będzie powiązany z ocenianym zbiorem danych. Nazwa tego zbioru danych ma prefix *SD* i ciąg losowo wybranych znaków alfanumerycznych.

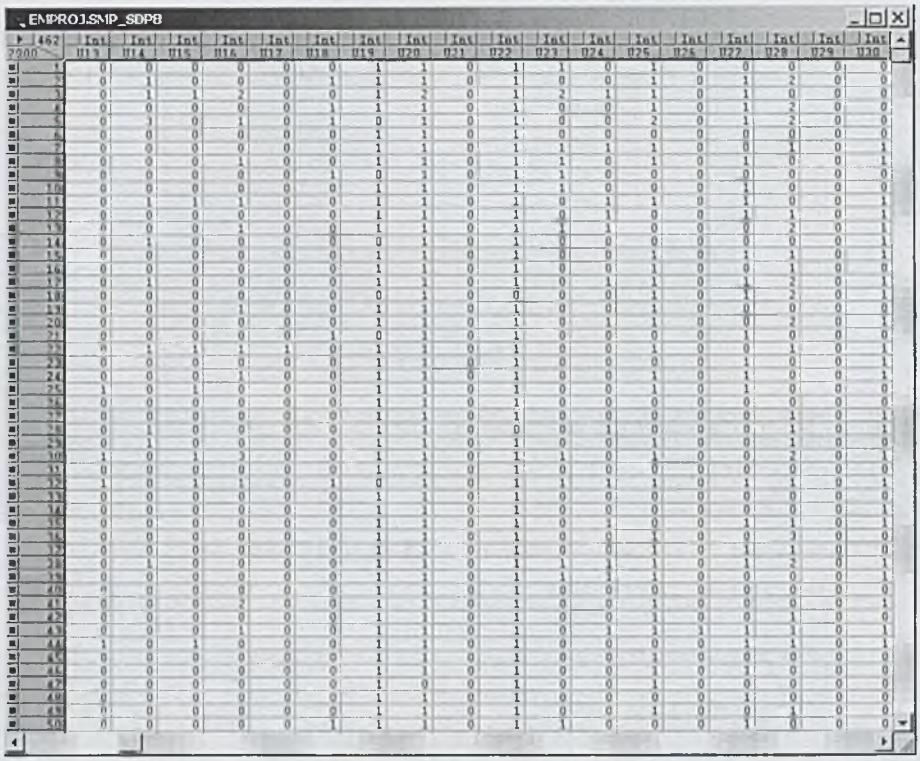
W naszym przypadku ma on nazwę *SD_M3S7Y*, rolę *SCORE* i opis *Score data*. Wybierzmy *OK*, aby powrócić do zakładki *Data* węzła *Insight*. Zamknijmy węzeł *Insight*, zapisując wprowadzone zmiany. Jeżeli pozostawiliśmy wybraną opcję *Insight based on Sample* po uruchomieniu węzła otrzymamy wyniki na podstawie próby danych.



Rys. 12. Ustalenia w węźle *Insight* dla analizy danych po skoringu w *Enterprise Miner*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Uruchamiamy węzeł *Insight* i obejrzyjmy otrzymane wyniki.

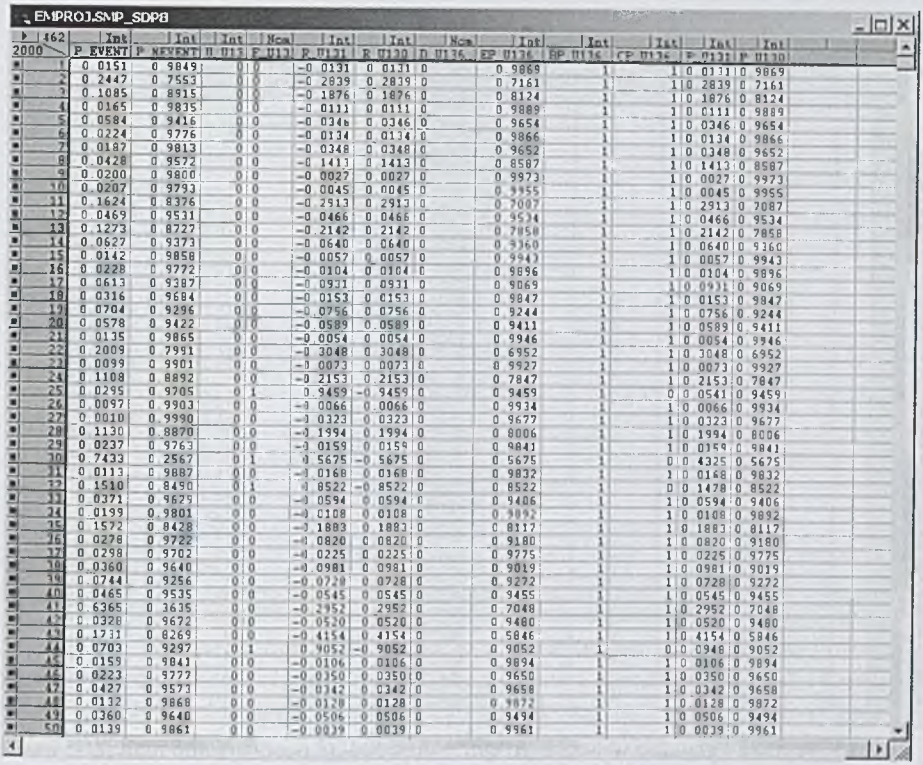


462	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
2000	U13	U14	U15	U16	U17	U18	U19	U20	U21	U22	U23	U24	U25	U26	U27	U28	U29	U30	
1	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	0	0	0	
2	0	1	0	0	0	1	1	1	0	1	0	0	1	0	1	2	0	0	
3	0	1	1	2	0	0	1	2	0	1	2	1	1	0	1	0	0	0	
4	0	0	0	0	0	1	1	1	0	1	0	0	1	0	1	2	0	0	
5	0	1	0	1	0	1	0	1	0	1	0	0	2	0	1	2	0	0	
6	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	1	1	0	1	1	1	1	0	0	1	0	1	
8	0	0	0	1	0	0	1	1	0	1	1	0	1	0	1	0	0	1	
9	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	
10	0	0	0	0	0	0	1	1	0	1	1	0	0	0	1	0	0	0	
11	0	1	1	1	0	0	1	1	0	1	0	1	1	0	1	0	0	1	
12	0	0	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	
13	0	0	0	1	0	0	1	1	0	1	0	1	0	0	0	2	0	1	
14	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	
15	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	1	
16	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	
17	0	1	0	0	0	0	1	1	0	1	0	1	1	0	1	2	0	1	
18	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	2	0	1	
19	0	0	0	1	0	0	1	1	0	1	0	0	1	0	0	0	0	0	
20	0	0	0	0	0	0	1	1	0	1	0	1	1	0	0	2	0	1	
21	0	0	0	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	
22	0	1	1	1	1	0	1	1	0	1	0	0	1	0	0	1	0	1	
23	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	1	
24	0	0	0	1	0	0	1	1	0	1	0	0	1	0	1	0	0	1	
25	1	0	1	0	0	0	1	1	0	1	0	0	1	0	1	0	0	0	
26	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
27	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	
28	0	1	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	1	
29	0	1	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	
30	1	0	1	3	0	0	1	1	0	1	1	0	1	0	0	2	0	1	
31	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
32	1	0	1	1	0	1	0	1	0	1	1	1	1	0	1	1	0	1	
33	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
34	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	
35	0	0	0	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	
36	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	
37	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	0	
38	0	1	0	0	0	0	1	1	0	1	1	1	1	0	1	2	0	1	
39	0	0	0	0	0	0	1	1	0	1	1	1	1	0	0	0	0	1	
40	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	
41	0	0	0	2	0	0	1	1	0	1	0	0	1	0	0	0	0	1	
42	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	
43	0	0	0	1	0	0	1	1	0	1	0	1	1	1	1	1	0	1	
44	1	0	1	0	0	0	1	1	0	1	0	0	0	0	1	1	0	1	
45	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	0	0	
46	0	0	0	0	0	0	1	1	0	1	0	0	1	0	1	0	0	0	
47	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	
48	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	
49	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	
50	0	0	0	0	0	1	1	1	0	1	1	0	0	0	1	0	0	0	

Rys. 13. Wyniki oglądane w węźle *Insight*

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Otrzymany zbiór danych po przeprowadzeniu skoringu ma o 14 zmiennych więcej niż przed przeprowadzeniem skoringu. Dodane zostały m.in. zmienne przedstawiające prognozę dla zmiennej *U13*. Są to dwie ostatnie kolumny o nazwach *P_U131* i *P_U130*.



	P_EVENT	P_NEVENT	U13	F_U13	P_U131	R_U131	R_U130	U136	P_U130	EP_U136	CP_U136	P_U131	P_U130
1	0.0151	0.9849	0	0	-0.0131	0.0131	0	0	0.9869	1	1	0.0131	0.9869
2	0.2447	0.7553	0	0	-0.2839	0.2839	0	0	0.7161	1	1	0.2839	0.7161
3	0.1085	0.8915	0	0	-0.1876	0.1876	0	0	0.8124	1	1	0.1876	0.8124
4	0.0165	0.9835	0	0	-0.0111	0.0111	0	0	0.9889	1	1	0.0111	0.9889
5	0.0584	0.9416	0	0	-0.0346	0.0346	0	0	0.9654	1	1	0.0346	0.9654
6	0.0224	0.9776	0	0	-0.0134	0.0134	0	0	0.9866	1	1	0.0134	0.9866
7	0.0187	0.9813	0	0	-0.0348	0.0348	0	0	0.9652	1	1	0.0348	0.9652
8	0.0428	0.9572	0	0	-0.1413	0.1413	0	0	0.8587	1	1	0.1413	0.8587
9	0.0200	0.9800	0	0	-0.0027	0.0027	0	0	0.9973	1	1	0.0027	0.9973
10	0.0207	0.9793	0	0	-0.0045	0.0045	0	0	0.9955	1	1	0.0045	0.9955
11	0.1624	0.8376	0	0	-0.2913	0.2913	0	0	0.7087	1	1	0.2913	0.7087
12	0.0469	0.9531	0	0	-0.0466	0.0466	0	0	0.9534	1	1	0.0466	0.9534
13	0.1273	0.8727	0	0	-0.2142	0.2142	0	0	0.7858	1	1	0.2142	0.7858
14	0.0627	0.9373	0	0	-0.0640	0.0640	0	0	0.9360	1	1	0.0640	0.9360
15	0.0142	0.9858	0	0	-0.0057	0.0057	0	0	0.9943	1	1	0.0057	0.9943
16	0.0228	0.9772	0	0	-0.0104	0.0104	0	0	0.9896	1	1	0.0104	0.9896
17	0.0613	0.9387	0	0	-0.0931	0.0931	0	0	0.9069	1	1	0.0931	0.9069
18	0.0316	0.9684	0	0	-0.0153	0.0153	0	0	0.9847	1	1	0.0153	0.9847
19	0.0704	0.9296	0	0	-0.0756	0.0756	0	0	0.9244	1	1	0.0756	0.9244
20	0.0578	0.9422	0	0	-0.0589	0.0589	0	0	0.9411	1	1	0.0589	0.9411
21	0.0135	0.9865	0	0	-0.0054	0.0054	0	0	0.9946	1	1	0.0054	0.9946
22	0.2009	0.7991	0	0	-0.3048	0.3048	0	0	0.6952	1	1	0.3048	0.6952
23	0.0099	0.9901	0	0	-0.0073	0.0073	0	0	0.9927	1	1	0.0073	0.9927
24	0.1108	0.8892	0	0	-0.2153	0.2153	0	0	0.7847	1	1	0.2153	0.7847
25	0.0295	0.9705	0	1	0.9459	-0.9459	0	0	0.9459	1	0	0.0541	0.9459
26	0.0097	0.9903	0	0	-0.0066	0.0066	0	0	0.9934	1	1	0.0066	0.9934
27	0.0010	0.9990	0	0	-0.0323	0.0323	0	0	0.9677	1	1	0.0323	0.9677
28	0.1130	0.8870	0	0	-0.1994	0.1994	0	0	0.8006	1	1	0.1994	0.8006
29	0.0237	0.9763	0	0	-0.0159	0.0159	0	0	0.9841	1	1	0.0159	0.9841
30	0.7433	0.2567	0	1	0.5675	-0.5675	0	0	0.5675	1	0	0.4325	0.5675
31	0.0113	0.9887	0	0	-0.0168	0.0168	0	0	0.9832	1	1	0.0168	0.9832
32	0.1510	0.8490	0	1	0.8522	-0.8522	0	0	0.8522	1	0	0.1478	0.8522
33	0.0371	0.9629	0	0	-0.0594	0.0594	0	0	0.9406	1	1	0.0594	0.9406
34	0.0199	0.9801	0	0	-0.0108	0.0108	0	0	0.9892	1	1	0.0108	0.9892
35	0.1572	0.8428	0	0	-0.1883	0.1883	0	0	0.8117	1	1	0.1883	0.8117
36	0.0278	0.9722	0	0	-0.0820	0.0820	0	0	0.9180	1	1	0.0820	0.9180
37	0.0298	0.9702	0	0	-0.0225	0.0225	0	0	0.9775	1	1	0.0225	0.9775
38	0.0360	0.9640	0	0	-0.0981	0.0981	0	0	0.9019	1	1	0.0981	0.9019
39	0.0744	0.9256	0	0	-0.0728	0.0728	0	0	0.9272	1	1	0.0728	0.9272
40	0.0465	0.9535	0	0	-0.0545	0.0545	0	0	0.9455	1	1	0.0545	0.9455
41	0.6365	0.3635	0	0	-0.2952	0.2952	0	0	0.7048	1	1	0.2952	0.7048
42	0.0328	0.9672	0	0	-0.0520	0.0520	0	0	0.9480	1	1	0.0520	0.9480
43	0.1731	0.8269	0	0	-0.4154	0.4154	0	0	0.5846	1	1	0.4154	0.5846
44	0.0703	0.9297	0	1	0.9052	-0.9052	0	0	0.9052	1	0	0.0948	0.9052
45	0.0159	0.9841	0	0	-0.0106	0.0106	0	0	0.9894	1	1	0.0106	0.9894
46	0.0223	0.9777	0	0	-0.0350	0.0350	0	0	0.9650	1	1	0.0350	0.9650
47	0.0427	0.9573	0	0	-0.0342	0.0342	0	0	0.9658	1	1	0.0342	0.9658
48	0.0132	0.9868	0	0	-0.0128	0.0128	0	0	0.9872	1	1	0.0128	0.9872
49	0.0360	0.9640	0	0	-0.0506	0.0506	0	0	0.9494	1	1	0.0506	0.9494
50	0.0139	0.9861	0	0	-0.0039	0.0039	0	0	0.9961	1	1	0.0039	0.9961

Rys. 14. Wyniki oglądane w węźle *Insight* - widoczne kolumny z ocenami prawdopodobieństw posiadania komputera z dostępem do Internetu (*P_U131*) oraz z ocenami prawdopodobieństw nie posiadania komputera z dostępem do Internetu (*P_U130*)

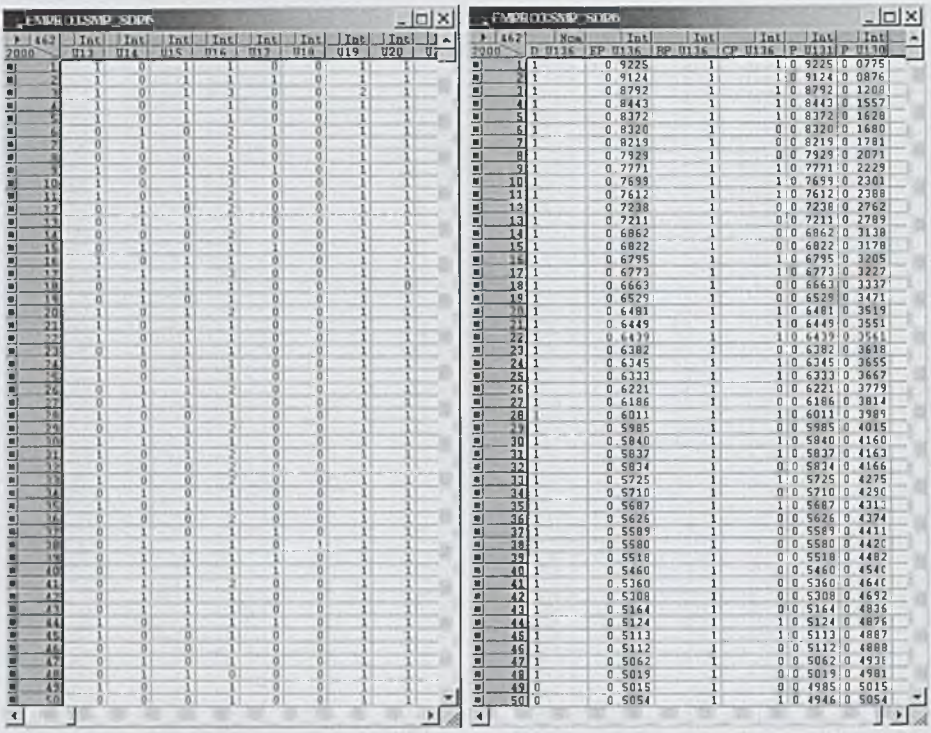
Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Można ograniczyć liczbę zmiennych, wskazując w węźle *Score* (zakładka *Output Variables*), zmienne, których nie chcemy zamieszczać w zbiorze danych otrzymanym po dokonaniu skoringu.

Otwarty zbiór danych ma inną nazwę niż nazwa wybranego przez nas zbioru w węźle *Insight*. *Insight* przedstawia próbę złożoną z 2000 obserwacji. *Prefix SMP_* w nazwie zbioru wskazuje, że jest to próba ze zbioru danych.

Możemy posortować zbiór według wartości zmiennej *U13*, wybrać gospodarstwa posiadające komputer z dostępem do Internetu, tj. takie dla których wartość zmiennej *U13* wynosi 1 i przeanalizować prognozy dotyczące posiadania komputera z Internetem dla tych gospodarstw.

Jeżeli posortujemy obserwacje malejąco według P_U131, to ułatwi nam znalezienie obserwacji, dla których prognoza o posiadaniu komputera z dostępem do Internetu była trafna, jak ilustruje to rysunek poniżej (rys. 15).



Rys. 15. Wyniki oglądane w węźle *Insight* – obserwacje posortowane malejąco według oceny prawdopodobieństwa posiadania komputera z dostępem do Internetu

Źródło: Opracowanie własne przy wykorzystaniu programu *SAS Enterprise Miner*

Odczytywanie wyników ułatwia ograniczenie liczby zmiennych (kolumn) w tablicy otrzymywanej z węzła *Insight*. Możemy to zrobić klikając na obszarze tablicy prawym klawiszem myszy i wybierając opcję *Extract*.

Rysunek przedstawiony poniżej ilustruje sytuację, gdy ograniczyliśmy wyniki w tablicy do przedstawiających numer gospodarstwa domowego (*NR*), wartości zmiennej *U13* oraz otrzymanego z modelu regresji logistycznej prawdopodobieństwa, że dane gospodarstwo domowe posiada komputer z dostępem do Internetu (*P_U131*) lub go nie posiada (*P_U130*). Dodatkowo dane zostały posortowane według malejących wartości prawdopodobieństwa posiadania komputera z dostępem do Internetu.

EMPROJ.SMP_SDG1

2000	4 NR	Nom U13	Int P U131	Int P U130	Int
1	11740064	1	0.9225	0.0775	
2	10342091	1	0.9124	0.0876	
3	10259064	1	0.8792	0.1208	
4	10079021	1	0.8443	0.1557	
5	10151011	1	0.8372	0.1628	
6	10142104	0	0.8320	0.1680	
7	20062054	0	0.8219	0.1781	
8	21136051	0	0.7929	0.2071	
9	20053064	1	0.7771	0.2229	
10	10204071	1	0.7699	0.2301	
11	10327031	1	0.7612	0.2388	
12	21220091	0	0.7238	0.2762	
13	21028061	0	0.7211	0.2789	
14	20054041	0	0.6862	0.3138	
15	10099124	0	0.6822	0.3178	
16	11304041	1	0.6795	0.3205	
17	10283114	1	0.6773	0.3227	
18	21034011	0	0.6663	0.3337	
19	10285051	0	0.6529	0.3471	
20	10327064	1	0.6481	0.3519	

Rys. 16. Ułatwienie oglądania wyników w węźle Insight – ograniczono liczbę wyświetlanych zmiennych do 4: numeru gospodarstwa domowego, wartości zmiennej U13, prawdopodobieństwa, że gospodarstwo ma komputer z dostępem do Internetu i prawdopodobieństwa, że komputera takiego nie posiada (4 kolumny z danymi); dodatkowo obserwacje zostały posortowane malejąco według oceny prawdopodobieństwa posiadania komputera z dostępem do Internetu

Źródło: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner

5. Zakończenie

W rozdziale staraliśmy się ukazać, jak po zbudowaniu modelu prognostycznego eksploracji danych *Data Mining* można go zastosować do przeprowadzenia prognozy wybranej zmiennej. W naszym przypadku było to przewidywanie za pomocą modeli, czy gospodarstwo domowe o określonych cechach (zmienne niezależne) posiada komputer z dostępem do Internetu (zmienna zależna).

W celu przewidywania wartości zmiennej zależnej na podstawie danych o wartościach zmiennych niezależnych posłużyliśmy się programem *Enterprise Miner* firmy *SAS*. Za jego pomocą nie tylko przeprowadziliśmy prognozy, ale także zbudowaliśmy wykorzystywane do tego prognozowania modele.

W programie *Enterprise Miner* na potrzeby prognozowania umieszczono specjalne narzędzie – węzeł *Score*. Węzeł ten umożliwia nie tylko uzyskanie prognozy na podstawie modeli, ale także zapamiętanie (zapis) kodu służącego tworzeniu prognoz (tzw. kodu skoringowego), który można wykorzystać poza

Enterprise Miner w *Base SAS* oraz łączenie kodu skoringowego z różnych modeli.

Podsumowując, jak przedstawiliśmy to w rozdziale, możliwe jest na potrzeby otrzymania prognozy, wykorzystanie zapamiętanego kodu skoringowego jako programu *SAS* w *Base SAS* poza *Enterprise Miner* lub przeprowadzenie prognozy w *Enterprise Miner*, posługując się „bezpośrednio" kodem wygenerowanym przez narzędzie *Score*.

Naszym podstawowym celem było pokazanie, przede wszystkim z uwagi na brak dokładniejszych, precyzyjniejszych, pozbawionych niepotrzebnych zawiłości opisów, jak w prosty sposób skonstruować narzędzie do oceny skoringowej obiektów za pomocą programu *SAS Enterprise Miner*.

LITERATURA

1. Applied Analytics Using SAS Enterprise Miner 5.3 Course Notes, SAS Institute Inc., Cary, NC, USA 2008.
2. Budżety gospodarstw domowych w 2007 r., Informacje i opracowania statystyczne, GUS, Warszawa 2008.
3. Matignon R., Data Mining Using SAS Enterprise Miner, John Wiley & Sons, New Jersey 2007.
4. Reference Help – Enterprise Miner 5.3., SAS Institute Inc., Cary, NC, USA 2007.
5. Supera J.: Skoring – podstawowe pojęcia i stosowana terminologia, materiały konferencji „Skoring w teorii i praktyce", Kazimierz Dolny 1998.

Rozdział 24

Harmonogramowanie realizacji programów w wieloprocesorowym systemie informatycznym

Zbigniew Buchalski
Politechnika Wrocławska
zbigniew.buchalski@pwr.wroc.pli

Streszczenie

W rozdziale przedstawiono zagadnienie czasowo-optymalnego przydziału n programów niezależnych niepodzielnych i stron pamięci operacyjnej do procesorów pracujących równolegle. Czas wykonania i-tego programu na k-tym procesorze określony jest przez pewną funkcję zależną od liczby stron pamięci operacyjnej przydzielonych k-temu procesorowi oraz od parametrów charakteryzujących wykonywany program. Dla zadanej funkcji czasu realizacji programów zaproponowano pewien algorytm heurystyczny wyznaczający czasowo-optymalne szeregowanie programów i przydział stron pamięci operacyjnej do procesorów w wieloprocesorowym systemie informatycznym. Przedstawiono wyniki eksperymentów obliczeniowych przeprowadzonych na tym algorytmie.

1. Wstęp

Intensywny rozwój równoległych systemów przetwarzania informacji pociągnął za sobą wzrost zainteresowania problematyką czasowo-optymalnego szeregowania zadań i rozdziału zasobów [1, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14]. Prezentowana praca bazuje na wynikach tych badań i jest kontynuacją wcześniejszych prac autora [5, 6, 7].

Problem szeregowania zadań na maszynach równoległych z równoczesnym rozdzia-łem zasobów bardzo często spotykany jest w różnego rodzaju złożonych procesach produkcyjnych. W wieloprocesorowych systemach komputerowych

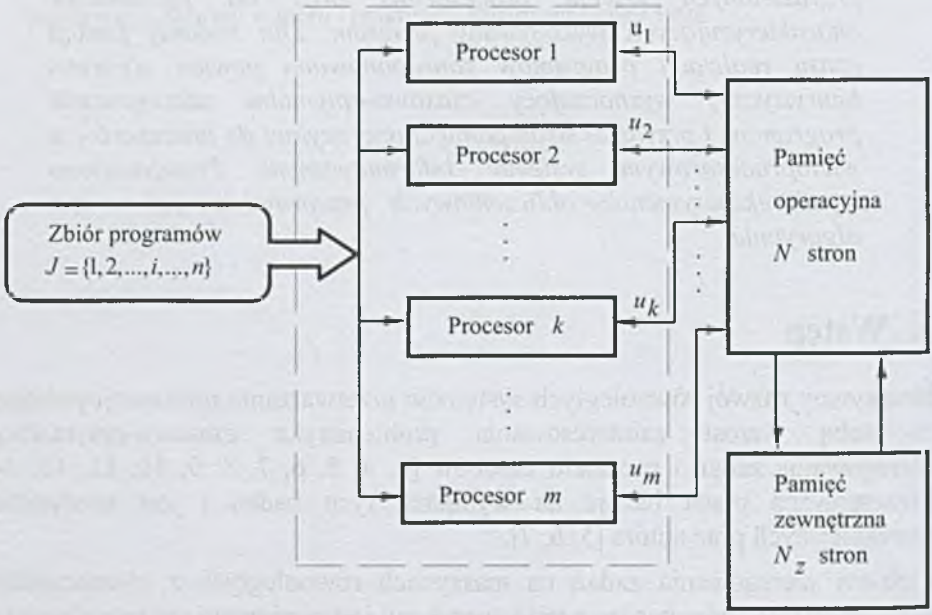
spotykamy się z szeregowaniem programów na procesorach oraz przydziałem zasobów w postaci stron pamięci operacyjnej do procesorów [2, 3, 5, 12].

Wyniki teorii złożoności obliczeniowej oraz rozmiar problemów praktycznych w sposób jednoznaczny eliminują z rozważań algorytmy dokładne, pozostawiając do zastosowania praktycznego jedynie algorytmy heurystyczne umożliwiające rozwiązanie postawionych problemów w krótkim czasie z zadowalającą dokładnością. Badania nad algorytmami heurystycznymi, dostarczającymi rozwiązań zagadnień, w których zastosowanie metod dokładnych jest nieefektywne lub wręcz niemożliwe, stanowią jedną z najszybciej rozwijających się gałęzi nauki.

W niniejszym rozdziale przedstawiono pewien algorytm heurystyczny wyznaczający czasowo-optymalne harmonogramowanie rozdziału n programów niezależnych niepo-dzielnych i stron pamięci operacyjnej do m procesorów w wieloprocesorowym syste-mie informatycznym. Przedstawiono wyniki badań numerycznych przeprowadzonych na tym algorytmie dla losowo generowanych danych.

2. Sformułowanie problemu optymalizacji

Rozpatrzmy wieloprocesorowy system informatyczny przedstawiony na poniższym rysunku:



Rys. 1. Wieloprocesorowy system informatyczny

Na wieloprocesorowy system informatyczny nakładamy następujące założenia:

- pamięć operacyjna składa się z N stron o jednakowej objętości,
- każdy procesor może wykonywać dowolny program i ma dostęp do dowolnej stro-ny pamięci operacyjnej,
- pamięć zewnętrzna zawiera N_z stron; $N_z > N$,
- liczba programów do wykonania jest większa od liczby procesorów; $n > m$ i każdy program wykonywany jest bez przerwań,
- strony pamięci operacyjnej są przydzielane procesorom i podczas wykonywania programów k-ty procesor może korzystać tylko z u_k stron pamięci operacyjnej jemu przydzielonych; $\sum_{k=1}^{m} u_k \le N, \quad u_k \ge 0, \quad 1 \le k \le m,$
- realizacja każdego z programów na procesorach musi następować niezwłocznie po zakończeniu wykonywania poprzedniego programu lub nastąpić w chwili zerowej, gdy program realizowany jest jako pierwszy na dowolnym z procesorów.

Niech $J = \{1,2,...,i,...,n\}$ oznacza zbiór programów, $P = \{1,2,...,k,...,m\}$ - zbiór procesorów, a $U = \{1,2,...,N\}$ - zbiór stron pamięci operacyjnej. Czas wykonywania i-te-go programu na k-tym procesorze określony jest przez następującą funkcję $T_i(u_k,k)$:

$$T_i(u_k,k) = a_{ik} + \frac{b_{ik}}{u_k}, \quad u_k \in U, \quad k \in P, \quad i \in J,$$

(1)

gdzie $a_{ik} > 0$, $b_{ik} > 0$ są parametrami charakteryzującymi i-ty program i k-ty procesor.

Przedstawiony do rozwiązania problem minimalizacyjny można scharakteryzować następująco: znajdź takie uszeregowanie programów na procesorach i taki przydział stron pamięci operacyjnej do procesorów, aby minimalizować czas T_{opt} zakończenia wykonywania całego zbioru programów J.

Jeżeli przez $J_1, J_2, ... , J_k, ... , J_m$ oznaczymy zbiory programów realizowanych od-powiednio na 1, 2, ... , k, ... , m procesorze, to problem polega na znalezieniu takich zbiorów $J_1, J_2, ... , J_k, ... , J_m$ i takich ilości stron pamięci operacyjnej u_1, $u_2, ... , u_k, ... , u_m$ przydzielonych poszczególnym procesorom, które minimalizują następujące kry-terium optymalizacji:

$$T_{opt} = \min_{\substack{J_1,J_2,...,J_m \\ u_1,u_2,...,u_m}} \max_{1 \le k \le m} \left\{ \sum_{i \in J_k} T_i(u_k,k) \right\}$$

(2)

przy następujących założeniach:

(*i*) $J_s \cap J_t = \emptyset;\quad s,t = 1,2,\ldots,m,\quad s \neq t,\quad \bigcup_{k=1}^{m} J_k = J,$

(*ii*) $\sum_{k=1}^{m} u_k \leq N;\quad u_k \in U,\quad k = 1,2,\ldots,m,$

(*iii*) $u_1, u_2, \ldots, u_m$ – calkowite dodatnie.

Założenie (iii) do kryterium (2) sprawia, że postawiony problem jest dość skomplikowany. W celu uproszczenia problemu przyjmiemy najpierw, że strony pamięci operacyjnej $u_1, u_2, \ldots, u_k, \ldots, u_m$ są typu ciągłego i przy tym założeniu będziemy wyznaczać rozwiązanie problemu. Zaokrąglimy następnie otrzymane wartości optymalne $u_1, u_2, \ldots, u_k, \ldots, u_m$ do najbliższych liczb naturalnych i wówczas rozpatrywany problem sprowadzi się do następującego problemu minimalizacji dyskretno-ciągłej:

$$T_{opt} = \min_{\substack{J_1, J_2, \ldots, J_m \\ u_1, u_2, \ldots, u_m}} \max_{1 \leq k \leq m} \left\{ \sum_{i \in J_k} \widetilde{T}_i(u_k, k) \right\}$$

(3)

przy następujących ograniczeniach:

(*i*) $J_s \cap J_t = \emptyset;\quad s,t = 1,2,\ldots,m,\quad s \neq t,\quad \bigcup_{k=1}^{m} J_k = J,$

(*ii*) $\sum_{k=1}^{m} u_k \leq N;\quad u_k \geq 0,\quad u_k \in U,\quad k = 1,2,\ldots,m,$

gdzie $\widetilde{T}_i : [0, N] \times \{1, 2, \ldots, m\} \to R^+$ jest rozszerzeniem następującej funkcji: $T_i : [1, 2, \ldots, N] \times \{1, 2, \ldots, m\} \to R^+$ i określony jest przez funkcję:

$$\widetilde{T}_i(u_k, k) = a_{ik} + \frac{b_{ik}}{u_k},\quad u_k \in [0, N],\quad k \in P,\quad i \in J.$$

(4)

Przez u_k^*, J_k^*, $k = 1, 2, \ldots, m$ oznaczmy rozwiązania zadania (3). W celu znalezienia tych rozwiązań pomocne będzie wykorzystanie poniższego **lematu**:

Lemat 1

Jeżeli u_k^*, J_k^*, $k = 1, 2, \ldots, m$ są rozwiązaniami optymalnymi zadania (3), to:

(*i*) $\sum_{k=1}^{m} u_k^* = N;\quad u_k^* > 0,\quad k : J_k^* \neq \emptyset,\quad k = 1,2,\ldots,m,$

$$u_k^* = 0,\quad k : J_k^* = \emptyset,\quad k = 1,2,\ldots,m,$$

(*ii*) $\sum_{i \in J_k^*} \widetilde{T}_i(u_k^*, k) = \text{const};\quad k : J_k^* \neq \emptyset,\quad k = 1,2,\ldots,m.$

Warunek (*i*) w **Lemacie 1** mówi, że w przydziale czasowo-optymalnym stron pamięci operacyjnej i programów do procesorów wykorzystuje się wszystkie N stron, a warunek (*ii*), że czasy pracy tych procesorów, które wykonują jakieś

programy są identycz-ne. Zdefiniujmy funkcję $F(J_1, J_2, \ldots, J_m)$ określoną dla zbiorów $J_1, J_2, \ldots, J_m$, dla których zachodzi ograniczenie (*i*) dla kryterium optymalizacji (3). Wartość tej funkcji jest rozwiązaniem następującego układu równań:

$$\begin{cases} \sum_{i \in J_k} a_{ik} + \dfrac{\sum_{i \in J_k} b_{ik}}{u_k} \triangleq F(J_1, J_2, \ldots, J_m); & k : J_k \neq \emptyset, \quad k = 1, 2, \ldots, m \\ \sum_{k : J_k \neq \emptyset} u_k = N; \quad u_k > 0, & k : J_k \neq \emptyset, \quad k = 1, 2, \ldots, m. \end{cases}$$

(5)

Wykorzystując **Lemat 1** oraz (5) zadanie (3) przyjmie ostatecznie następującą postać:

$$T_{opt} = \min_{J_1, J_2, \ldots, J_m} F(J_1, J_2, \ldots, J_m)$$

(6)

przy następujących ograniczeniach:

(*i*) $J_s \cap J_t = \emptyset; \quad s, t = 1, 2, \ldots, m, \quad s \neq t,$

(*ii*) $\bigcup_{k=1}^{m} J_k = J; \quad k = 1, 2, \ldots, m.$

Jeżeli $J_1^*, J_2^*, \ldots, J_m^*$ jest rozwiązaniem zadania (6), to $u_k^*, J_k^*, \quad k = 1, 2, \ldots, m$, gdzie:

$$u_k^* = \begin{cases} \dfrac{\sum_{i \in J_k^*} b_{ik}}{F(J_1^*, J_2^*, \ldots, J_m^*) - \sum_{i \in J_k^*} a_{ik}}; & k : J_k^* \neq \emptyset, \quad 1 \leq k \leq m \\ 0 \qquad ; & k : J_k^* = \emptyset, \quad 1 \leq k \leq m \end{cases}$$

(7)

jest rozwiązaniem zadania (3).

3. Algorytm heurystyczny

Procesory wchodzące w skład wieloprocesorowego systemu informatycznego różnią się pod względem szybkości wykonywania programów. Decyduje o tym liczba stron pamięci operacyjnej przydzielonych poszczególnym procesorom. Dlatego też *k*-ty procesor będzie tym szybszy, im więcej stron pamięci operacyjnej u_k zostanie mu przydzielonych.

Poniżej przedstawiony zostanie algorytm heurystyczny, który najpierw szereguje programy na jednakowych procesorach, tj. takich, do których przydzielona została jednakowa liczba dostępnych stron $u_k = \dfrac{N}{m}, k \in P$. Po tym uszeregowaniu następuje zróżnicowanie procesorów pod względem liczby

przydzielonych im stron pamięci operacyjnej i sprawdzenie, czy skrócony został czas zakończenia wykonywania wszystkich programów $T_{opt.}$.

Strony pamięci operacyjnej przydzielone zostają do procesorów w następujący sposób:

- miarą szybkości realizacji i-tego programu przez k-ty procesor jest tzw. współczynnik podziału stron pamięci operacyjnej β; $\beta > 1$,
- zakładamy, że procesorem najszybszym jest procesor pierwszy, a procesorem najwolniejszym jest procesor m-ty.

Jeżeli procesorowi najwolniejszemu przydzielimy u_m stron pamięci operacyjnej, to do pozostałych procesorów przydział tych stron będzie wyglądał następująco:

$$u_1 = (m-1)\cdot\beta\cdot u_m$$
$$u_2 = (m-2)\cdot\beta\cdot u_m$$
$$\vdots$$
$$u_k = (m-k)\cdot\beta\cdot u_m$$
$$\vdots$$
$$u_{m-2} = [m-(m-2)]\cdot\beta\cdot u_m = 2\beta\cdot u_m$$
$$u_{m-1} = [m-(m-1)]\cdot\beta\cdot u_m = \beta\cdot u_m .$$

Jak wiadomo, pojemność pamięci operacyjnej wynosi N stron, a zatem:

$$\sum_{k=1}^{m} u_k = N.$$

(8)

Rozwijając sumę (8) oraz wprowadzając do niej parametr β otrzymujemy:

$$(m-1)\cdot\beta\cdot u_m + (m-2)\cdot\beta\cdot u_m + \ldots\ldots + (m-k)\cdot\beta\cdot u_m + \ldots + 2\cdot\beta\cdot u_m +$$
$$+\beta\cdot u_m + u_m = N.$$

(9)

Z zależności (9) wyliczamy wartość u_m dla m-tego procesora, czyli procesora najwolniejszego:

$$u_m + \sum_{k=1}^{m-1}[(m-k)\cdot\beta\cdot u_m] = N,$$

a zatem procesorowi m-temu przydzielimy następującą liczbę stron pamięci operacyjnej:

$$u_m = \frac{N}{1+\sum_{k=1}^{m-1}[(m-k)\cdot\beta]}.$$

(10)

Pozostałe procesory otrzymują liczbę stron pamięci operacyjnej określoną następującą zależnością:

$$u_k = (m-k)\cdot\beta\cdot u_m, \quad k = 1,2,\ldots,m-1.$$

(11)

Przedstawiony powyżej sposób przydziału stron pamięci operacyjnej do procesorów wykorzystany zostanie w algorytmie heurystycznym, którego kolejne kroki są następujące:

Krok 1. Oblicz czasy wykonywania programów na poszczególnych procesorach $T_i(u_k,k) = a_{ik} + \frac{b_{ik}}{u_k}$, $i \in J$, $k \in P$ dla zadanej wartości $u_k = \frac{N}{m}$ i losowo ge- nerowanych parametrów a_{ik}, b_{ik}.

Krok 2. Uszereguj malejąco czasy wykonywania poszczególnych programów i utwórz listę L tych programów.

Krok 3. Oblicz średni czas $T_{śr}$ wykonywania programów przez każdy z procesorów wg wzoru:

$$T_{śr} = \frac{\sum_{i=1}^{n} T_i(u_k,k)}{m}; \quad i \in J, k \in P, u_k = \frac{N}{m}.$$

Krok 4. Przydzielaj kolejno najdłuższe programy z listy L do kolejnych procesorów (od pierwszego poczynając) aż do momentu, gdy suma czasów wykonywania programów przydzielonych kolejnym procesorom nie przekroczy czasu $T_{śr}$. Przydzielone programy usuń z listy L.

Krok 5. Jeżeli lista L się jeszcze nie wyczerpała to przydziel na przemian najkrótszy program z listy L do procesora, który ma najdłuższy czas wykonywania programów i najdłuższy program z listy L do procesora, który ma najkrótszy czas wykonywania programów jemu przydzielonych. Usuń te dwa ostatnio przydzielone programy z listy L.

Krok 6. Jeżeli lista L nie została jeszcze wyczerpana to wróć do **Kroku 5**. W przeciwnym wypadku przejdź do **Kroku 7**.

Krok 7. Oblicz czas zakończenia wykonywania wszystkich programów T_{opt} dla uszeregowania programów na procesorach utworzonego w **Krokach 4÷6** i dla $u_k = \frac{N}{m}$.

Krok 8. Oblicz sumaryczne czasy wykonywania programów uszeregowanych na po-szczególnych procesorach.

Krok 9. Usuń najkrótszy program z procesora o najdłuższym sumarycznym czasie wykonywania zadań i przydziel go do procesora o najkrótszym sumarycz-nym czasie wykonywania programów.

Krok10. Oblicz czas zakończenia wykonywania wszystkich programów T_{opt}. po zamianie programów w **Kroku 9**. Jeżeli czas ten ulegnie skróceniu, to wróć do **Kroku 8**. W przeciwnym wypadku anuluj ostatnio wykonaną czynność w **kroku 9** i zakończ szeregowanie programów na procesorach.

Krok11. Dla zadanego współczynnika β przydziel strony u_k, $k \in P$ poszczególnym procesorom wyliczone z zależności (10) i (11).

Krok12. Dla uszeregowania programów na procesorach utworzonego w **Krokach 4÷10** i dla liczby stron u_k przydzielonych procesorom w **Kroku 11** oblicz czas zakończenia wykonywania wszystkich programów T_{opt}.

Krok13. Powtórz **Krok 11** i **Krok 12** dla następnych siedmiu zwiększających się kolejno wartości współczynnika β. Po zakończeniu tych prób przejdź do **Kroku 14.**

Krok14. Porównaj wartości czasów zakończenia wykonywania programów T_{opt} z
kolejnych prób i wybierz najkrótszy z tych czasów.

Krok15. Wyznacz dyskretne liczby stron $\hat{u}_k$, $k \in P$ według zależności:

$$\hat{u}_{\alpha(k)} = \begin{cases} \lfloor u_{\alpha(k)} \rfloor + 1; & k = 1, 2, \ldots, \Delta, \\ \lfloor u_{\alpha(k)} \rfloor \quad ; & k = \Delta + 1, \Delta + 2, \ldots, m, \end{cases}$$

gdzie $\Delta = N - \sum_{j=1}^{m} \lfloor u_j \rfloor$ oraz α jest permutacją elementów zbioru $P = \{1,2,\ldots,m\}$ taką, że $u_{\alpha(1)} - \lfloor u_{\alpha(1)} \rfloor \geq u_{\alpha(2)} - \lfloor u_{\alpha(2)} \rfloor \geq \ldots \geq u_{\alpha(m)} - \lfloor u_{\alpha(m)} \rfloor$. Jeżeli istnieją takie procesory, którym przydzielono zerowe liczby stron pamięci operacyjnej, to przydziel każdemu z tych procesorów po jednej stronie pobierając je z kolejnych procesorów poczynając od procesora, któremu przydzielono największą liczbę stron pamięci operacyjnej.

4. Wyniki eksperymentów obliczeniowych

Przedstawiony w pracy algorytm heurystyczny poddano ocenie dla ośmiu zwiększających się kolejno wartości współczynnika podziału stron pamięci operacyjnej β ze zbioru {2.5, 5.0, 7.5, ... , 20.0}. Parametry a_{ik}, b_{ik} charakteryzujące i-ty program i k-ty procesor wylosowane zostały ze zbioru {1.5, 2.0, ... , 10.0} przez generator o jednostajnym rozkładzie prawdopodobieństwa. Zadano liczbę programów n = 25, 50, 75, 100 i liczbę procesorów m = 2, 4, 6, 8, 10, 12 oraz liczbę stron pamięci operacyjnej N = 1000. Dla każdej kombinacji n i m wygenerowano 30 instancji. Rezultaty analizy porównawczej algorytmu heurystycznego przedstawionego w niniejszej pracy i znanego z literatury algorytmu *LPT* przedstawione zostały w Tab.1.

Tabela 1. Wyniki analizy porównawczej algorytmu heurystycznego i algorytmu *LPT*

n/m	Liczba instancji, dla których:			Δ^H	S^H	S^{LPT}
	$T_{opt}^H < T_{opt}^{LPT}$	$T_{opt}^H = T_{opt}^{LPT}$	$T_{opt}^H > T_{opt}^{LPT}$	%	sek	sek
25/2	16	1	13	1,8	1,8	1,5
50/2	16	0	14	2,8	2,7	2,3

75/2	17	2	11	3,3	4,4	4,2
100/2	19	1	10	3,9	7,1	6,7
25/4	16	0	14	1,7	2,4	2,1
50/4	15	2	13	2,8	3,9	3,3
75/4	18	2	10	3,2	5,3	4,7
100/4	20	1	9	3,9	8,2	7,6
25/6	16	0	14	2,0	2,6	2,5
50/6	15	2	13	2,8	4,4	3,9
75/6	19	0	11	3,6	5,6	5,2
100/6	20	1	9	4,9	8,7	7,7
25/8	15	1	14	2,6	3,4	2,9
50/8	16	3	11	2,9	4,8	4,4
75/8	17	1	12	3,2	7,5	6,6
100/8	21	0	9	4,4	10,4	8,8
25/10	17	1	12	2,6	3,9	3,2
50/10	19	2	9	3,3	5,7	4,6
75/10	20	2	8	3,9	9,6	7,8
100/10	22	1	7	4,9	13,8	11,9
25/12	18	0	12	2,6	3,8	3,2
50/12	19	1	10	3,9	6,5	5,8
75/12	20	2	8	4,7	9,9	7,8
100/12	22	1	7	5,3	14,8	12,7

W Tab. 1. występują następujące wielkości:

m – liczba procesorów,

n – liczba programów,

T_{opt}^{H} – czas zakończenia wykonywania wszystkich programów ze zbioru J przy wykorzystaniu algorytmu heurystycznego,

T_{opt}^{LPT} – czas zakończenia wykonywania wszystkich programów ze zbioru J przy wykorzystaniu algorytmu *LPT*,

Δ^{H} – średnia procentowa poprawa czasu T_{opt}^{H} w stosunku do czasu T_{opt}^{LPT} wyrażona następującym wzorem: $\Delta^{H} = \frac{T_{opt}^{LPT} - T_{opt}^{H}}{T_{opt}^{H}} \cdot 100\%$,

S^{H} – średni czas obliczeń dla algorytmu heurystycznego,

S^{LPT} – średni czas obliczeń dla algorytmu *LPT*.

5. Podsumowanie

Przedstawione w poprzednim rozdziale pracy eksperymenty obliczeniowe wykazały, że efektywność szeregowania programów na równoległych procesorach na bazie zaproponowanego w pracy algorytmu heurystycznego uległa poprawie w stosunku do szeregowania za pomocą znanego z literatury algorytmu *LPT*. Kilkuprocentowa poprawa czasu T_{opt}^{H} w stosunku do T_{opt}^{LPT} może być zachętą do dalszych prac nad efektywnymi algorytmami heurystycznymi.

Zastosowanie przedstawionego algorytmu heurystycznego jest wskazane przede wszystkim dla systemów a dużej liczbie programów, gdyż wówczas średnia procentowa poprawa Δ^{H} jest największa. Zaproponowany w pracy algorytm może służyć zarówno do szeregowania programów w wieloprocesorowych systemach komputerowych, jak i do rozdziału operacji na stanowiska produkcyjne wyposażone w odpowiednie maszyny w dyskretnym systemie produkcyjnym.

LITERATURA

1. Błażewicz J.: Złożoność obliczeniowa problemów kombinatorycznych. WNT, Warszawa 1988.
2. Błażewicz J., Dell'Olmo P., Drozdowski M., Speranza M. G. : Scheduling multiprocessor tasks on three dedicated processors. Information Processing Letters 41, 1992, pp.275-280.
3. Błażewicz J., Drabowski M., Węglarz J.: Scheduling multiprocessor tasks to minimize schedule length. IEEE Transactions on Computers C-35, 1986, pp.389-393.
4. Boctor F. F.: A new and efficient heuristic for scheduling projects will resources restrictions and multiple execution models. European Journal of Operational Research, vol. 90, 1996, pp. 349-361.
5. Buchalski Z.: Optimization of programs scheduling and primary memory allocation in multiprocessing computer systems. Information Systems Architecture and Technology ISAT'98, Wrocław 1998, pp.246-253.
6. Buchalski Z.: Application of heuristic algorithm for the tasks scheduling on parallel machines to minimize the total processing time. Proceedings of the 15th International Conference on Systems Science , vol. 2, Wrocław 2004, pp.235-242.
7. Buchalski Z.: Minimising the Total Processing Time for the Tasks Scheduling on the Parallel Machines System. Proc. of the 12th IEEE International Conference on Methods and Models in Automation and Robotics, Domek S., Kaszyński R. (Eds.), Międzyzdroje, MMAR 2006, 28-31 August 2006, s1081-1084.
8. Janiak A.: Single machine scheduling problem with a common deadline and resource dependent release dates. European Journal of Operational Research, vol. 53, 1991, pp.317-325.
9. Janiak A., Kovalyov M.: Single machine scheduling subject to deadlines and resources dependent processing times. European Journal of Operational Research, 1996, vol. 94, pp.284-291.

10. Józefowska J., Węglarz J.: On a methodology for discrete-continuous scheduling. European Journal of Operational Research, vol. 107, 1998, pp.338-353.
11. Józefowska J., Mika M., Różycki R., Waligóra G., Węglarz J.: Rozwiązywanie dyskretno-ciągłych problemów rozdziału zasobów przez dyskretyzację zasobu ciągłego. Zeszyty Naukowe Politechniki Śląskiej Nr 1474, seria Automatyka, Gliwice 2000, z.129, s.221-229.
12. Krawczyk H., Kubale M.: An approximation algorithm for diagnostic test scheduling in multicomputer systems. IEEE Trans. Comp., C-34, 1985, pp.869-872.
13. Ng C.T., Cheng E.T.C., Janiak A., Kovalyov M.Y.: Group scheduling with controllable setup and processing times: minimizing total weighted completion time. Ann. Oper. Res., 133, 2005, pp.163-174.
14. Nowicki E., Smutnicki C.: The flow shop with parallel machines. A tabu search approach, European Journal of Operational Research 106, 1998, pp.226-253.

Rozdział 25

Zastosowanie narzędzi informatycznych w obliczaniu charakterystyki energetycznej budynków

Stefan Nowak
Politechnika Częstochowska
nowakstef@gmail.com

Streszczenie

Wprowadzenie obowiązku wykonywania świadectw charakterystyki energetycznej dla budynków spowodowało znaczne zainteresowanie producentów specjalistycznego oprogramowania. W niniejszym rozdziale przedstawiono podstawy dotyczące certyfikacji obiektów oraz podjęto próbę klasyfikacji i weryfikacji merytorycznej proponowanych na rynku aplikacji.

1. Wprowadzenie wymogu certyfikacji energetycznej budynków w Polsce

Do początku lat dziewięćdziesiątych problem izolacyjności termicznej budynków praktycznie nie istniał. Dostęp do taniego źródła energii w postaci węgla kamiennego i brak odpowiednich materiałów izolacyjnych w przeszłości sprawił, że większość społeczeństwa do dziś termomodernizację traktuje jako zbędny wymysł administracji. Brakiem motywacji jest przede wszystkim brak odczuwalnej potrzeby zmian. Zainstalowana moc grzewcza urządzeń jest na tyle duża, że w obiektach takich jak bloki mieszkalne temperatura w pomieszczeniach znacznie przekracza minimalne wartości normatywne - po co więc docieplać?

Poglądy te jednak zaczynają ewaluować w kierunku rozwiązań obniżających koszty eksploatacji obiektów. Przyczyną tego jest znaczny wzrostem cen surowców energetycznych w ostatnich latach oraz stosowanie indywidualnego rozliczania kosztów ogrzewania w obiektach wielorodzinnych na przykład na podstawie podzielników ciepła[2].

Światowe tendencje ochrony środowiska skierowane między innymi na ograniczanie emisji CO^2 i poprawę efektywności energetycznej wymuszają zmiany w postaci obniżenia energochłonności budynków, administracyjnie wymuszając spełnianie surowych norm zużycia energii, końcowej jak i pierwotnej, w odniesieniu na metr kwadratowy powierzchni użytkowej danego obiektu.

ŚWIADECTWO CHARAKTERYSTYKI ENERGETYCZNEJ dla budynku mieszkalnego nr 1/8/2009

Ważne do: 21.08.2019r.

Budynek oceniany

Rodzaj budynku	Wielorodzinny budynek mieszkalny
Adres budynku	Częstochowa, ul. xxxxxxxxxx
Całość/część budynku	Całość budynku
Rok zakończenia budowy/rok oddania do użytkowania	1954
Rok budowy instalacji	1954/2004
Liczba lokali mieszkalnych	40
Powierzchnia użytkowa (A_k, m^2)	2261,36

Cel wykonania świadectwa: [ ] budynek nowy [X] budynek istniejący [ ] najem/sprzedaż [ ] rozbudowa

Obliczeniowe zapotrzebowanie na nieodnawialną energię pierwotną[1)]

EP - budynek oceniany

416,54 kWh/(m²·rok)

0 50 100 150 200 250 300 350 400 450 500 >500

Wg wymagań WT2008[2)] budynek nowy

Wg wymagań WT2008[2)] budynek przebudowany

Świadczenie dotrzymania wymagań wg. WT2008 **Budynek nie spełnia wymagań wg.WT2008**

Zapotrzebowanie na energię pierwotną (EP)			Zapotrzebowanie na energię końcową (EK)		
budynek oceniany	416,54	kWh/m²·rok	budynek oceniany	328,20	kWh/m²·rok
budynek wg WT2008	89,63	kWh/m²·rok			

[1)] Charakterystyka energetyczna budynku określana jest na podstawie porównania jednostkowej ilości nieodnawialnej energii pierwotnej EP niezbędnej do zaspokajania potrzeb energetycznych budynku w zakresie ogrzewania, chłodzenia, wentylacji i ciepłej wody użytkowej (efektywność całkowita) z odpowiednią wartością referencyjną.

[2)] Rozporządzenie Ministra Infrastruktury z dnia 12 kwietnia 2002r. W sprawie warunków technicznych jakim powinny odpowiadać budynki i ich usytuowanie (Dz. U. Nr 75, poz. 690, z późn. zm.). spełnienie warunków jest wymagane tylko dla budynku nowego lub przebudowanego.

Uwaga: charakterystyka energetyczna określana jest dla warunków klimatycznych odniesienia - stacja **Częstochowa** oraz dla normalnych warunków eksploatacji budynku podanych na str. 2

Sporządzający świadectwo:

Imię i Nazwisko: dr inż. Stefan Nowak

Nr. Uprawnień: PK/10499/2009

Data wystawienia: 21.05.2009r.

dr inż. Stefan Nowak
Uprawniony do sporządzania świadectw charakterystyki energetycznej
PK/10499/2009

Data Pieczątka i podpis

Rys. 1. Pierwsza strona świadectwa charakterystyki energetycznej budynku - opracowanie własne

Polska wypełniając zapisy Dyrektywy Unii Europejskiej 2009/91/EC z dnia 16.12.2002r. dotyczącej jakości energetycznej budynków[7] wprowadziła w życie Rozporządzenie Ministra Infrastruktury z dnia 6 listopada 2008 r.[6] w sprawie metodologii obliczania charakterystyki energetycznej budynku i lokalu mieszkalnego lub części budynku stanowiącej samodzielną całość techniczno-użytkową oraz sposobu sporządzania i wzorów świadectw ich charakterystyki energetycznej. Stanowi ono podstawę do wykonywania świadectw charakterystyki energetycznej budynków.

Świadectwo charakterystyki energetycznej określa zapotrzebowanie danego obiektu na energię – niezbędną ilość energii, którą w warunkach standardowych trzeba dostarczyć, aby utrzymać odpowiednią temperaturę w pomieszczeniach. Określa tym samym jego jakość energetyczną pośrednio wynikającą z technologii wykonania, użytych materiałów budowlanych i izolacyjnych, jakości wykonania i stopnia zużycia obiektu.

Pierwsza strona świadectwa charakterystyki energetycznej budynku zawiera szereg danych określających badany obiekt oraz jego zdjęcie (rysunek 1). Za najważniejszy element świadectwa należy jednak uznać poziom prezentowanych wskaźników energetycznych, które charakteryzują dany obiekt. W Rozporządzeniu Ministra Infrastruktury z dnia 6 listopada 2008 r. przyjęto prezentacje wskaźników energetycznych w postaci wyniku określonego w kWh/m^2 na rok, usytuowanego na skali dodatkowo wybarwionej dla podkreślenia ilości zużywanej energii. Stanowi to odejście od sposobu prezentacji wyników jakości energetycznej znanych choćby z określania sprzętów AGD, gdzie obowiązują klasy efektywności energetycznej w skali od A do G, gdzie A to urządzenie najbardziej efektywne a G najmniej efektywne. Dla urządzeń chłodniczych dodano jeszcze klasę A+ oraz A++, które oznaczają bardzo wysoką efektywność energetyczną urządzenia.

Prezentowane na pierwszej stronie świadectwa wartości: EP – wskaźnik nakładu nieodnawialnej energii pierwotnej (górna strzałka na skali) oraz EK – wskaźnik nakładu nieodnawialnej energii końcowej wskazują ilość energii niezbędnej do zapewnienia odpowiednich warunków bytowych w danym budynku. Wskaźnikowi te zestawione są z wyliczoną dla danego obiektu ilością energii jaką powinien on maksymalnie zużyć (dolne wskazania na skali). Określa się je dla danego obiektu na podstawie równań zawartych w Rozporządzeniu Ministra Infrastruktury z dnia 6 listopada 2008 r. nowelizującym rozporządzenie w sprawie warunków technicznych, jakim powinny odpowiadać budynki i ich usytuowanie[5]. Określono w ten sposób poziom zużycia energii w zależności od współczynnika kształtu obiektu (A/V).

Maksymalne wartości rocznego wskaźnika obliczeniowego zapotrzebowania na nieodnawialną energię pierwotną do ogrzewania i wentylacji (EP[$kWh/(m^2*rok)$]) określono na poziomie[5]:

$$A/V \leq 0{,}2;\ EP = 73\ [kWh/(m^2*rok)],$$

$0{,}2 \leq A/V \geq 1{,}05$; $EP = 55 + 90 * (A/V)$ [kWh/(m^2*rok)],

$A/V \geq 1{,}05$; $EP = 149{,}5$ [kWh/(m^2*rok)],

gdzie:

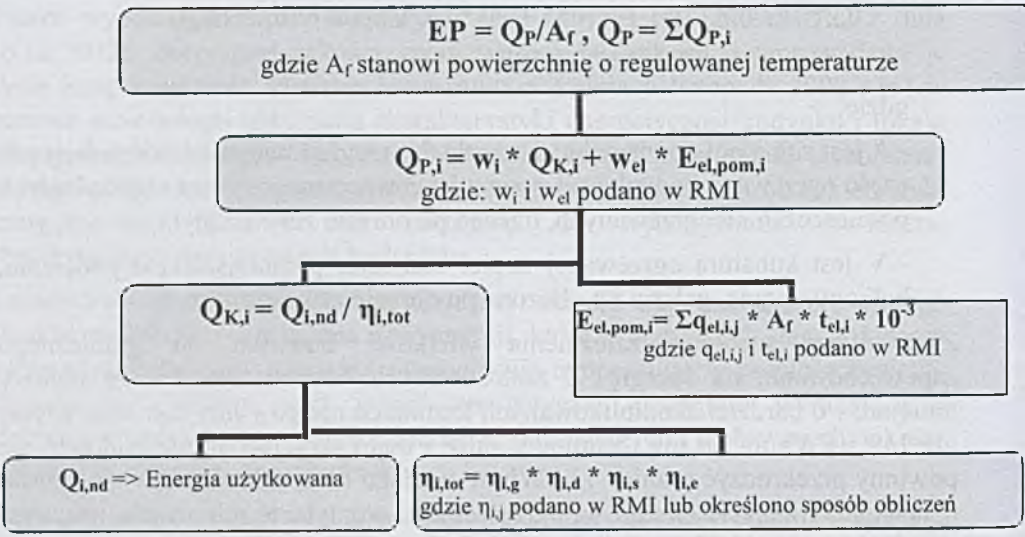
- A jest sumą pól powierzchni wszystkich przegród budynku, oddzielających część ogrzewaną budynku od powietrza zewnętrznego, gruntu i przyległych pomieszczeń nieogrzewanych, liczona po obrysie zewnętrznym,
- V jest kubaturą ogrzewanej części budynku, pomniejszoną o podcienia, balkony, loggie, galerie itp., liczoną po obrysie zewnętrznym.

Zastosowana metoda uzależnienia wielkości budynku od granicznego zapotrzebowania na energię. Z zastosowanych wzorów wynika, że obiekty mniejsze i o bardziej skomplikowanych kształtach nie powinny zużywać więcej niż 149,5 kWh/m^2 na rok. Natomiast duże i mało skomplikowane budowle nie powinny przekroczyć granicy 73 kWh/m^2 w ciągu roku. Określone w ten sposób maksymalne zużycie energii na ogrzewanie i wentylacje odpowiada obecnym standardom energochłonności i jest wskaźnikiem jakości obiektów.

Takie zestawienie graficzne od razu wskazuje na jakość energetyczna danego obiektu uświadamiając możliwości ograniczenia zużycia energii, a przez to możliwości zmniejszenia kosztów eksploatacji obiektu.

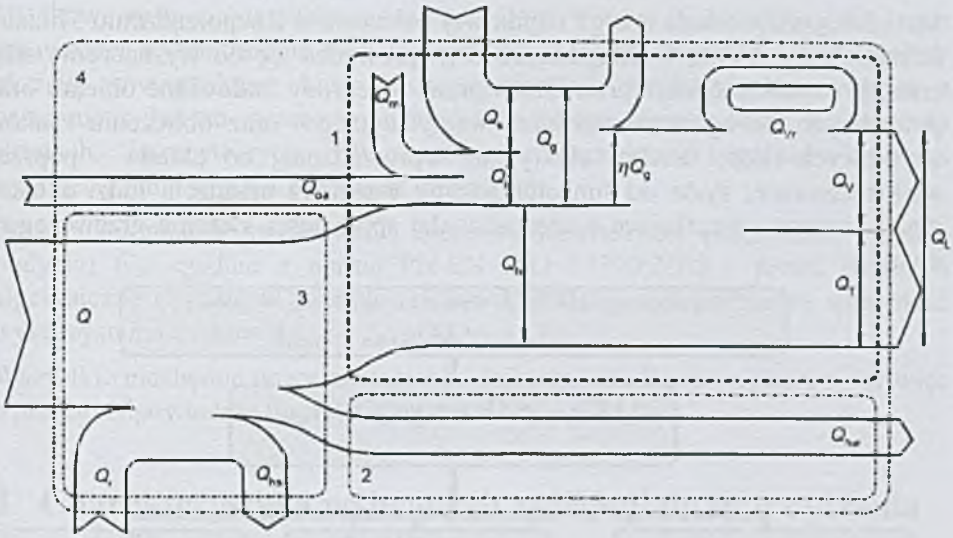
2. Metodologia obliczania charakterystyki energetycznej obiektów

Wskaźnik zapotrzebowania na energię pierwotną EP wskazywany na pierwszej stronie świadectwa charakterystyki energetycznej stanowi wartość sum ilości energii chemicznej używanego paliwa niezbędnej do zaspokojenia potrzeb energetycznych obiektu. Zależny jest zatem od zapotrzebowania na zużywaną energią końcową $Q_{K,i}$ i wskaźników nakładu nieodnawialnej energii pierwotnej w_i oraz energii pomocniczej $w_{el} * E_{el,pom,i}$ (rysunek 2).



Rys. 2. Schemat wyznaczania wskaźnika EP według Rozporządzenia Ministra Infrastruktury z dnia 6 listopada 2008 r. opracowanie własne na podstawie [6]

Należy jednak wspomnieć, iż wszelkie niezbędne wskaźniki potrzebne do obliczenia wartości EP, jak choćby w_i – współczynnik nakładu nieodnawialnej energii pierwotnej na wytworzenie i dostarczenie nośnika energii końcowej do analizowanego budynku, zawarte są w Rozporządzeniu Ministra Infrastruktury [6]. Istotne zatem w obliczeniach roczne zapotrzebowania na energię końcową $Q_{K,i}$ uzyskuje się dzieląc energię użytkową $Q_{i,nd}$ przez iloczyn sprawności cząstkowych systemu grzewczego $\eta_{i,tot}$, którego wartości również można odczytać z Rozporządzenia Ministra Infrastruktury [6]. Opisana więc w Rozporządzeniu procedura wyznaczania wskaźnika EP oparta jest na kilku prostych działaniach algebraicznych i zastosowaniu odpowiednich wskaźników. Wpisując daną w postaci wartości zapotrzebowania na energią użytkową $Q_{i,nd}$, można w popularnym arkuszu kalkulacyjnym opisać prosty algorytm wyliczający wskaźnik EP.



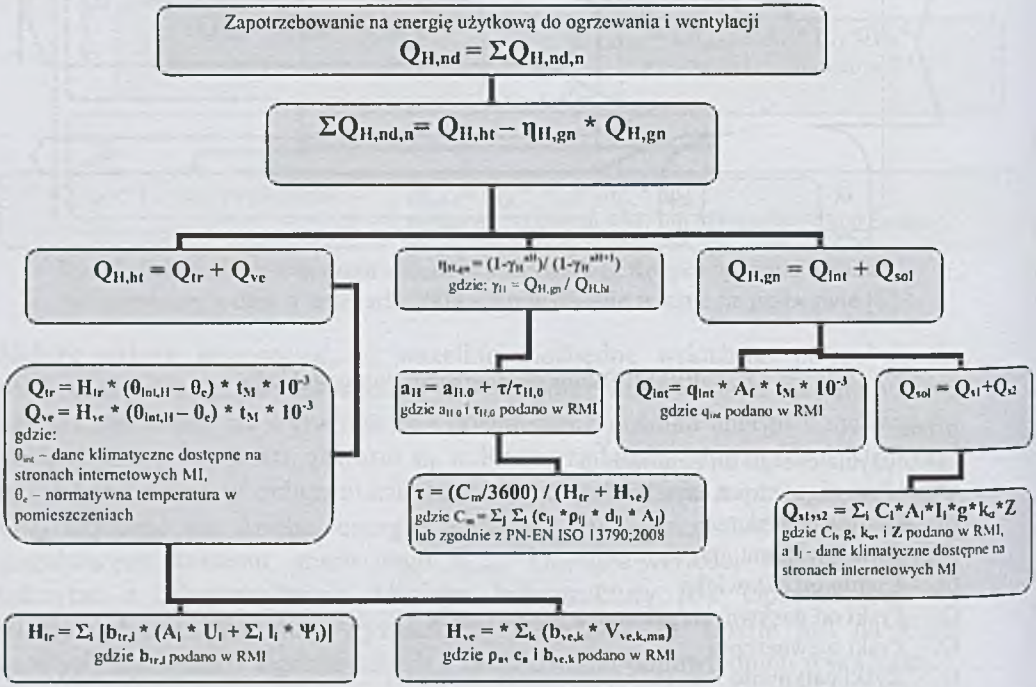
Rys. 3. Bilans energetyczny budynku[11]

gdzie:

Q – Zużycie energii do ogrzewania
Q_{oa} – Energia z innych urządzeń
Q_r – Energia odzyskana
Q_{ns} – Straty systemu grzewczego
Q_m – Ciepło od człowieka
Q_s – Zyski od pasywnych systemów słonecznych
Q_i – Zyski wewnętrzne
Q_g – Zyski całkowite
ηQ_g – Zyski użyteczne
Q_h – Zużycie ciepła
Q_V – Straty ciepła na podgrzanie powietrza wentylacyjnego
Q_{Vr} – Odzysk ciepła z powietrza wentylacyjnego
Q_T – Straty ciepła przez przenikanie
Q_{hw} – Ciepło potrzebne do przygotowania ciepłej wody użytkowej
Q_L – Całkowite straty ciepła
1 – Granica strefy ogrzewanej
2 – Granica systemu przygotowania ciepłej wody użytkowej
3 – Granica instalacji grzewczej
4 – Granica budynku

Zawarty w Rozporządzeniu schemat obliczeń opiera się na bilansie energetycznym budynku i jest zbliżony do schematu zawartego w normie EN ISO 13790:2004 (rysunek 3)

Metodologia obliczania energii użytkowej wskazana w Rozporządzeniu Ministra Infrastruktury z dnia 6 listopada 2008 r. sprowadza się do wyznaczenia strat energii, (wynikających z przenikania przez przegrody budowlane obiektu oraz potrzebnych do ogrzania powietrza wentylacyjnego) oraz obliczeniu zysków użytkowych (ilości energii efektywnie wprowadzonej do układu – poprzez nasłonecznienie, zyski od funkcjonujących wewnątrz urządzeń, ludzi a także zyski ze strat - wynikające z niedoskonałej sprawności systemu grzewczego i cwu).



Rys. 4. Schemat obliczeń zapotrzebowania na energię użytkową do ogrzewania i wentylacji według Rozporządzenia Ministra Infrastruktury z dnia 6 listopada 2008 r. opracowanie własne na podstawie [6]

Metodologia obliczania energii użytkowej, jak wynika z rysunku 4 sprowadza się, w przypadku obliczania strat ciepła przez przenikanie i wentylację $Q_{H,ht}$, do obliczania współczynników strat ciepła - odpowiednio H_{tr} i H_{ve} - pozostałe wartości można wyliczyć stosując odpowiednie formuły w arkuszu kalkulacyjnym. Współczynniki te są charakterystyczne dla konkretnych obiektów, więc powinno się je obliczać na podstawie projektu technicznego obiektu lub obmiarów wizji lokalnej. Analizując wzory wyliczania H_{tr} i H_{ve} można uznać, iż wyliczenia te nie powinny przysporzyć żadnych problemów osobom wykonującym świadectwa, choć mogą się one wydawać żmudne, ze względu na dodawanie poszczególnych powierzchni i długości liniowych mostków cieplnych.

Użytkowe zyski ciepła wyznacza się jako iloczyn sumy zysków $Q_{int} + Q_{sol}$ oraz sprawności wykorzystania tych zysków $\eta_{H,gn}$. Wyznaczając Q_{int} i Q_{sol} należy określić powierzchnie okien i ich usytuowanie względem stron świata a następnie z danych meteorologicznych dla danego regionu (dane dostępne na stronach internetowych Ministerstwa Infrastruktury) pobrać wartości nasłonecznienia w poszczególnych miesiącach.

Przyjmując metodę oszacowania wartości wewnętrznej pojemności cieplnej budynku C_m zgadnie z normą PN-EN ISO 13790;2008 - proste działania algebraiczne opisane w formule arkusza kalkulacyjnego wyznaczą sprawność wykorzystania zysków $\eta_{H,gn}$.

Wszystkie niezbędne pozostałe wartości zawarte są w Rozporządzeniu, tak więc wpisując odpowiednie formuły otrzymamy żądany wynik.

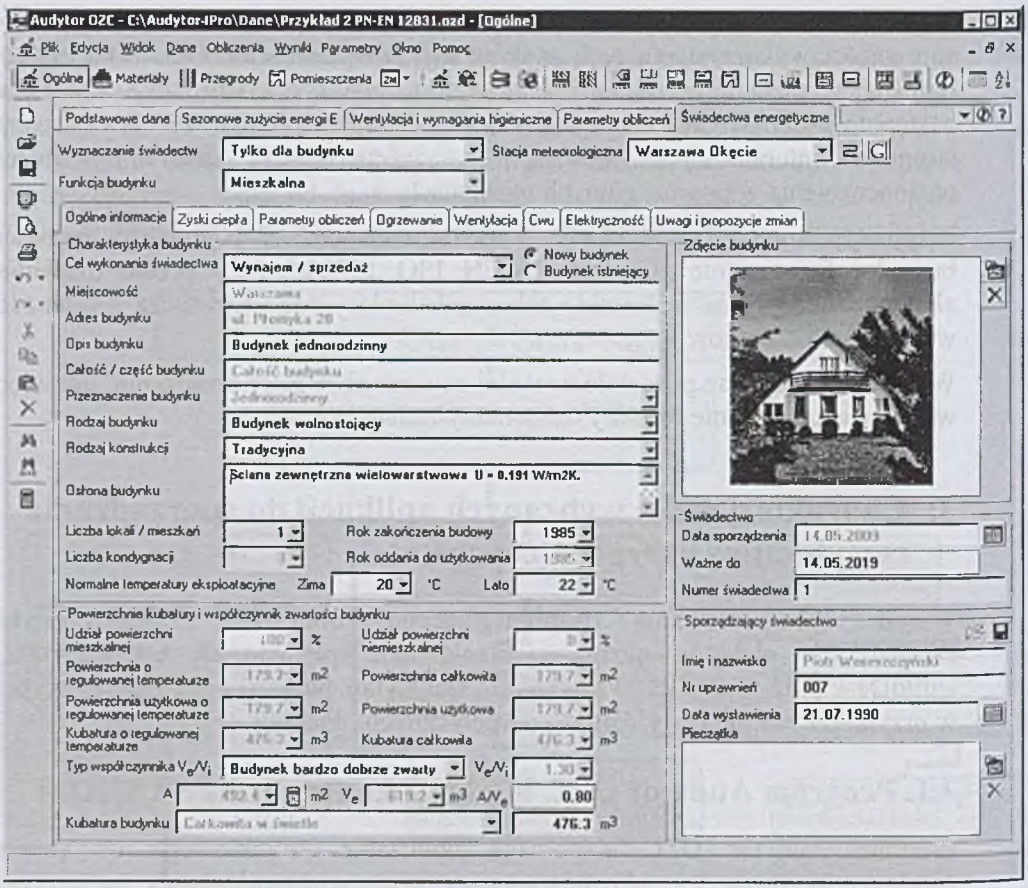
3. Charakterystyka wybranych aplikacji do sporządzania certyfikatów energetycznych

Wśród oprogramowania wspomagającego obliczenia charakterystyki energetycznej obiektów można wymienić takie aplikacje jak Audytor OZC, BuildDesk oraz ArCADia. Programy te stanowią najczęściej używane narządza w przygotowywaniu certyfikatów energetycznych obiektów.

3.1. Program Audytor OZC w wersji 4.6 Pro firmy SANKOM

Program Audytor OZC wspomaga obliczanie projektowego obciążenia cieplnego pomieszczeń, określanie sezonowego zapotrzebowania na energię cieplną do ogrzania budynków oraz wykonywania Świadectw Energetycznych budynków oraz ich poszczególnych części [8]. Do najważniejszych cech aplikacji można zaliczyć[8]:

- Obliczenia do Projektu centralnego ogrzewania, Audytu Energetycznego i Świadectwa Energetycznego w jednym programie.
- Dopracowana metodyka obliczeń uwzględniająca niezbędne korekty błędów występujących w normach i rozporządzeniach - program został opracowany przez firmę od lat zajmującą się tworzeniem oprogramowania z branży sanitarnej i audytingu energetycznego.
- Przejrzystość wyników i możliwość ich szczegółowej analizy.
- Intuicyjna obsługa.
- Automatyczne generowanie danych do obliczeń sezonowego zużycia energii.
- Duża elastyczność konfiguracji obliczeń projektu dzięki zastosowaniu danych domyślnych i dziedziczenia parametrów w strukturze budynku.



Rys. 5. Dialog z wprowadzonymi danymi podstawowymi dla świadectw energetycznych [8]

Podczas wykonywania obliczeń wykonywana jest pełna kontrola wprowadzonych danych umożliwiająca wykrycie błędnie wprowadzonych informacji. Diagnozowane są również otrzymane wyniki obliczeń. Dzięki rozbudowanej diagnostyce znacznie zostaje ograniczona liczba przypadkowych błędów popełnionych podczas wprowadzania danych[8].

3.2. BuildDesk Energy Certificate Professional

BuildDesk Energy Certificate Professional (BDEC PRO) jest programem służącym do analizy energetycznej budynków[9]. Przygotowuje on świadectwo energetyczne zgodne z Rozporządzeniem Ministra Infrastruktury z dnia 6 listopada w sprawie metodologii obliczania charakterystyki energetycznej budynku, projektowaną charakterystykę energetyczną zgodnie z Rozporządzeniem w sprawie zakresu i formy projektu budowlanego z dnia 6 listopada oraz opracowanie dotyczące właściwości cieplno-wilgotnościowych zarówno dla poszczególnych przegród jak i całego budynku. Program

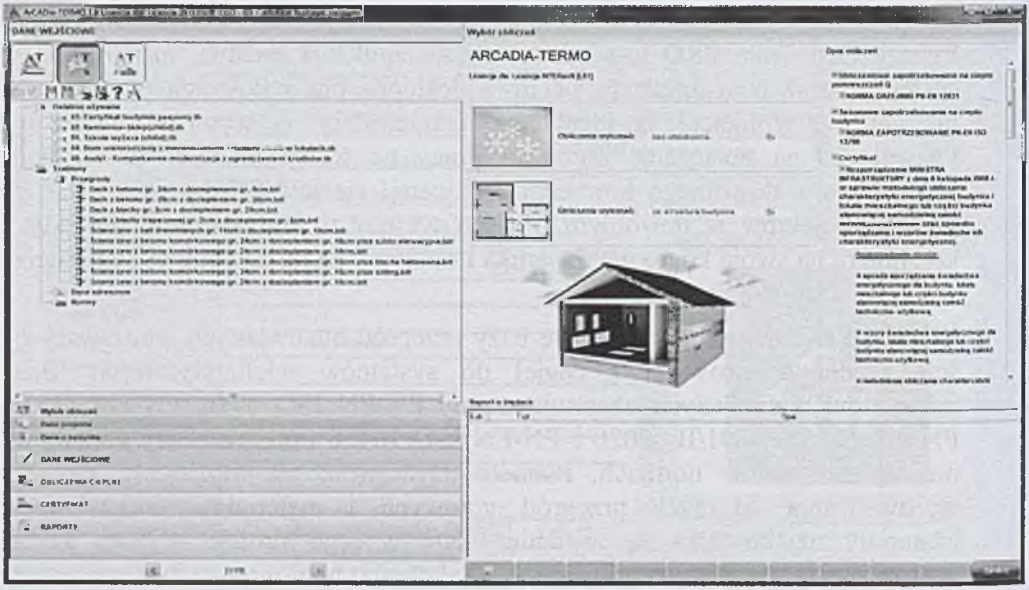
przygotowuje opracowania w formie elektronicznej (plik PDF)[9]. BuildDesk Energy Certificate PRO jest dostępne jako aplikacja lokalna, instalowana w systemie, oraz jako aplikacja sieciowa dostępna przez dowolną przeglądarkę internetową. Aplikacje są identyczne i pozwalają na wymianę projektów. Pozwala to na połączenie komfortu pracy na wersji lokalnej ze swobodą korzystania z dowolnego komputera w wersji sieciowej. Aby skorzystać ze swoich projektów w dowolnym miejscu wystarczy zaimportować projekty z komputera na swoje konto użytkownika i otworzyć je na dowolnym komputerze w wersji sieciowej programu[9].

BDEC PRO zawiera rozbudowane bazy przegród budowlanych, począwszy od ścian jednowarstwowych z cegieł do systemów wielkopłytowych. Bazy materiałów budowlanych obejmują normy PN-EN ISO 6946, PN-EN 12524, PN-EN 1745, PN-91/B-02020 i PN-EN ISO 10456 oraz materiały dodatkowe nie występujące w normach. Ponadto użytkownik ma możliwość zarówno wprowadzania własnych przegród gotowych i materiałów budowlanych. Materiały użytkownika są zapamiętywane w jego profilu i mogą zostać wykorzystane w kolejnych projektach. Dodatkowo w programie zawarte zostały normowe katalogi mostków termicznych[9].

BuildDesk zapewnia wsparcie dla użytkowników systemu. Dedykowane forum dyskusyjne - http://forum.builddesk.pl, gdzie zarówno konsultanci BuildDesk jak i użytkownicy programu wymieniają się doświadczeniem i wiedzą jest doskonałą platformą komunikacji[9].

3.3. ArCADia – Termo firmy INTERSOFT

ArCADia - Termo to program do kompleksowych obliczeń cieplnych budynku. Program pozwala na wykonywanie obliczeń współczynnika przenikania ciepła U, wymiany ciepła przez grunt, mostków cieplnych, sezonowego zapotrzebowania na ciepło budynku. Wykonywanie certyfikatu charakterystyki energetycznej (certyfikat energetyczny, świadectwo energetyczne) wg rozporządzenia Ministra Infrastruktury z dn. 6.11.2008 r[10].



Rys. 7. Dialog z wprowadzonymi danymi podstawowymi dla świadectw energetycznych [10]

Program współpracując z edytorem graficznym CAD pobiera z rzutu architektonicznego stworzonego w systemie ArCADia (lub przeniesionego do ArCADii z innych programów architektonicznych posiadających interfejs IFC np. takich jak program Allplan, ArchiCAD, Revit lub z programu ArCon) geometrię budynku wraz z niezbędnymi danymi i w zasadzie za kliknięciem myszki wykonuje obliczenia cieplne pozwalające na oszacowanie strat w pomieszczeniach, określenie sezonowego zapotrzebowania na ciepło, charakterystyki cieplnej (certyfikat) oraz audytu energetycznego budynku. W takim przypadku ArCADia-TERMO pozwala na bardzo szybkie sporządzenie certyfikatu oraz dokonywanie analiz i obliczeń cieplnych. Otwartość systemu, wymiana danych pomiędzy aplikacjami i współpraca z innymi programami CAD jest przewagą programu ArCADia-TERMO nad produktami konkurencyjnymi[10].

4. Podsumowanie

Obecnie na rynku funkcjonuje duża ilość programów wspomagających generowanie świadectw charakterystyki energetycznej budynków. Przytoczone w niniejszej publikacji aplikacje jak Audytor OZC, BuildDesk oraz ArCADia należą do najpopularniejszych, ale w gruncie rzeczy bardzo są do siebie podobne. Należy jednak zwrócić uwagę na fakt, iż wyznaczenie rocznego zapotrzebowania energii użytkowej, a następnie końcowej i pierwotnej jest precyzyjnie opisane w Rozporządzeniu z dnia 6 listopada 2008 r. w sprawie

metodologii obliczania charakterystyki energetycznej budynków. Używając odpowiednich formuł w popularnych arkuszach kalkulacyjnych, po wpisaniu wartości charakteryzujących analizowany obiekt, można uzyskać dane do wypełnienia świadectwa lub wręcz gotowe świadectwo. Niezbędne jest do tego oczywiście posiadanie odpowiedniej wiedzy inżynierskiej z zakresu fizyki budowli i energetyki, ale w korzystaniu z komercyjnych aplikacji ta wiedza również jest potrzebna. Niewątpliwą i największą zatem zaletą korzystania z własnych kalkulatorów jest pełna znajomość wpisanych formuł i wprowadzanych wartości, oraz możliwość prawidłowej interpretacji uzyskanych wyników.

LITERATURA

1. Laskowski L.: Ochrona cieplna i charakterystyka energetyczna budynku. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2008
2. Życzyńska A., Dyś G.: Finansowanie przedsięwzięć termo modernizacyjnych.[w] Ocena energetyczna budynków. Dostosowanie nieruchomości do wymagań Dyrektywy 2002/91/WE, nowe standardy projektowania i obliczania charakterystyki energetycznej." Wydawnictwo Forum. Wiedza na usługach rynku. Wrzesień 2008 r.
3. Robakiewicz M.: Ocena cech energetycznych budynku, Wymagania, Dane, Obliczenia, Poradnik. Biblioteka Fundacji Poszanowania Energii, Warszawa 2009
4. Rozporządzenie Ministra Infrastruktury z dnia 17 marca 2009 r. w sprawie szczegółowego zakresu i form audytu energetycznego oraz części audytu remontowego, wzorów kart audytów, a także algorytmu oceny opłacalności przedsięwzięcia termomodernizacyjnego
5. Rozporządzenie Ministra Infrastruktury z dnia 17 grudnia 2008 r. w sprawie zmiany rozporządzenia zmieniającego rozporządzenie w sprawie warunków technicznych, jakim powinny odpowiadać budynki i ich usytuowanie
6. Rozporządzenie Ministra Infrastruktury z dnia 6 listopada 2008 r. w sprawie metodologii obliczania charakterystyki energetycznej budynku i lokalu mieszkalnego lub części budynku stanowiącej samodzielną całość techniczno-użytkową oraz sposobu sporządzania i wzorów świadectw ich charakterystyki energetycznej
7. DIRECTIVE 2002/91/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 2002
8. http://www.sankom.pl/s,audytor-ozc4-6-swiadectwa-energetyczne
9. http://www.builddesk.pl
10. http://www.intersoft.com.pl
11. EN ISO 13790:2004
12. PN-EN 12831:2006

Rozdział 26

Metody redukcji szumu w obrazie filmowym oraz ich wpływ na powstawanie zakłóceń pofiltracyjnych

Jakub Kościelny
Politechnika Częstochowska
qbakos@gmail.com

Streszczenie

W rozdziale przedstawiono metody redukcji szumu w obrazie filmowym.

1. Wstęp

Z zakłóceń i zniekształceń mogących wystąpić w obrazie jednym z najczęściej pojawiających się jest właśnie szum (lub jego odmiana – ziarno). Przyczyny jego występowania mogą być różne: najczęściej pojawia się on przy cyfrowej rejestracji słabej jakości sygnału analogowego lub powstaje podczas nagrywania obrazu przez urządzenia dysponujące matrycami CCD lub CMOS w warunkach słabego oświetlenia.

Można przyjąć, że każdy obraz rejestrowany cyfrowo w pewnym stopniu obarczony jest defektem szumu. Tak więc skuteczne metody filtracji obrazu, które będą usuwały lub redukowały szum są ciągle rozwijane. Obecnie nie istnieje jednak metoda filtracji doskonałej, która usuwałaby całkowicie ten defekt bez choćby minimalnego zniekształcenia treści obrazu.

W pracy zajęto się przybliżeniem podstawowych metod filtracji szumów oraz opisano i zwizualizowano rodzaje zakłóceń jakie mogą zostać wprowadzone przez filtry. Przeanalizowano przyczyny powstawania tych zakłóceń oraz możliwe metody ich uniknięcia.

2. Podstawowe metody filtracji szumów

W filtracji obrazów można wyróżnić dwie podstawowe metody działania: filtrację
w dziedzinie przestrzennej oraz w dziedzinie częstotliwości.

2.1. Filtracja w dziedzinie przestrzennej

Najprostszy i zarazem najmniej skuteczny sposób filtracji obrazu w dziedzinie przestrzennej uzyskuje się wykorzystując operację splotu. Nowa wartość piksela jest obliczana na podstawie wartości pikseli sąsiadujących. Każda wartość piksela sąsiadującego jest wagowana wg wartości z macierzy filtra i wpływa na końcową wartość piksela po filtracji. Działanie to można opisać wzorem [1]:

$$P_i = \frac{\sum_{k=1}^{K}\sum_{l=1}^{K} P_{kl} \times F_{kl}}{N} \tag{1}$$

gdzie:

P_i – wartość piksela po filtracji

K – rząd macierzy filtra (zazwyczaj 3 – macierz 3x3)

P_{kl} – wartość piksela obrazu oryginalnego

F_{kl} – wartość wagi filtra

N – suma wartości wag filtra, lub 1 gdy suma wynosi 0

Metoda takiej filtracji zwana jest również filtracją konwolucyjną. Filtr 3-rzędu o każdej wadze równej 1 powoduje uśrednienie wszystkich kolejnych pikseli obrazu. Średnia arytmetyczna 9 składników dla każdego piksela powoduje więc wygładzenie obrazu – usunięcie elementów o wysokiej częstotliwości. Filtr taki działa na zasadzie filtra dolnoprzepustowego. Poniżej najczęściej spotykane macierze wag rzędu trzeciego dla filtrów dolnoprzepustowych:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

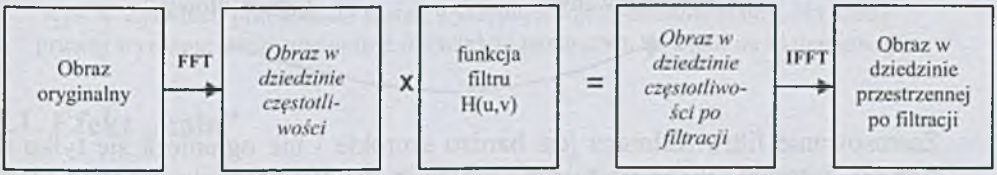
Rys. 1. Najczęściej spotykane macierze wag rzędu trzeciego dla filtrów dolnoprzepustowych

Największą wadą opisywanych filtrów jest to, że filtrowanie przebiega w sposób liniowy. Oznacza to, ze filtry te w jednakowy sposób wpływają na szum, jak i na szczegóły obrazu oryginalnego, powodując jego mniejsze lub większe rozmycie.

Aby uzyskać efekt filtracji z ograniczoną utratą szczegółów obrazu należy posłużyć się filtrami nieliniowymi, a dokładniej filtrem medianowym. W tego rodzaju filtracji uśrednianie wartości pikseli zastąpiono medianą wartości pikseli sąsiednich. Mediana jest to wartość środkowa, czyli liczba, która leży pośrodku ciągu wartości uporządkowanych od najmniejszej do największej. Filtry tego rodzaju usuwają szum bez zamazywania krawędzi obiektów. Filtry medianowe nie są jednak doskonałe. Obraz poddany takiej filtracji jest zniekształcony w szczególny sposób – drobne detale obrazu zlewają się ze sobą.

2.2. Filtracja w dziedzinie częstotliwości

Przejście z przestrzennej dziedziny obrazu x, y do dziedziny częstotliwości u, v umożliwia transformata Fouriera, a dokładniej jej odmiana – szybka transformata Fouriera (FFT). Wykorzystując takie przekształcenie znajduje się dla badanego obrazu $f(x,y)$, jego widmo $F(u,v)$. Tak wyznaczone widmo ma jednak mało wygodną postać do dalszych przekształceń, dlatego dokonuje się jego przesunięcia, polegającego na przestawieniu ćwiartek obrazu widma: lewej-górnej z prawą-dolną i prawej-górnej z lewą-dolną. Taki sposób przedstawienia widma nazywa się centrycznym. Niskie częstotliwości w obrazie źródłowym są reprezentowane przez punkty w środkowej części widma, a wysokie częstotliwości przez punkty na obrzeżach obrazu widma. Następnie usuwa się z widma te elementy, które podejrzewamy że zawierają informację o zakłóceniach. Dokonuje się tego przemnażając (w dziedzinie częstotliwości mnożenie dwóch transformat – obrazu i filtru, jest równoważne splotowi obrazu z filtrem w dziedzinie przestrzennej) odpowiadające sobie elementy funkcji widma i odpowiednio dobranej funkcji filtru $H(u,v)$. Powstaje w ten sposób nowa zmodyfikowana funkcja widma $F'(u,v) = F(u,v) \cdot H(u,v)$. Następnie należy zastosować do niej odwrotną transformację Fouriera by uzyskać obraz $f'(x,y)$ wolny od zakłóceń. Zmodyfikowaną macierz widma należy najpierw przegrupować, przesuwając ćwiartki w odpowiednie miejsca.



Rys. 2. Schemat działania filtracji w dziedzinie częstotliwości

W metodach częstotliwościowych, podobnie jak w metodach przestrzennych można stosować filtracje dolnoprzepustowe, górnoprzepustowe oraz pasmowe. Najczęściej stosowane metody filtracji to filtry Butterwortha, Wienera i Kalmana.

Bardzo dobre rezultaty przy filtracji szumu daje zastosowanie filtru Wienera, który w ostatecznej postaci dany jest wzorem [6]:

$$\hat{F}(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + \dfrac{S_\eta(u,v)}{S_f(u,v)}} \tag{2}$$

gdzie: $|H(u,v)|^2 = H^*(u,v) \times H(u,v)$

$S_\eta(u,v)$ - gęstość widmowa mocy składowej zakłócenia

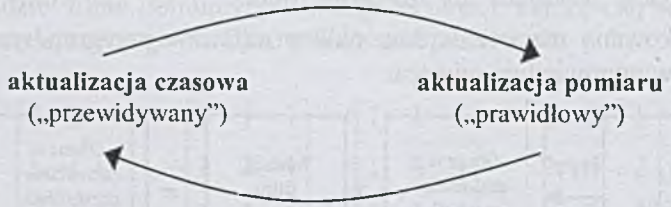
$S_f(u,v)$ - gęstość widmowa mocy obrazu wzorcowego

niech $\dfrac{S_\eta(u,v)}{S_f(u,v)} = K$

dla $K \to 0$ filtr Wienera jest filtrem odwrotnym,

dla $K \to \infty$ filtr Wienera ma charakter filtru dolnoprzepustowego.

Filtr Kalmana to zbiór równań matematycznych, które dostarczają wydajnego rekursywnego sposobu do wyestymowania stanu procesu, w sposób który minimalizuje błąd średniokwadratowy. Można go stosować zarówno do estymacji przeszłych, teraźniejszych jak i przyszłych stanów. Estymacja procesu zachodzi na zasadzie pewnej formy sprzężenia zwrotnego. Filtr estymuje stan procesu w pewnej chwili czasu, po czym otrzymuje informację zwrotną w postaci zaburzonego pomiaru.



Zastosowanie filtru Kalmana jest bardzo szerokie i nie ogranicza się tylko do filtracji szumów. Wykorzystuje się go m.in. w ekonomii do przewidywania mierników ekonomicznych, w inżynierii – w urządzeniach służących do

namierzania i śledzenia obiektów, w urządzeniach typu autopilot; w grafice czasu rzeczywistego itp.

2.3. Podział filtrów ze względu na sposób filtracji obrazu filmowego

W przypadku obrazu filmowego dochodzi dodatkowy problem analizy pikseli w sąsiadujących ze sobą klatkach obrazu. Dlatego też filtry usuwające szum możemy podzielić na 3 podstawowe grupy:

- przestrzenne – analizują piksele wewnątrz jednej klatki obrazu
- czasowe – analizują piksele pomiędzy sąsiadującymi klatkami
- czasowo-przestrzenne – używają kombinacji dwóch powyższych metod (zazwyczaj najbardziej efektywne).

3. Najczęstsze zakłócenia wprowadzane podczas filtracji szumu

Występuje wyłącznie w filtracji czasowej lub czasowo-przestrzennej. Przy zmianach scen lub bardzo szybkich ruchach sąsiadujące ze sobą klatki materiału filmowego znacząco różnią się od siebie. W związku z tym zawarte na nich informacje nie są powiązane w sposób przydatny dla filtracji szumu. Jeśli filtracja czasowa nie zostanie wyłączona np. wskutek błędnego szacowania różnic sąsiednich klatek lub braku takiego szacowania wystąpi zjawisko przenikania sąsiednich klatek. Na takich klatkach pojawią się zakłócenia wynikające z częściowego nałożenia się niektórych składowych kolorów na siebie.

Zjawisko to ilustruje przykład poniżej (rys. 4):

Rys. 4. Zjawisko przenikania klatek występujące przy zmianie sceny. Na klatce prawej wyraźnie widoczne smugi o kształcie pasującym do cieni na klatce lewej.

3.1. Efekt „halo”

Efekt ten powstaje wyłącznie na skutek zbyt silnej filtracji obrazu w dziedzinie częstotliwości. Objawia się tym, że wokół konturów obiektów powstają

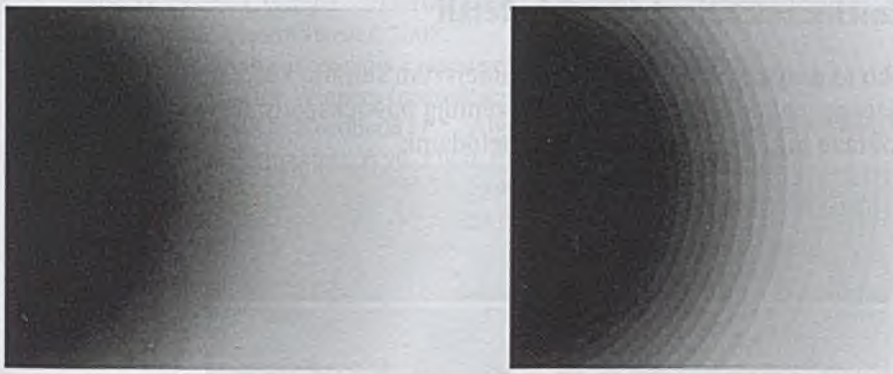
zniekształcenia przypominające kształtem jasne smugi zanikające wraz ze zwiększaniem się odległości od obiektu. Efekt ten jest tym wyraźniejszy im bardziej kontrastowy jest obiekt w stosunku do tła, które go otacza. Powstawanie efektu „halo" można porównać z efektem aliasingu przy filtracji idealnym filtrem dolnoprzepustowym. Efekt ten jest typowym efektem tzw. przesterowania filtra i łatwo go uniknąć dobierając optymalne parametry filtracji.

Efekt „halo" ilustruje przykład poniżej (rys. 5):

Rys. 5. Ilustracja efektu „halo". Klatka lewa zawiera szum, prawa – po filtracji

3.2. Efekt bandingu

Efekt ten powstaje częściej wskutek filtracji w dziedzinie częstotliwości, ale w niektórych przypadkach może też powstawać po filtracji przestrzennej. Objawia się on tym, że w miejscach gdzie powinny występować płynne przejścia barw/jasności (gradienty) powstają zgrupowania jednobarwnych obszarów tworzące wyraźne pasma. Efekt bandingu nie jest związany wyłącznie z filtracją. Powstaje często po kompresji obrazu ze zbyt małą wartością bitrate lub na skutek konwersji przestrzeni barw. Niektóre filmy są szczególnie podatne na powstawanie tego zniekształcenia: filmy animowane techniką komputerową (generalnie występuje duża liczba płynnych przejść tonalnych, cieniowanie) oraz filmy kręcone w warunkach słabego oświetlenia.

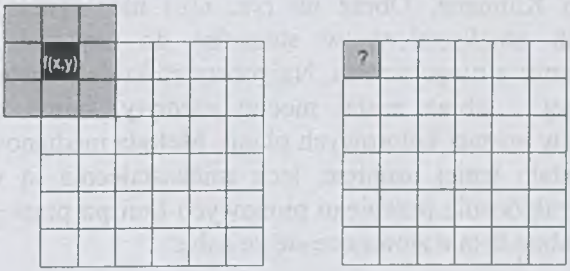


Rys. 6. Ilustracja efektu bandingu. Klatka lewa – bez zniekształcenia, prawa – z wyraźnym efektem bandingu

3.3. Efekt brzegowy i efekt krańcowych klatek

Występowanie efektu brzegowego dotyczy większości filtrów działających w dziedzinie przestrzennej. Polega on na tym, że analizowany pierwszy i ostatni wiersz oraz pierwsza i ostatnia kolumna punktów obrazu nie posiadają 8 sąsiadów, przez co nie mogą być równoważnie analizowane. Najprostszym rozwiązaniem tego problemu jest pomijanie tych wierszy i kolumn w procesie filtracji. Jednak przy materiale o dużej rozdzielczości efekt ten jest niezauważalny.

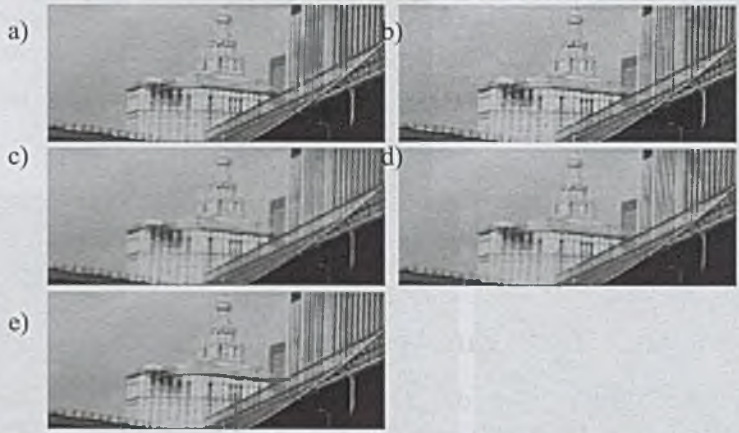
Efekt krańcowych klatek odpowiada efektowi brzegowemu dla filtrów działających metodą czasową. Pierwsza i ostatnia klatka materiału filmowego nie posiada jednej ze swoich klatek sąsiednich, dlatego też filtrację zaczyna się od drugiej i kończy na przedostatniej klatce materiału.



Rys. 7. Ilustracja efektu brzegowego

3.4. Zniekształcenia drobnych detali

Zjawisko to dotyczy w większym lub mniejszym stopniu każdego filtra usuwającego szum. Zdjęcia poniżej prezentują powiększony fragment tej samej klatki obrazu przefiltrowaną różnymi metodami:



Rys. 8. (a) klatka oryginalna; (b) z dodanym szumem; (c) filtrowanie konwolucyjne; (d) filtrowanie medianowe; (e) filtrowanie w dziedzinie częstotliwości z użyciem filtrów Wienera/Kalmana.

4. Podsumowanie

Podsumowując, subiektywnie najlepszy efekt usuwania szumu osiągnięto metodą filtracji w dziedzinie częstotliwości z użyciem zaawansowanych filtrów Wienera lub Kalmana. Obraz na rys. 6(e) nie wykazuje prawie żadnych zauważalnych zniekształceń w stosunku do oryginału przy całkowitym wyeliminowaniu z niego szumu. Najgorszy efekt dała metoda liniowej filtracji konwolucyjnej – obraz został mocno rozmyty, mimo to szum był nadal zauważalny (w postaci kolorowych plam). Metoda medianowa dała lepszy efekt – ziarno zostało lepiej usunięte, lecz zniekształcenia są wyraźnie widoczne, szczególnie zakłócenie przebiegu pionowych linii po prawej stronie klatki oraz widoczne drobne detale zlewające się ze sobą.

LITERATURA

1. Rumiński J.: Metody reprezentacji, przetwarzania i analizy obrazów w medycynie: http://astrophysics.fic.uni.lodz.pl/medtech/dodatki/metpo.html
2. Strumiłło P.: Filtracja obrazu (prezentacja); Politechnika Łódzka, Instytut Elektroniki, 2008.
3. Dadura P., Koc T.: Filtr Kalmana na podstawie "Modelling the equity beta risk of australian financial sector companies."

4. Bołdak C,: Cyfrowe Przetwarzanie Obrazów – Przetwarzanie obrazów w dziedzinie częstotliwości, 2008.
5. Góra P. F.: Analiza szeregów czasowych: Filtr Wienera, 2007.
6. Makowski R.: Filtr Wienera (prezentacja)
7. Tadeusiewicz R., Korohoda P.: Algorytmy i metody komputerowej analizy i przetwarzania obrazów, Wyd. Fund. Post. Telekom. Kraków, 1997.
8. Doom9.net - The definitive DVD backup resource: http://forum.doom9.org/showthread.php?t=108681

Rozdział 27

Analiza skuteczności filtrów redukcji szumu w obrazie filmowym działających w środowisku avisynth

Jakub Kościelny
Politechnika Częstochowska
qbakos@gmail.com

Streszczenie

W rozdziale zaprezentowano wyniki analizy skuteczności filtrów redukcji szumu w obrazie filmowym.

1. Wstęp

Zjawisko szumu jest nierozerwalnie związane z obrazem rejestrowanym w sposób cyfrowy i wiąże się ze specyfiką działania matryc CCD / CMOS. Na złączu półprzewodnikowym elementu światłoczułego oprócz generowanych pod wpływem światła ładunków dochodzi także do samoistnego pojawiania się niepożądanych ładunków elektrycznych zwanych prądem ciemnym. Jego natężenie zależy przede wszystkim od temperatury otoczenia – im jest ona wyższa, tym więcej zostaje wytworzonych przypadkowych ładunków. Prąd ciemny losowo zmienia wielkość ładunku elektrycznego generowanego przez światło, przez co zmniejsza czułość detektora. W wyniku tego procesu powstaje właśnie szum, objawiający się pojawieniem przypadkowo rozmieszczonych różnokolorowych punktów. Zniekształcenia te są tym mniejsze, im wyższy jest stosunek prądu generowanego przez padające fotony do prądu ciemnego (stosunek sygnał/szum). Parametr ten ulega pogorszeniu wraz ze wzrostem prądu wzmocnienia, który jest m.in. większy przy większych wartościach czułości ISO.

W pracy zajęto się analizą sposobu działania wybranych filtrów działających w środowisku AviSynth – pre-procesora obrazu filmowego, pełniącego też funkcję serwera klatek dla innych aplikacji. W dalszej części pracy przeprowadzono badanie skuteczności zaimplementowanych w filtrach metod redukcji szumu

przy użyciu obiektywnych wskaźników jakości, takich jak: PSNR, SSIM, wskaźnik rozmycia.

2. Analiza działania wybranych filtrów pracujących w środowisku AviSynth

Do analizy opisywanych filtrów potrzebny będzie schemat:

n(1)	n(2)	n(3)
n(4)	c	n(5)
n(6)	n(7)	n(8)

c to piksel centralny, czyli aktualnie poddawany badaniu
n(1) ... n(8) to piksele sąsiadujące

2.1. UnDot

Prosty filtr przestrzenny usuwający pojedyncze piksele ziarna występujące na obrazie. Nie jest skuteczny w przypadku szumu występującego na całym obrazie, tak więc należy go stosować wyłącznie do materiału dobrej jakości.

Działanie filtra można opisać wzorem[6]:

$$c = \max[\min(c, MAX(n(1),\dots n(8))), MIN(n(1),\dots n(8))]$$

Inaczej: jeśli badany piksel c ma największą lub najmniejszą wartość spośród 8 sąsiadujących z nim pikseli to zostanie zastąpiony najbliższą mu wartością wybraną spośród pikseli sąsiednich.

Filtr ten nie został uwzględniony w badaniu ze względu na jego małą przydatność w usuwaniu silnego szumu.

2.2. RemoveGrain

Filtr przestrzenny w którym zaimplementowano ponad 20 różnych trybów działania.

Z uwagi na ich dużą liczbę zajęto się najbardziej efektywnymi metodami usuwania szumu. Tryb 4 filtruje obraz standardową metodą medianową z badaniem wszystkich pikseli sąsiadujących. Tryby 5-9 wprowadzają inne podejście, a mianowicie nie są badane wszystkie piksele sąsiadujące, a tylko te które tworzą z badanym pikselem parę ułożoną w linię prostą. Możemy wyróżnić 4 takie pary pikseli: poziomą – n(4), n(5); pionową – n(2), n(7) i 2 ukośne – n(1), n(8) oraz n(3), n(6).

Niech n1 i n2 oznaczają parę pikseli ułożoną w linię. W zależności od wybranego trybu działania filtra, wybierane są te pary pikseli dla których różnica |n1 – n2| jest najmniejsza (tryb 9) lub różnica między pikselem centralnym, a pikselami w linii jest najmniejsza[6]:

|c – max(min(c, MAX(n1,n2)), MIN(n1,n2))| (tryb 5)

Pozostałe tryby 6 – 8 stanowią kombinację dwóch powyższych metod. Tryb 5 jest najmniej skuteczny w usuwaniu szumu, ale wprowadza też najmniej zniekształceń obrazu. Tryb 9 odznacza się największą skutecznością, ale najbardziej zniekształca obraz. Za najbardziej skuteczny i stosunkowo mało agresywny dla obrazu uważa się tryb 17, którego działanie można opisać jako:

m1 = max[min(n1,n2)]
m2 = min[max(n1,n2)]

gdzie:
m1, m2 to wybrane pary pikseli, które nie muszą stanowić linii z pikselem centralnym;
n1, n2 są wyszukiwane spośród wszystkich par pikseli tworzących linię z pikselem centralnym[6].

2.3. Convolution3D

Jest to filtr czasowo-przestrzenny działający na zasadzie prostej filtracji konwolucyjnej. Różnica polega na tym, że zamiast filtracji jednej klatki obrazu za pomocą macierzy wag 3x3 wykorzystywana jest macierz 3x3x3. Tak więc brane są pod uwagę również klatki sąsiadujące z badaną. Filtracja czasowa nie jest jednak przeprowadzana jeśli na sąsiadujących klatkach nie ma informacji przydatnych w procesie redukcji szumu. Sytuacja taka występuje przy gwałtownych ruchach, zmianach scen lub sekwencjach ściemniania/rozjaśniania. Filtr ma możliwość ustalenia wartości progowej po przekroczeniu której filtracja czasowa nie będzie przeprowadzana. Można to zapisać instrukcją programu:

if ($|Y_{0,k} - Y_{0,k-1}| + |Y_{0,k} - Y_{0,k+1}| + |Y_{1,k} - Y_{1,k-1}| + |Y_{1,k} - Y_{1,k+1}|$) > limit
then
filtracja przestrzenna (tylko macierz 3x3)
else
filtracja czasowo-przestrzenna (macierz 3x3x3)

gdzie: $Y_{i,k}$ – piksel o współrzędnych i w klatce k (i=0 – piksel badany, i=1 – piksel kolejny)[7].

2.4. FluxSmooth

Filtr czasowo-przestrzenny, którego głównym założeniem jest twierdzenie, że szum jest zjawiskiem losowym, natomiast ruch nie. Analizowane są piksele na klatkach sąsiadujących z badaną oraz w otoczeniu punktu badanego – łącznie 10 pikseli. Autor nie podaje danych o zastosowanym algorytmie filtracji. Tak jak w opisywanym powyżej filtrze Convolution3D filtracja czasowa jest wyłączana po przekroczeniu określonej wartości różnicy między pikselami na klatkach sąsiednich.

2.5. DeGrainMedian

Filtr czasowo-przestrzenny rozwijający podejście filtracji metodą par pikseli ułożonych w linię prostą (zaczerpniętą z filtra RemoveGrain). Rozwinięcie to polega na tym, że DeGrainMedian używa również informacji z klatek sąsiednich. Rozpatrywany jest więc blok pikseli 3x3x3 na którym można opisać 13 par pikseli tworzących linię. Wybierana jest optymalna para pikseli, która w zależności od wybranego trybu pracy filtra musi spełniać określone warunki, a następnie jest ona używana do uśrednienia wartości piksela centralnego metodą medianową (z ograniczeniami). Wprowadzono 6 trybów, które wyznaczają wagę (limit) na podstawie której dobierane są optymalne pary pikseli. Tryb 0 jest najbardziej skuteczny w usuwaniu szumu, ale powoduje też największe zniekształcenia filtrowanego obrazu. Kolejną zmianą jest fakt, że przekroczenie progowej wartości różnicy pikseli na sąsiadujących klatkach nie wyłącza filtracji czasowej, a jedynie zmniejsza różnicę o jaką filtrowany piksel zostanie zmieniony. Dzięki temu sekwencje zmiany scen czy bardzo szybkie ruchy nie nasilają szumu w tak dużym stopniu.

2.6. FFT3d

Filtr czasowo-przestrzenny działający w dziedzinie częstotliwości. Jego działanie można podzielić na kilka etapów. Najpierw, w zależności od ustawień oraz metody filtracji pobieranych jest od jednej do maks. pięciu klatek obrazu (badana + dwie następne i dwie poprzedzające). Filtr dzieli pobrane klatki filmowe na mniejsze bloki wykorzystując przy tym tzw. efekt zakładkowania – bloki częściowo zachodzą na siebie, co pozwala uniknąć powstawania efektu blokowego i innych zniekształceń często powstających na krawędziach bloków. Następnie dla każdego z tych bloków wykonywana jest transformacja do dziedziny przestrzennej za pomocą FFT (szybka transformata Fouriera). W kolejnym etapie wykonywana jest zaawansowana filtracja widma za pomocą filtrów Wienera lub Kalmana. W przypadku filtracji metodą Kalmana wzmocnienie filtracji jest na bieżąco korygowane w zależności od wahań widma szumu występującego pomiędzy kolejnymi klatkami. W tym więc przypadku można powiedzieć, że na filtrację bieżącej klatki mają wpływ wszystkie wcześniej przefiltrowane klatki. Po tym etapie opcjonalnie wykonywane jest

wyostrzenie obrazu przez wzmocnienie określonych częstotliwości widma. Ostatni etap to wykonanie odwrotnej transformaty Fouriera dla każdego bloku w celu przywrócenia odwzorowania przestrzennego oraz wyrównanie zachodzących na siebie fragmentów bloków.

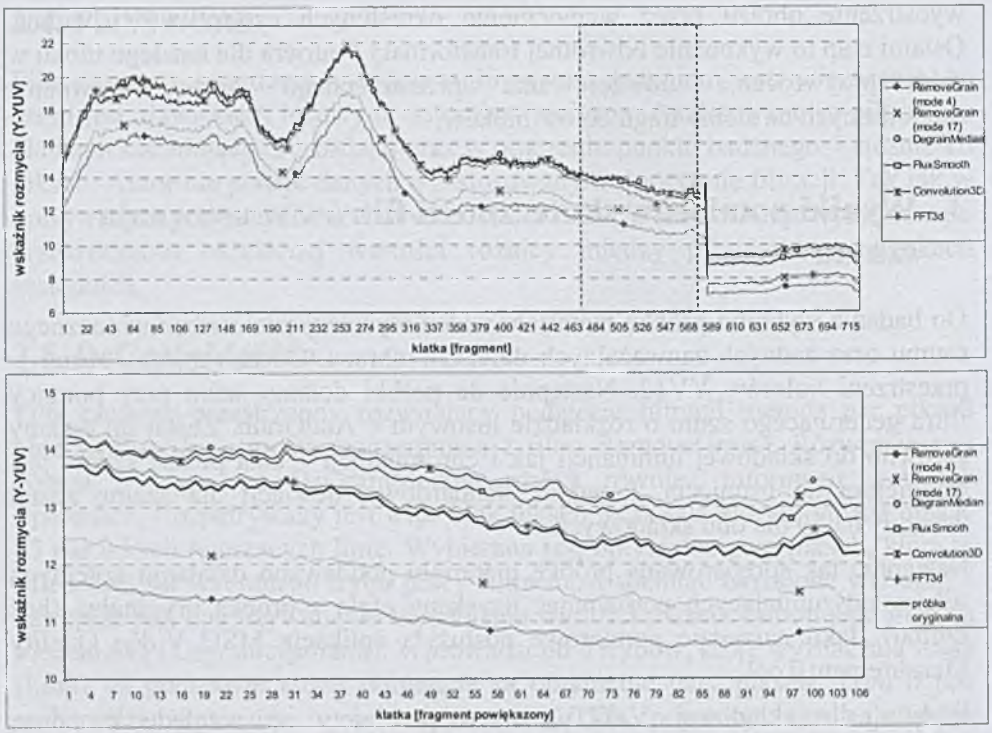
3. Wyniki pomiarów skuteczności filtrów w usuwaniu szumu

Do badania wybrano próbkę materiału wideo wysokiej jakości bez widocznego szumu oraz żadnych zauważalnych defektów obrazu. Obraz zapisany został w przestrzeni kolorów YV12. Następnie do próbki dodano szum przy pomocy filtra generującego szum o rozkładzie losowym – AddGrain. Został on dodany zarówno do składowej luminancji jak i chrominancji – taka postać szumu jest trudniejsza do usunięcia. Parametr standardowej dewiacji dla szumu został ustalony na 10 dla obu składowych.

Następnie tak spreparowaną próbkę materiału poddawano działaniu kolejnych filtrów odszumiających porównując uzyskany efekt z próbką oryginalną (bez szumu). Jako narzędzie pomiarowe posłużyła aplikacja MSU Video Quality Measurement Tool[1].

Wykres dla składowej V-YUV został pominięty ze względu na duże podobieństwo wyników do składowej U-YUV. Wykres dla wskaźnika rozmycia obrazu należy interpretować następująco: im niższa wartość tym większe rozmycie obrazu.

[1] MSU Graphics&Media Lab, Video Group, MSU filters and codecs, http://www.compression.ru/video/



Rys. 1. Zestawienie wyników pomiarów wykonane wskaźnikami SSIM, PSNR dla składowej luminancji (Y) oraz chrominancji (U) oraz wskaźnikiem rozmycia obrazu.

4. Podsumowanie

4.1. Ogólny efekt redukcji szumu

Nie wszystkie filtry poradziły sobie dobrze z redukcją szumu typu losowego występującego w składowej luminancji i chrominancji. Najgorszy efekt osiągnął filtr Convolution3D, po działaniu którego szum pozostał nadal wyraźnie widoczny, jedynie zmniejszyło się jego nasilenie. Szczególnie słabe efekty osiągnął przy filtracji szumu ze składowej koloru. Filtr RemoveGrain we wszystkich trybach działania dobrze usuwał szum, lecz powodował też dość duże zniekształcenia obrazu – szczególnie tryb 4 (standardowa filtracja medianowa) – wynik widać wyraźnie na wykresie przedstawiającym pomiar wskaźnikiem rozmycia. Najlepszy efekt filtracji mierzonej zarówno wskaźnikiem SSIM, jak i PSNR osiągnął filtr FFT3d. Usunął on całkowicie szum nie powodując przy tym praktycznie żadnych zauważalnych zniekształceń obrazu. Filtry DeGrainMedian i FluxSmooth z uwagi na podobny sposób działania osiągnęły zbliżone rezultaty z niewielką przewagą filtra FluxSmooth.

Powodowały one bardzo małe zniekształcenia obrazu, jednak nie udało im się całkowicie usunąć szumu.

4.2. Porównanie efektu filtracji metodą przestrzenną i czasowo-przestrzenną

Filtry opierające się wyłącznie na filtracji przestrzennej miały tendencję do spadku jakości filtracji w obszarach dużego zagęszczenia szczegółów na obrazie – dobrze widoczne przy pomiarze wskaźnikiem PSNR dla składowej Y (klatki ok. 990 – 1100).

Filtry czasowo-przestrzenne osiągały lepsze efekty filtracji. Obraz tracił szczegóły w niewielkim stopniu i było to widoczne jedynie w obszarach o płynnych przejściach barw lub dużej ilości drobnych szczegółów, które miały tendencję do zlewania się. Piksele ziarna nie były rozmywane, ale traciły intensywność przez co szum stawał się mniej widoczny. Przy zmianach scen wyraźnie widoczne są spadki jakości filtracji spowodowane przekroczeniem limitu różnicy pikseli na klatkach sąsiednich, przy czym filtr DeGrainMedian odznaczał się bardziej płynnymi przejściami między scenami.

4.3. Porównanie efektu filtracji w zależności od składowej przestrzeni barw

Wszystkie filtry osiągnęły lepsze efekty filtracji szumu występującego w składowej koloru. Wartość PSNR dla wszystkich próbek przefiltrowanych znajduje się powyżej wartości osiągniętej przez próbkę zaszumioną. Dla składowej jasności filtr RemoveGrain we wszystkich badanych trybach osiągnął średnie wartości PSNR poniżej wartości próbki zaszumionej. Oznacza to, że zniekształcenia, które spowodowała filtracja były większe niż zysk jakości uzyskany dzięki usunięciu szumu.

4.4. Porównanie wyników pomiarów uzyskanych wskaźnikami SSIM i PSNR oraz ocena ich korelacji z subiektywnymi wrażeniami percepcyjnymi odbiorcy

Wyniki uzyskane z pomiarów wskaźnikiem SSIM w większości przypadków nie pokrywają się z pomiarami wskaźnikiem PSNR. Szczególnie duża rozbieżność wystąpiła w przypadku filtra RemoveGrain tryb 4 i tryb 17. Dla badania wskaźnikiem SSIM osiągnięto bardzo wysoki wynik średni, z kolei dla badania PSNR wyniki były jedne z najniższych. Można przypuszczać, że rozmycie obrazu (ogólna cecha filtracji przez RemoveGrain) przy wyliczaniu wartości wskaźnika SSIM jest traktowane jako mniej znaczące niż zmiany jasności i koloru wywołane przez szum (ogólnie większe przy filtrach DeGrainMedian i FluxSmooth).

Bardzo ważna jest obserwacja zmienności wartości obu wskaźników dla próbki z dodanym szumem (szum został dodany sztucznie za pomocą filtra, więc wartość ta powinna być bliska stałej). W przypadku wskaźnika SSIM wahania jego wartości w obrębie całej próbki są bardzo duże, natomiast w przypadku wskaźnika PSNR wartość ta jest niemal stała. Zmienność wartości wskaźnika PSNR dla próbki z dodanym szumem w stosunku do zmienności wartości wszystkich filtrowanych próbek wynosi zaledwie 0,26%; dla wskaźnika SSIM różnica ta wynosi aż 64,02%. Wartość SSIM rośnie znacząco w miejscach dużego nasilenia drobnych detali – wtedy szum percepcyjnie jest mało widoczny, a maleje w miejscach gdzie znajdują się duże powierzchnie o jednolitej barwie – szum jest bardziej widoczny. Jest to dowód na to, że wskaźnik SSIM nie bada deformacji obrazu wynikającej z intensywności zaszumienia, a koreluje się bardziej z ogólnymi wrażeniami percepcyjnymi określającymi jakość obrazu.

Aby wyciągnąć ostateczny wniosek, który z tych dwóch użytych do pomiarów wskaźników jest bardziej miarodajny w badaniu aspektu redukcji szumu należałoby wykonać dodatkowe pracochłonne badania polegające na uśrednieniu ocen subiektywnych dużej liczby badanych i odniesieniu ich do wyników uzyskanych wskaźnikami SSIM i PSNR.

LITERATURA

1. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P.: Image quality assessment: from error visibility to structural similarity; IEEE Trans. Image Processing, vol. 13, no.4; 2004.
2. Wang Z., Sheikh H. R., Bovik A. C.: Objective video quality assessment, Chapter 41 in The Handbook of Video Databases: Design and Applications, B. Furht and O. Marqure, ed., CRC Press, pp. 1041-1078; 2003.
3. Dadura P., Koc T.: Filtr Kalmana na podstawie "Modelling the equity beta risk of australian financial sector companies."
4. MSU Graphics & Media Lab (Video Group): http://www.compression.ru/video/quality_measure/info_en.html ; 2007.
5. Barry T.: Dokumentacja filtra „AddGrainC"; 2003; http://avisynth.org/warpenterprises/files/addgrainc_25_dll_20060610.zip
6. Wittmann R.: Dokumentacja filtra "RemoveGrain"; 2005; http://avisynth.org/warpenterprises/files/removegrain_25_dll_20050501.zip
7. Lucas S.: Dokumentacja filtra "Convolution3D"; 2002; http://avisynth.org/warpenterprises/files/convolution3dyv12_25_dll_20030329.zip
8. Balakhnin A. G.: Dokumentacja filtra "DeGrainMedian"; 2005; http://avisynth.org/warpenterprises/files/degrainmedian_20061008.zip
9. Thomas R.: Dokumentacja filtra "FluxSmooth"; 2004; http://avisynth.org/warpenterprises/files/fluxsmooth_25_dll_20040729.zip

10. Balakhnin A. G.: Dokumentacja filtra "FFT3d"; 2007; http://avisynth.org/warpenterprises/files/fft3dfilter_20070220.zip.

Rozdział 28

Zastosowanie systemu wspomagania decyzji w zakresie doboru i parametryzacji urządzeń audio w pojazdach samochodowych

Zbigniew Buchalski
Politechnika Wrocławska
zbigniew.buchalski@pwr.wroc.pl

Streszczenie

W rozdziale przedstawiono pewną koncepcję systemu ekspertowego o naz-wie AUDIOS wspomagającego proces decyzyjny w zakresie doboru i parametryzacji urządzeń audio w pojazdach samochodowych. Po przedstawieniu struktury organizacyjnej tego systemu podano jego implementację komputerową. Ponieważ system AUDIOS jest systemem modulowym opisano szczegółowo każdy z pięciu modułów wchodzących w skład systemu.

1. Wstęp

Podejmowanie decyzji oznacza akt wyboru jednej możliwości (kierunku) działania spośród pewnego ich zestawu. Wybór ten może być wykonywany na podstawie określonej sekwencji działań, które prowadzą do wyselekcjonowania najkorzystniejszej (optymalnej) alternatywy. Istotną rolę we wspomaganiu procesu decyzyjnego odgrywają inteligentne systemy informatyczne jakimi są systemy ekspertowe [1, 2, 6, 7, 9, 10, 12, 13, 14].

Systemy ekspertowe pozwalają na transfer wiedzy od ekspertów-specjalistów z zakresu problematyki objętej dziedziną systemu do użytkowników tych systemów często znających zagadnienia tej dziedziny, ale w stopniu nie pozwalającym na samodzielne podejmowanie odpowiedzialnych decyzji.

Systemy ekspertowe potrafią doskonalić się na podstawie wiedzy zawartej w bazie wiedzy tych systemów, czyli są systemami samouczącymi się. Pozyskiwanie wiedzy do bazy wiedzy jest procesem bardzo pracochłonnym i budowa systemu ekspertowego ma sens wówczas, gdy będzie on

wykorzystywany przez długi okres czasu i przez wielu użytkowników [2, 3, 4, 8, 11].

Możliwości zastosowań tej nowoczesnej technologii informatycznej są ogromne; począwszy od medycyny, poprzez geologię, technikę aż do zastosowań w dziedzinie wspomagania podejmowania decyzji gospodarczych i finansowych. Trudno podać ściśle wyznaczniki tej technologii, gdyż jest ona w trakcie intensywnego rozwoju.

Celem niniejszego opracowania jest przedstawienie pewnego systemu ekspertowego, naz-wanego AUDIOS, jako narzędzia wspomagającego proces decyzyjny w zakresie doboru i parametryzacji urządzeń audio w pojazdach samochodowych. Podano strukturę organizacyjną tego systemu opisując każdy z pięciu modułów wchodzących w skład systemu. Zaprezentowano implementację komputerową systemu AUDIOS.

2. Założenia i cel budowy systemu AUDIOS

Dobór i konfiguracja sprzętu audio w samochodach jest dziedziną, która do niedawna była skierowana do bardzo wąskiego grona ludzi, jednak sytuacja ta zmieniła się i cieszy się ona stale rosnącą popularnością. System ekspertowy AUDIOS jest systemem budowanym z myślą o osobach stawiających pierwsze kroki w kierunku uzyskania dobrego brzmienia jak i tych, profesjonalnie zajmujących się instalacją sprzętu audio w samochodzie. Przy budowie tego systemu przyjęto takie rozwiązania, które pozwolą zaspokoić merytoryczne potrzeby każdej z wymienionych powyżej grup odbiorców.

Funkcjonalność systemu powinna być złożona, wielopłaszczyznowa oraz uwzględniać wszystkie fazy, które składają się na proces tworzenia wyposażenia audio i jego późniejszą eksploatację. Dzięki takiemu rozwiązaniu uzyskane narzędzie informatyczne poprowadzi użytkownika z punktu początkowego, którym jest dobór poszczególnych komponentów toru elektroakustycznego do punktu finalnego, którym jest usuwanie zakłóceń pojawiających się w sprzęcie audio. Użytkownik systemu otrzyma zintegrowaną platformę świadczącą usługi na wielu płaszczyznach w dziedzinie wyposażenia samochodów w sprzęt audio.

Podstawową funkcją systemu AUDIOS jest wspomaganie w zakresie podejmowania decyzji związanych z doborem elementów toru audio oraz jego parametryzacją. Działanie systemu nie będzie ograniczać się tylko do tego zadania. System będzie pełnił jeszcze jedną bardzo ważną funkcję. Na podstawie informacji podanych przez użytkownika dokona niezbędnych obliczeń, bez których stworzenie dobrego brzmienia audio w samochodzie nie jest możliwe. Funkcjonalność ta będzie dodatkowym atutem systemu, podwyższającym jego znaczenie oraz utwierdzającym w przekonaniu o profesjonalizmie narzędzia informatycznego.

Technologie wykorzystane do budowy systemu AUDIOS umożliwiają zbudowanie narzędzia informatycznego, które może być eksploatowane bez

konieczności posiadania specjalistycznego i trudno dostępnego oprogramowania. Dzięki temu rozwiązaniu system może być rozpowszechniany bez większych ograniczeń. Korzystanie z niego jest możliwe na każdym personalnym komputerze, będącym w posiadaniu przyszłego, potencjalnego użytkownika. Ze względu na dużą popularność pakietu Microsoft Office zdecydowano się na wykorzystanie jego elementów w celu implementacji systemu ekspertowego AUDIOS. Microsoft Access i język Visual Basic for Applications, będące składnikami wspomnianego powyżej oprogramowania, zostaną zastosowane do zrealizowania systemu AUDIOS.

Nawigacja w systemie powinna być intuicyjna. Sposób organizacji poszczególnych jego modułów pozwoli na szybką identyfikację składników, które będą w stanie wesprzeć działania w zakresie wybranego zagadnienia. Dzięki temu sposób korzystania z AUDIOS już po kilku chwilach stanie się zrozumiały i nie będzie wymagał od użytkownika zagłębiania się w wielostronicowe instrukcje. System AUDIOS integruje wiedzę dotyczącą wielu zagadnień z dziedziny wyposażenia samochodów w sprzęt audio w jednym miejscu prezentując ją w przejrzystej i ciekawej formie.

3. Organizacja systemu AUDIOS

AUDIOS jest systemem wielomodułowym, w którym poszczególne elementy działają niezależnie od siebie. Głównym wyznacznikiem przy dokonywaniu podziału była idea zastosowania bloków tematycznych. Każdy z pięciu modułów pełni określoną rolę istotną na poszczególnych etapach wykonywania instalacji audio. Moduły te są następujące:

- Wybór Miejsc Montażowych,
- Dobór Konfiguracji Instalacji Audio,
- Parametry Komponentów Instalacji Audio,
- Strojenie Instalacji Audio,
- Eliminacja Zakłóceń.

Zastosowanie wielomodułowej konstrukcji systemu niesie ze sobą wiele korzyści. Użytkownik już na samym początku otrzymuje informacje na temat przekroju zagadnień poruszanych przez system AUDIOS z zakresu wyposażenia samochodu w sprzęt audio. Dzięki takiemu rozwiązaniu korzystanie z systemu ogranicza się tylko do wybranego w danej chwili tematu przez co skraca się czas dostępu do informacji. Dodatkowym atutem zastosowania struktury o opisanym charakterze jest łatwa nawigacja pomiędzy blokami tematycznymi.

3.1. Wybór Miejsc Montażowych

Jednym z podstawowych pytań jakie należy sobie zadać przed przystąpieniem do projektowania instalacji audio jest kwestia wyboru miejsc montażowych dla

głoś-ników. Stosowane są dwa podejścia do tego problemu. Pierwszym jest wybór standar-dowych miejsc przewidzianych przez producentów samochodów. Drugie podejście ma charakter indywidualny, zależny od preferencji właściciela pojazdu i zaleceń za-kładu montażowego.
Tworząc system AUDIOS zdecydowano się na zastosowanie rozwiązania wykorzystującego standardowe miejsca montażu głośników. Jednym z czynników decydującym o wyborze był fakt, iż większość projektowanych instalacji w zakładach montażowych wykorzystuje oryginalnie przygotowane otwory montażowe. Zaadaptowanie specjalnych panelów głośnikowych wymaga za każdym razem specjalistycznego i indywidualnego podejścia.
System AUDIOS jest narzędziem, które dostarcza wiedzę na temat lokalizacji głośników w wielu różnych samochodach. W bazie wiedzy przechowującej tą informację znajduje się 1829 rekordów, w których dane podzielone są ze względu na markę, model oraz rocznik samochodu. Dzięki tak obszernej bazie wiedzy istnieje możliwość pozyskania informacji podstawowej i zarazem niezbędnej na etapie projektowania instalacji.
System AUDIOS umożliwia również akwizycję wiedzy poprzez dodanie nowych modeli samochodów. Operację tą wykonuje się klikając przycisk *Dodaj samochód*. Dodanie nowego rekordu do bazy wiedzy polega na wprowadzeniu kilku niezbędnych danych takich, jak: marka, model, rocznik, miejsca montażowe, średnica głośnika.
Finalizowanie zaktualizowania bazy wiedzy o nowe pozycje dokonuje się klikając przycisk *Dodaj miejsce montażowe*. Dzięki zastosowanemu rozwiązaniu AUDIOS jest systemem, który może być na bieżąco aktualizowany w miarę pojawiania się coraz to nowych modeli samochodów. Funkcja ta pozwoli systemowi rozwijać się tak dynamicznie, jak będzie tego wymagał rynek motoryzacyjny.

3.2. Dobór Konfiguracji Instalacji Audio

Tor elektroakustyczny składa się z urządzeń elektronicznych mających na celu przetworzenie fali akustycznej na sygnał elektryczny, kolejno poddanie go wstępnemu wzmocnieniu oraz korekcji częstotliwościowej. Następnie zmiksowanie sygnałów oraz ostatecznie wzmocnienie mocy i ponowne przetworzenie sygnału elektrycznego na falę akustyczną. Instalacja audio w samochodzie jest częścią opisanego toru, odpowiedzialną za przekształcenie sygnału elektrycznego, pochodzącego ze źródła dźwięku na falę akustyczną reprodukowaną przez głośniki.

Projektując zestaw audio należy zadbać o to, aby żaden z elementów nie został pominięty. Istnieje wiele urządzeń, które mogą pełnić szereg funkcji, wpływając w różnym stopniu na jakość odtwarzanego dźwięku. Ich zastosowanie w dużej mierze zależne jest od warunków z jakimi zetknie się projektant systemu. Stopień skomplikowania zależny jest od modelu samochodu, oczekiwań brzmieniowych i możliwości finansowych klienta. Charakterystyka

komponentów mających wpływ na tor elektroakustyczny przedstawiona została poniżej:

- **radioodtwarzacz** – jest to jeden z najistotniejszych elementów toru elektroakustycznego w instalacji nagłaśniającej samochodu. Stanowi on źródło sygnału przesyłanego do głośników, z których słyszymy dźwięk. Dobierając model radioodtwarzacza należy określić jakie parametry i funkcje powinien posiadać oraz z jakimi urządzeniami zewnętrznymi powinien współpracować,
- **głośnik** – jest to najsłabszy element wchodzący w skład instalacji audio. Membrana głośnika wprawiana jest w ruch ponieważ przez połączoną z nią mechanicznie cewkę płynie prąd, który wytwarza modulowane pole magnetyczne znajdujące się w polu magnetycznym magnesu stałego. Jednym ze sposobów podziałów głośników jest klasyfikacja ze względu na pasmo przenoszonych częstotliwości. Wyróżnia się następujące typy: niskotonowe, średniotonowe, wysokotonowe [5],
- **wzmacniacze mocy** – są to urządzenia, które wzmacniają sygnał pochodzący ze źródła dźwięku. Wykorzystanie wzmacniacza w instalacji audio umożliwia uzys-kanie lepszego brzmienia oraz dźwięku o wyższym poziomie głośności,
- **zwrotnica** – jest to zestaw filtrów wykorzystywany do podziału pasma częstotliwości. Pojedynczy głośnik nie jest w stanie przetworzyć całego zakresu, należy zatem zastosować zwrotnicę do jego podziału. Jest to element mający bardzo duży wpływ na ostateczne brzmienie systemu audio. Umożliwia on odfiltrowanie rejestrów, które nie mogą zostać poprawnie odtworzone przez zainstalowane głośniki.

Korzystając z modułu Dobór Konfiguracji Instalacji Audio określa się własności opisanych powyżej urządzeń toru elektroakustycznego. Dostępność poszczególnych typów komponentów zależna jest od wyborów dokonywanych przez użytkownika w wyżej wymienionej części systemu AUDIOS. Istnieje możliwość realizacji wielu różnych koncepcji w procesie, którego finalnym produktem jest skompletowana instalacja audio.

System AUDIOS na podstawie analizy wybranych sekcji głośnikowych przedstawia rozwiązania, które mogą być dostępne w następnym etapie. Wpływ na pojawienie się kolejnego formularza z określonymi typami wzmacniaczy mają następujące czynniki:

- ilość sekcji głośnikowych w systemie,
- typy systemów głośnikowych,
- ilość sekcji zasilanych przez wzmacniacze mocy.

W przypadku, gdy instalacja audio jest pozbawiona wzmacniaczy system AUDIOS zakłada, że głośniki będą podłączone bezpośrednio do radioodtwarzacza.

Ostatnim etapem w pracy z modułem Dobór Konfiguracji Instalacji Audio jest określenie sposobu podziału pasma częstotliwości. Użytkownik ma do wyboru trzy możliwości:

- podział pasywny,
- podział aktywny,
- podział pasywny i aktywny.

Dostępność poszczególnych wariantów uwarunkowana jest od wyborów dokonywanych w poprzednich krokach. Jest to ważna część tego modułu, ponieważ sposób dzielenia pasma częstotliwości ma znaczący wpływ na możliwość strojenia systemu audio.

3.3. Parametry Komponentów Instalacji Audio

Wybranie podstawowych urządzeń wchodzących w skład instalacji audio to dopiero połowa drogi do sukcesu. Kolejnym etapem jest określenie parametrów komponentów, bez których system nie może poprawnie funkcjonować. Wszystkie urządzenia należy połączyć przewodami, które umożliwią transmisję sygnału pomiędzy nimi. Dbając o bezpieczeństwo pasażerów podróżujących w samochodzie oraz o sam pojazd należy podjąć wszelkie możliwe kroki, które zabezpieczą przed pożarem spowodowanym zwarciem w instalacji. Ważnym elementem jest również gwarancja dostarczenia odpowiedniej ilości prądu do wzmacniaczy w przypadku, gdy akumulator nie będzie w stanie podołać temu zadaniu. Rozwiązanie opisanych powyżej zagadnień możliwe jest dzięki modułowi Parametry Komponentów Instalacji Audio.

3.4. Strojenie Instalacji Audio

Ostatecznym etapem realizowanym po zamontowaniu wszystkich komponentów jest strojenie systemu. Etap ten ma ogromne znaczenie ponieważ nawet najlepszej jakości urządzenia niepoprawnie zestrojone nie pozwolą cieszyć się doskonałą jakością brzmienia. W zależności od konfiguracji i możliwości komponentów instalacji proces ten dzieli się na trzy części:

- ustawienie podziału pasma częstotliwości,
- regulacja poziomów sygnału,
- ustawienie opóźnień czasowych.

3.5. Eliminacja Zakłóceń

Instalacja wykonana zgodnie z założeniami nawet najdoskonalszego projektu nie gwarantuje perfekcyjnego brzmienia wolnego od zakłóceń. W przypadku instalacji małej mocy prawdopodobieństwo pojawienia się zniekształceń sygnału jest znacznie mniejsze niż dla znacznie bardziej skomplikowanych konfiguracji. W takich sytuacjach należy poświęcić więcej uwagi każdemu z etapów tworzenia instalacji audio w samochodzie. Istnieje

wiele sposobów zabezpieczenia się przed ryzykiem pojawienia się niepożądanych dźwięków w torze akustycznym. Ochronę systemu można zapewnić poprzez dobór komponentów odpornych na zakłócenia takich jak ekranowane przewody, zbalansowane przewody sygnałowe.

Bardzo ważna jest również regulacja poziomów sygnału oraz rozmieszczanie okablowania zgodnie z obowiązującymi regułami. Może się jednak zdarzyć, że pomimo zastosowania się do wszystkich zasad w systemie słyszalne są zakłócenia. Powodem ich powstawania może być sam samochód, który posiada przykładowo uszkodzony zapłon lub brakujące czy uszkodzone fabryczne filtry przeciwzakłóceniowe. System AUDIOS umożliwia rozwiązywanie problemów związanych ze zniekształceniami sygnału.

4. Implementacja komputerowa systemu AUDIOS

System ekspertowy AUDIOS został zaimplementowany w postaci programu o takiej samej nazwie w Microsoft Access z wykorzystaniem języku Visual Basic for Applications (VBA). Access jest jednym z elementów pakietu Microsoft Office, który przeznaczony jest do zarządzania relacyjnymi bazami danych. Umożliwia on projektowanie i zarządzanie uporządkowanymi strukturami danych.

W systemie AUDIOS zastosowano dwa obiekty: tabele i formularze. Pierwszym z nich są dwie tabele o nazwach: marka i samochody. Zawarte w nich dane wykorzys-tywane są do wskazania miejsc montażowych oraz średnic montowanych głośników zależnie od określonej marki samochodu. Drugim rodzajem obiektów wykorzystanych w systemie AUDIOS są formularze. To właśnie na tych obiektach pojawiają się pyta-nia dla użytkownika i dostępne opcje wyboru. W systemie wykorzystano 38 formu-larzy, które stanowią interfejs do komunikowania się użytkownika z systemem. Na formularzach umieszczone zostały elementy, które umożliwiają obsługę systemu. Należą do nich:

- przyciski poleceń,
- pola wyboru,
- etykiety,
- pola tekstowe.

Dostęp do danych, zawartych w opisanych powyżej obiektach, możliwy jest dzięki wykorzystaniu metody ActiveX Data Objects (ADO). Technologia ta jest pewnego ro-dzaju interfejsem dostępu do danych, który zawiera sześć obiektów i dwie kolekcje. W systemie AUDIOS wykorzystane zostały dwa najważniejsze obiekty: Connection i Recordset. Pierwszy z nich definiuje sesję aplikacji ze źródłem, kolejny wykorzystywany jest do reprezentacji pełnego zestawu rekordów zwróconych z tabeli bazy danych.

System AUDIOS został podzielony na kilka głównych modułów, różniących się od siebie m. in. liczbą reguł, zależną od funkcji jaką pełni dany element. W poszczególnych modułach znajdują się następujące liczby reguł:

- Dobór Konfiguracji Instalacji Audio – 62,
- Parametry Komponentów Instalacji Audio – 65,
- Strojenie Instalacji Audio – 18,
- Eliminacja Zakłóceń – 10.

Podsumowując, system AUDIOS zbudowany jest ze 155 reguł, które wspierają proces decyzyjny w zakresie doboru i parametryzacji sprzętu audio w samochodzie.

W kodzie programu oprócz wymienionych powyżej obiektów zastosowano dodatkowo okna typu MsgBox, które wykorzystywane są do udzielania informacji dla użytkownika systemu. Ich rola polega na wyświetlaniu różnego typu komunikatów w trakcie operacji wnioskowania oraz na informowaniu o jego wynikach.

Podstawowym elementem służącym do manipulowania danymi w MS Access, znajdującymi się w tabelach, są kwerendy. Umożliwiają one wyświetlanie danych według określonych kryteriów, które w dowolny sposób mogą być definiowane. Z powodu skomplikowanej procedury pozyskiwania danych z tabel w systemie AUDIOS konieczne było wykorzystanie innej technologii. Doskonałym rozwiązaniem tego problemu było zastosowanie języka Structured Query Language (SQL). Język SQL według American National Standard Institute jest standardowym językiem komputerowym przeznaczonym do dostępu do danych oraz do manipulacji danymi zgromadzonymi w relacyjnych bazach danych. Wyrażenia SQL wykorzystywane są do pobierania i aktualizacji danych w tabelach systemu AUDIOS. Technologia ta została wykorzystana głównie w module Wybór Miejsc Montażowych.

Wykorzystanie wszystkich powyższych technologii pozwoliło na zbudowanie systemu AUDIOS, który pomimo zastosowania kilku różnych rozwiązań informatycznych tworzy dobrze współpracującą całość. Technologie te pozwoliły na zbudowanie systemu o przyjaznej szacie graficznej oraz łatwej obsłudze. Dzięki temu udało się spełnić określone wcześniej założenia systemu.

5. Uwagi końcowe

Przedstawiony w niniejszym rozdziale system ekspertowy AUDIOSYS jest potwierdzeniem możliwości wykorzystania metod sztucznej inteligencji do zastosowań praktycznych. Wszystkie opisane w pracy moduły pozwalają na przeprowadzenie procesu, którego finalnym produktem jest realizacja od podstaw instalacji audio w samochodzie. Uzyskiwana informacja może być wykorzystywana przez profesjonalne zakłady montażowe jak i osoby amatorsko zajmujące się wyposażaniem samochodów w sprzęt audio.

Organizacja systemu AUDIOS oraz jego przejrzystość umożliwiają uzyskanie klarownej i łatwo przyswajalnej informacji, która z powodzeniem trafia do wszystkich.

System AUDIOS powstał przy wykorzystaniu języka Visual Basic for Applications i oprogramowania Microsoft Access. Rozwiązanie to nakłada pewnego rodzaju ograniczenie.

Korzystanie z systemu wymaga od użytkownika posiadania jednego z elementów pa-kietu Microsoft Office, bez którego jego działanie nie jest możliwe. W przyszłości bę-dzie można zastosować inny język programowania, który uczyni system bardziej uniwersalnym. Istnieje wiele narzędzi programistycznych, które mogą zostać wykorzystane do przeniesienia wiedzy zawartej w bazie wiedzy systemu AUDIOS na inną platformę.

Przeprowadzenie testów systemu AUDIOS pozwoliło na sprawdzenie poprawności i skuteczności jego funkcjonowania. Na podstawie analizy wykonanych testów można stwierdzić, że powyższe cechy systemu rekomendują go do zastosowań praktycznych.

LITERATURA

1. Buchalski Z.: Komputerowe wspomaganie podejmowania decyzji z wykorzystaniem regułowego systemu ekspertowego. W: Komputerowo zintegrowane zarządzanie, tom 1, R. Knosala (red.), WNT, Warszawa 2004, s.156-164.
2. Buchalski Z.: Knowledge Management of Expert System Based on the Symbolic Repres-entation of Natural Language Sentences. W: Information Systems Architecture and Techno-logy, L. Borzemski, A. Grzech, J. Świątek, Z. Wilimowska (eds.). Oficyna Wydawnicza Poli-techniki Wrocławskiej, Wrocław 2006, pp.75-85.
3. Buchalski Z.: Zarządzanie wiedzą w podejmowaniu decyzji przy wykorzystaniu systemu eks-pertowego. W: Bazy danych. Struktury, algorytmy, metody. WKiŁ, Warszawa 2006, s.471-478.
4. Buchalski Z.: The Role of Symbolic Representation of Natural Language Sentences in Know-ledge Acquisition for Expert System, Polish Journal of Environmental Studies, Vol.16, No 4A, 2007, pp.40-43.
5. Krajewski J.: Głośniki i zestawy głośnikowe, WKiŁ, Warszawa 2003.
6. Krishnamoorthy C.S., Rajeev S.: Artificial Inteligence and Expert Systems for Engineers. CRC Press, London 1994.
7. Liebowitz L.: The Handbook of Applied Expert Systems. CRC Press, London 1996.
8. Niederliński A.: Regułowo-modelowe systemy ekspertowe. Pracownia Komputerowa Jacka Skalmierskiego, Gliwice 2006.
9. Owoc M.: Elementy systemów ekspertowych, cz.1: Sztuczna inteligencja i systemy ekspertowe. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2006.
10. Radzikowski W.: Komputerowe systemy wspomagania decyzji. PWE, Warszawa 1990.
11. Rutkowski L.: Metody i techniki sztucznej inteligencji, PWN, Warszawa 2006.

12. Stefanowicz B.: Systemy eksperckie. Przewodnik. PWN, Warszawa 2003.
13. Twardowski Z.: Inteligentne systemy wspomagania decyzji w strategicznym zarządzaniu organizacją gospodarczą. Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice 2007.
14. Zieliński J.: Inteligentne systemy w zarządzaniu. Teoria i praktyka. PWN, Warszawa 2000.

Rozdział 29

Zastosowanie pakietu Microsoft Expression Studio w dydaktyce informatyki

Iwona Iskierka, Sławomir Iskierka
Politechnika Częstochowska
iwona.iskierka@el.pcz.czest.pl

Streszczenie

W pracy przedstawiono możliwości pracy z pakietem Microsoft Expression Studio w działalności dydaktycznej w zakresie przedmiotów związanych z grafiką komputerową (Microsoft Expression Design), przetwarzaniem obrazów, animacją (Microsoft Expression Blend) oraz technologiami internetowymi (Microsoft Expression Web). Przedstawiono możliwości współpracy programu z innymi technologiami w procesie tworzenia różnorodnych aplikacji.

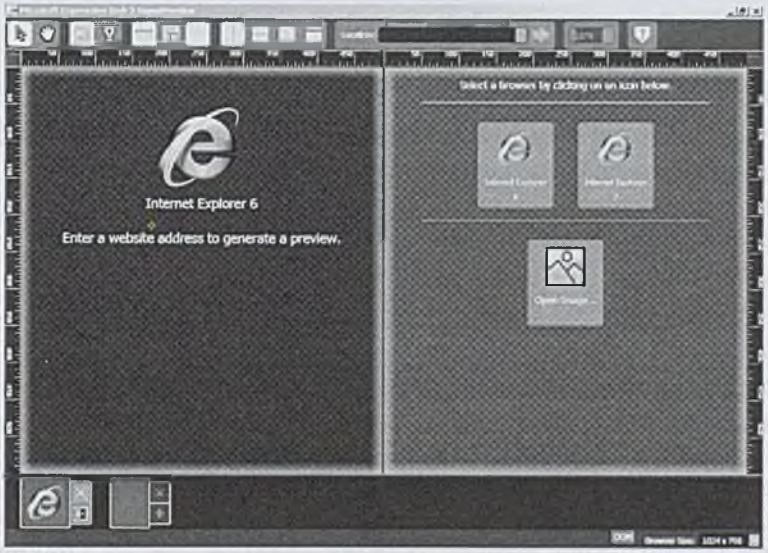
1. Pakiety Expression Studio 2.0 i Expression Studio 3.0

Firma Microsoft od połowy 2006 roku publikuje wersje trial swoich najnowszych produktów dla osób zajmujących się interaktywnym tworzeniem stron WWW i aplikacji webowych. Zestaw proponowanych narzędzi nosi nazwę Expression Studio. W procesie tworzenia interaktywnych stron WWW istotnymi czynnikami wpływającymi na końcowy efekt pracy są: wybór odpowiedniej technologii, możliwości współpracy grafika i programisty, możliwości współpracy wielu narzędzi i sposób wymiany danych między nimi. W pakiecie Expression Studio uwzględniono najistotniejsze elementy wpływające na proces budowy i strukturę tworzonych aplikacji. W skład pakietu wchodzą następujące produkty: Expression Web, Expression Blend, Expression Design, Expression Media, Expression Media Encoder [1]. Obecnie Microsoft udostępnił kilka nowych wersji testowych programów, które wejdą w skład pakietu Expression Studio 3.0 - zaawansowanego zestawu do tworzenia witryn internetowych i aplikacji webowych. Wśród nich znalazły się narzędzie Expression Web SuperPreview i nowa edycja Microsoft Blend.

Microsoft Expression Studio 2.0 zawiera wszystkie produkty rodziny Expression oraz Microsoft Visual Studio 2005 Standard Edition. W pakiecie tym udostępniono projektantom profesjonalne narzędzia projektowe i programistyczne, potrzebne do tworzenia aplikacji nowej generacji dla systemu Microsoft Windows i aplikacji internetowych [2].

Microsoft Expression Web jest przeznaczony do tworzenia zgodnych ze standardami stron internetowych. Jest to narzędzie do budowy stron internetowych wywodzące się z produktu FrontPage, ale zawiera wiele nowych funkcjonalności oraz możliwości programowania. W programie Microsoft Expression Web witryny są tworzone z wykorzystaniem aktualnie obowiązujących standardów, w tym XHTML, CSS, XML i XSLT. Expression Web, w początkowej fazie projektu znany jako Expression Web Designer, to zaawansowane narzędzie do projektowania i rozwijania serwisów internetowych z wykorzystaniem takich technologii i języków, takich jak XHTML (Extensible Hybertext Markup Language), XML (Exensible Markup Languge), CSS (Cascading Style Sheets), JavaScript i ASP.NET 2.0. W Expression Web dane mogą być pobierane z plików XML i przetwarzane za pomocą złożonych zapytań XPath (XLM Path Language) oraz transformacji XSLT (eXtensible Stylesheet Language Transformations) dla zapewnienia jak najlepszego sposobu ich prezentacji na stronie. Wydany pod koniec grudnia 2006 roku Expression Web działa w systemach operacyjnych Windows XP z Service Pack 2 lub w Windows Vista. Wymagana jest też instalacja środowiska .NET Framework wersja przynajmniej 2.0. Do wydajnej pracy przydadzą się ponadto komputer z wydajnym procesorem i co najmniej 512 MB pamięci operacyjnej [3].

W pakiecie Expression Studio 3.0 udostępniono ponadto Expression Web SuperPreview. Aplikacja zainteresuje wszystkich deweloperów, którzy chcą wyświetlać strony zgodnie z trzema ostatnimi wersjami przeglądarki Internet Explorer (6, 7 i 8). Narzędzie ma wbudowaną funkcję obsługi innych przeglądarek: Google Chrome, Firefoksa i Safari. Istnieje też możliwość nakładania otwartych stron jedna na drugą, co pozwala wychwycić różnice [4]. SuperPreview pobrać można z witryny Microsoftu (rozmiar pliku 250 MB).



Rys. 1. Expression Web SuperPreview

W tworzonych obecnie witrynach internetowych warstwa graficzna została oddzielona od zawartości kodu, co jest możliwe dzięki wykorzystaniu arkuszy styli kaskadowych. W programie Expression Web zaproponowane narzędzia projektowe dają użytkownikowi pełną kontrolę nad budową spójnie wyglądających witryn. Można wykorzystać możliwości związane z użyciem przeciąganego i upuszczanego marginesu graficznego. Zaproponowano kontrolę dopełniania oraz hierarchię wizualną aplikacji do obróbki stylistycznej. Wymienione elementy pozwalają w prosty sposób opracowywać projekty graficzne [5].

Microsoft Expression Blend jest profesjonalnym narzędziem do tworzenia aplikacji nowej generacji dla systemu Windows. Tworzone aplikacje mogą w pełni wykorzystać możliwości technologii Windows Presentation Foundation, która jest podstawowym składnikiem .NET Framework 3.0.

Microsoft Expression Blend pozwala na tworzenie interfejsów użytkownika dla aplikacji nowej generacji, łączących najlepsze zalety aplikacji internetowych z możliwościami wykorzystania pełnego potencjału komputera osobistego. Interfejsy użytkownika tworzy się z wykorzystaniem różnorodnych obiektów. Jest to grafika wektorowa i rastrowa, pliki, dźwiękowe, wideo, tekst, grafika 3D i animacja. W programie Expression Blend wbudowano możliwość tworzenia interaktywnych projektów bez pisania konieczności pisania nawet linijki kodu. Kontrolki pozwalają na tworzenie różnorodnych elementów od prostych przycisków po kontrolki wyświetlania danych. Użytkownik może tworzyć powiązania pomiędzy kontrolkami oraz pomiędzy kontrolkami i źródłami danych. W pakiecie Expression Studio 3.0 wprowadzono wiele nowości dotyczących Microsoft Blend 3.0. Są to wsparcie dla Silverlighta 3.0 beta, możliwość importu plików z programów Adobe PhotoShop (.psd) i Adobe

Illustrator (.ai), usprawnienia narzędzi gradientowych oraz operacji dostosowywania wyglądu. Udoskonalono także operacje na tekście oraz animacjach.

Za pomocą Microsoft Expression Design użytkownik może tworzyć grafikę wektorową i ilustracje do wykorzystania w aplikacjach dla Microsoft Windows i aplikacjach internetowych. Utworzoną w Microsoft Expression Design grafikę wektorową można następnie wykorzystać i animować w programie Microsoft Expression Blend. Utworzone elementy za pomocą Expression Design, eksportuje się do postaci XAML i animuje za pomocą interaktywnych funkcji Expression Blend.

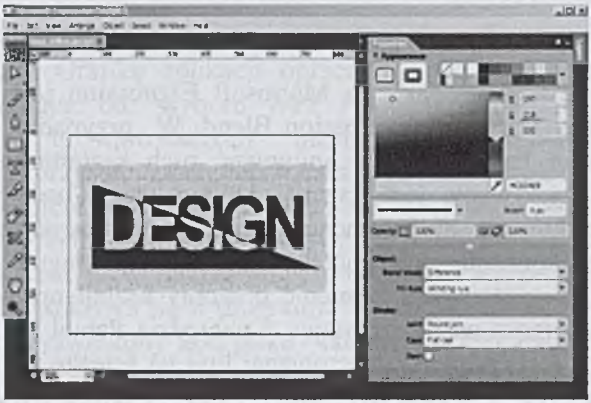
Program Microsoft Expression Media pozwala na zarządzanie wszystkimi zasobami, z których użytkownik korzysta podczas pracy w Microsoft Expression Studio. Użytkownik ma możliwości importu plików w ponad 100 różnych formatach. Można przechowywać pliki w dowolnym miejscu - mogą to być foldery udostępnione, płyty CD, dyski twarde , płyty DVD. Zintegrowane narzędzie wyszukiwania pozwoli odnaleźć pliki w bardzo krótkim czasie, niezależnie od miejsca w którym się znajdują. Użytkownik ma możliwości realizacji podstawowych zadań edycji i korekcji plików graficznych (takie jak zmiana rozmiaru czy usuwania efektu czerwonych oczu) z wnętrza Expression Media. Przetwarzanie wsadowe pozwala na łatwą obróbkę wielu plików jednocześnie [6].

Microsoft Expression Media to profesjonalne narzędzie zarządzania zasobami, pozwalające na graficzne katalogowanie i organizowanie obiektów cyfrowych w sposób umożliwiający łatwe ich odnalezienie i wykorzystanie. Wymiana opinii z innymi użytkownikami oprogramowania możliwa jest za pośrednictwem grupy dyskusyjnej Media Encoder. Przystępny i przyjazny interfejs użytkownika Expression Media Encoder umożliwia projektantom, edytorom i producentom generowanie materiałów w formacie VC-1, zapisanych jako prezentacje Silverlight. Silnik kodujący może być także sterowany z poziomu wiersza polecenia, co pozwala na wykorzystanie go w czasochłonnych i powtarzalnych pracach wsadowej kompresji plików. Należy zaznaczyć, że w pakiecie Microsoft Expression Studio definicja interfejsu opisywana jest językiem XAML (Extensible Application Markup Language), który bazuje na XML-u. Obsługa zdarzeń może być napisana w jednym z dwóch języków: C# lub VB.NET.

XAML jest metajęzykiem opisującym wygląd aplikacji. Dzięki językowi XAML budowanie interfejsu użytkownika jest bardzo proste i tworzone w sposób deklaratywny. Może być użyty dla dowolnego obiektu. Obiekty stworzone za pomocą XAML mogą być renderowane w przeglądarce lub też jako aplikacja okienkowa. Wszystkie pliki z definicją interfejsu stworzone w XAML mają generalnie rozszerzenie *.xaml.

2. Przegląd możliwości pakietu Microsoft Expression Studio

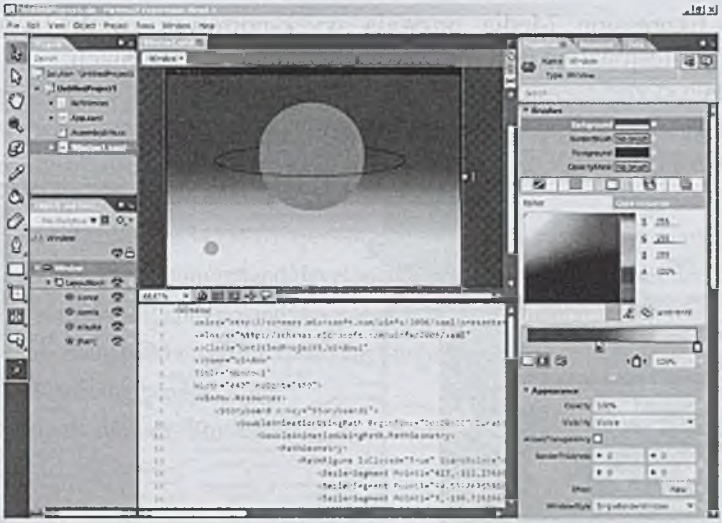
Microsoft Design pozwala na swobodne tworzenie grafiki wektorowej. Użytkownik może rysować dowolne kształty z wykorzystaniem elastycznych i intuicyjnych narzędzi, takich jak krzywe B-sklejane, aerograf czy krzywe Béziera. Może także tworzyć interesujące projekty hybrydowe, wykorzystując zaawansowane narzędzia edycyjne, operatory ścieżek i funkcje transformacji kształtów, które pozwalają na precyzyjne edytowanie grafiki. Obiekty graficzne utworzone w programie Microsoft Design użytkownik możne wykorzystywać w Expression Web, Microsoft Visual Studio i innych narzędziach programistycznych. Ważną funkcją jest możliwość dostosowania środowiska roboczego. W tym celu użytkownik może wykorzystać innowacyjne funkcje, które pozwalają nawet na dowolne obracanie ram rysunku na ekranie.



Rys. 2. Środowisko pracy Expression Design

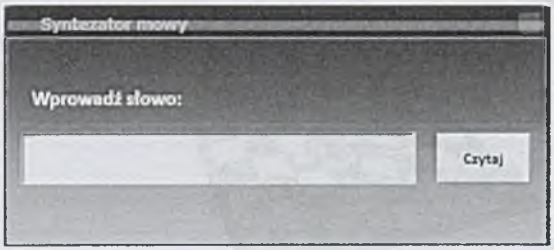
Rys. 3. Przykłady grafiki w programie Expression Design

Przykładem animacji w pakiecie Microsoft Expression jest animacja ruchu wzdłuż ścieżki w Microsoft Expression Blend. W przypadku takiego zadania narzędzie Motion Paths pozwala animować ruch kształtu wzdłuż dowolnej ścieżki. Wykorzystywany jest kształt, którego ruch będzie animowany. Kształtowi użytkownik może nadać nazwę. Animować można dowolny element, nawet kontrolkę lub kontener wypełniony innymi elementami. W ścieżkę ruchu również można zamienić dowolny kształt np. prostokąt, elipsę. Ścieżkę można narysować korzystając z narzędzia Pencil (ołówek). Istotnym etapem animacji jest zamiana narysowanej linii na ścieżkę ruchu Motion Path. Polecenie służące do tego celu to Object > Path > Convert To Motion Path. Po wybraniu poleceniea Convert To Motion Path, pojawi się okno dialogowe, w którym należy zaznaczyć obiekt, który ma być animowany. Poniżej zamieszczono rysunek, na którym widoczny jest panel linii czasu, edytor kodu XAML i struktura aplikacji. Domyślnie animacja ruchu wzdłuż ścieżki trwa 2 sekundy, można jednak zmienić ten czas, klikając i przemieszczając uchwyty wyświetlane na końcach ścieżki ruchu. Kształt linii został zapisany wewnątrz ścieżki ruchu, można więc usunąć narysowaną linię, co nie wpłynie to na sposób przemieszczania się animowanego obiektu. Po użyciu narzędzia bezpośredniej selekcji (biała strzałka w pasku narzędzi) i zaznaczeniu animowanego obiektu pojawią się dodatkowe elementy określające tor ruchu obiektu.



Rys. 4. Animacja ruchu wzdłuż ścieżki w Microsoft Expression Blend

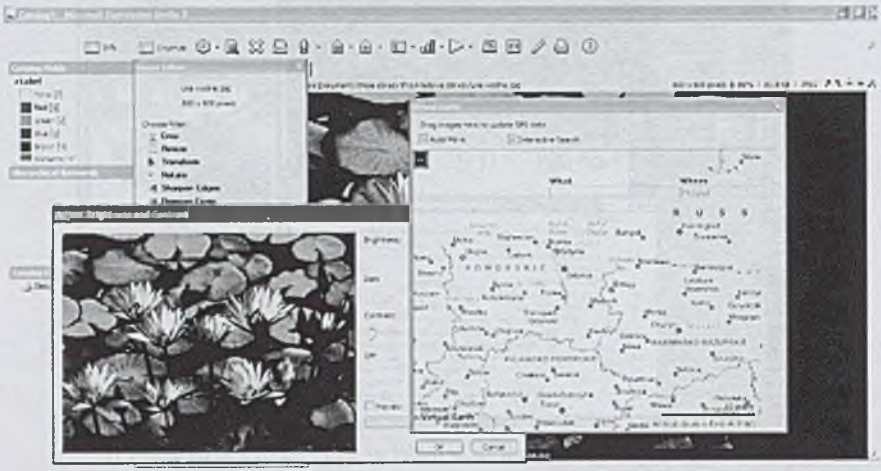
Przykładem ciekawej aplikacji jest syntezator mowy wykonany w pakiecie Microsoft Studio. Grafikę aplikacji opracowuje się w programie Microsoft Blend, rozpoczynając od wyboru WPF Application (Windows Presentation Foundation). Jest to wprowadzony mechanizm prezentacji grafiki WPF, pozwalający wykorzystać zalety nowoczesnego hardware'u graficznego oraz obszernego kompletu zarządzanych klas do tworzenia bogatych wizualnie aplikacji. Ponadto WPF wprowadza nowy język interfejsu użytkownika XAML, który bazuje na XML'u. W trakcie procesu tworzenia aplikacji dalsza praca, po utworzeniu grafiki odbywa się w programie Visual Studio. W programie Visual Studio zostanie wyświetlony kod pliku ***.xaml.cs**. Zadaniem użytkownika jest odpowiednia modyfikacja kodu pliku w programie Visual Studio (dodanie odpowiednich bibliotek, uzupełnienie kodu).



Rys. 5. Przykład aplikacji w pakiecie Microsoft Studio

W programie Microsoft Expression Media użytkownik może realizować podstawowe zadania edycji plików, takie jak obracanie, przycinanie i zmiana rozmiaru, korzystając z rozbudowanych narzędzi tej aplikacji. Odwzorowanie kolorów zostanie zachowane dzięki obsłudze profili zarządzania kolorami. Użytkownik może szybko dopasować poziomy, ostrość, jasność i balans

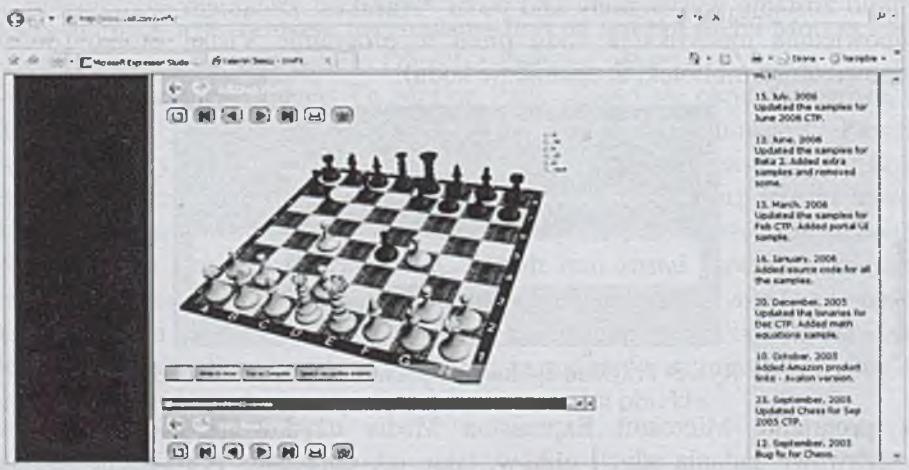
kolorów. Expression Media pozwala wyeksportować zasoby w oparciu o dziesiątki profesjonalnie przygotowanych szablonów prezentacji.



Rys. 6. Edycja plików oraz narzędzie Virtual Earth w Microsoft Expression Media

3. Podsumowanie

Mimo iż technologia jest nowa, można w Sieci znaleźć przykłady jej zastosowań oraz wykorzystania języka XAML [7]. Jednym z najbardziej spektakularnym przykładów jest aplikacja Turning the Pages™ 2.0, która została zbudowana na potrzeby British Library [8].



Rys. 7. Przykład wykorzystania technologii Microsoft Expression [8]

Do komercyjnych zastosowań nowych technologii zaliczyć można czytnik wiadomości z „New York Timesa" [9], szachy zrobione przy użyciu Blend oraz Visual Studio, dostępne wraz z kodami źródłowymi [8].

LITERATURA

1. http://www.idg.pl/news/342397/Expression.Studio.3.0.pierwsze.bety.za.plotem.html
2. http://www.microsoft.com/poland/developer/expression/studio/przewodnik.mspx
3. http://windowshosting.pl/Wyraz.swoje.emocje..Pracujemy.z.Microsoft.Expression.We b
4. http://www.idg.pl/news/342397/Expression.Studio.3.0.pierwsze.bety.za.plotem.html
5. http://www.microsoft.com/poland/developer/expression/web_designer/default.mspx
6. http://www.microsoft.com/poland/developer/expression/media/default.mspx
7. http://webhosting.pl/Microsoft.Expression.Studio._.co.to.takiego.cz..II
8. http://www.turningthepages.com/
9. http://firstlook.nytimes.com/?category_name=times%20reader

Rozdział 30

Systemy informatyczne zarządzania w gospodarce odpadami przedsiębiorstw handlowych w aspekcie realizacji idei rozwoju zrównoważonego

Tomasz Lis, Marek Lis, Konrad Sztumski

Politechnika Częstochowska

tomlis1@wp.pl, lism@el.pcz.czest.pl, konrad@gmail.com

Streszczenie

Niniejszy rozdział przedstawia zagadnienie zastosowania systemów informatycznych zarządzania w gospodarce odpadami przedsiębiorstw handlowych w aspekcie realizacji idei zrównoważonego rozwoju.

1. Uwagi wstępne

Działania związane z organizacją oraz doskonaleniem procesów obsługi klienta, także w sferze obsługi mającej miejsce po dokonaniu przez klienta transakcji zakupu (obsługa posprzedażna), podejmowane są przez coraz większą liczbę przedsiębiorstw działających na rynku. Dzieje się tak z jednej strony, ze względów ekonomicznych, ponieważ polepszenie jakości sprzedaży przyczynia się do poprawy pozycji konkurencyjnej przedsiębiorstwa, a z drugiej strony, ze względów ekologicznych, gdyż ekologia zdobywa coraz większe znaczenie w kształtowaniu wizerunku przedsiębiorstwa. Posiadanie w ofercie towarów nazywanych potocznie „ekologicznymi", czyli przyjaznych dla środowiska, a także umieszczanie ich w opakowaniach „ekologicznych", jest obecnie jednym z istotnych czynników przyciągania klienteli. Ekologiczność firmy jest jakby magnesem przyciągającym klientów.

Równie ważna w tym względzie jest umiejętność organizowania sprawnej obsługi posprzedażnej przez przedsiębiorstwo. Chodzi o to, by zapewniało ono jak najszybsze gospodarowanie odpadami, do których zalicza się opakowania oraz towary zepsute. Obsługa posprzedażna obejmuje cały szereg czynności

związanych z kontaktami przedsiębiorstwa handlowego z klientami, dostawcami, przedsiębiorstwami przetwarzającymi odpady oraz firmami transportowymi. Mają one na celu doprowadzenie do jak najszybszego przekazywania odpadów do odpowiednich odbiorców – albo do dostawcy przedsiębiorstwa handlowego, albo bezpośrednio do firmy zajmującej się przetwarzaniem odpadów.

Warto nadmienić, że „ekologiczne" zarządzanie odpadami, podobnie jak selektywna zbiórka odpadów, wpisuje się w program koncepcji rozwoju zrównoważonego jako jeden z warunków koniecznych do jego realizacji. Wymóg prowadzenia wszelkiej działalności gospodarczej w harmonii z przyrodą w taki sposób, aby nie spowodować w przyrodzie nieodwracalnych zmian wpisany jest w ideę rozwoju zrównoważonego[1]. Właśnie, sprawna i szybka utylizacja opakowań i towarów zepsutych, które mają znaczny udział w zaśmiecaniu naszego otoczenia, zapobiega dokonywaniu się nieodwracalnych zmian w środowisku przyrodniczym. Idea rozwoju zrównoważonego zakłada też zmniejszenie produkcji opakowań do niezbędnego minimum oraz o bezodpadową produkcję dóbr konsumpcyjnych. W pierwszym przypadku chodzi przede wszystkim o zmniejszenia zużycia surowców potrzebnych do produkcji opakowań, a w drugim o to, by zapobiec marnotrawstwu tego, co się już wytworzyło i na co zużyto surowce i energię. Realizacja postulatów wynikających z idei rozwoju zrównoważonego wymusiła uchwalenie szeregu aktów prawnych obowiązujących w Unii Europejskiej. Warunkiem koniecznym do wejścia Polski do Unii Europejskiej było dostosowanie naszego prawa w tym zakresie do przepisów unijnych[1]. Coraz bardziej rygorystyczne prawo ochrony środowiska wymusiło na przemyśle konieczność ograniczania ilości wytwarzanych odpadów, ich utylizacji i zagospodarowania[2].

Dążenie przedsiębiorstw do uzyskania jak najlepszej pozycji konkurencyjnej, powoduje konieczność ciągłego obserwowania rynku przez kadry zarządzające.

---

[1] Np. Ustawa o odpadach z 21 kwietnia 2001 roku zdefiniowała szereg pojęć z zakresu gospodarki odpadami, ale przede wszystkim nałożyła określone obowiązki na wytwórców i posiadaczy odpadów i ustanowiła system sankcji za ich nie przestrzeganie. Ujednolicono kodyfikację odpadów, wprowadzono konieczność ich ewidencjonowania oraz ustalono procedury postępowania z poszczególnymi rodzajami odpadów. Równolegle z tymi przepisami wprowadzono Ustawę z 11 maja 2001 roku nakładającą na producentów produktów w opakowaniach odpowiedzialność za odpady powstające z ich opakowań. Zgodnie z zasadą „zanieczyszczający płaci" przedsiębiorcy zobligowani są do zapewnienia recyklingu i odzysku odpadów powstających z ich opakowań. Ustawa ta wprowadziła roczne poziomy recyklingu i odzysku dla każdego rodzaju opakowania.

[2] W celu uzmysłowienia sobie skali problemy warto przypomnieć, że w 2006 roku w Polsce wprowadzono na rynek ponad 3,5 mln ton opakowań. W większości po produktach szybko zbywalnych, a życie takiego opakowania trwa zwykle kilka miesięcy.

Obserwacje te mają dotyczyć zarówno nowości asortymentowych (nowoczesnych towarów pojawiających się na rynku) jak i działań marketingowych, podejmowanych przez konkurujące podmioty gospodarcze. Analizując możliwość ujęcia nowych towarów we własnej ofercie asortymentowej, decydenci zwracają coraz większą uwagę nie tylko na ich cenę oraz jakość, ale również na materiały, z których są one wykonane ze względu na możliwość ich przetwarzania. To odnosi się szczególnie do opakowań, w jakie pakuje się towary.

Jednym z istotnych czynników, od których zależy skuteczne gospodarowanie odpadami jest właściwe zarządzanie informacjami. Zbyt powolny przepływ informacji pomiędzy organizacjami współpracującymi przy realizacji wybranego zadania często bywa przyczyną poważnych opóźnień. W celu optymalizacji szybkości tego procesu powszechnie korzysta się z systemów informatycznych, które najczęściej połączonych ze sobą siecią elektroniczną.)

2. Gospodarka odpadami w zarządzaniu przedsiębiorstwem

Rozdziały Gospodarka odpadami jest w przedsiębiorstwach ściśle powiązana z tematyką logistyki odwrotnej. „Logistyka odwrotna to proces planowania, wdrażania i kontroli efektywności, efektywności kosztowej przepływu surowych materiałów, zapasów w produkcji, wyrobów gotowych i informacji z nim powiązanych począwszy od punktu konsumpcji do punktu pierwotnego w celu odzyskania wartości lub prawidłowego zagospodarowania[2]." Gospodarka odpadami przynosi przedsiębiorstwom szereg korzyści. Zalicza się do nich między innymi oszczędności kosztowe, ale również polepszenie wizerunku firmy u klientów. Zacieśnienie kontaktów z klientami, a także wyrobienie odpowiedniej renomy są kluczowymi czynnikami pozwalającymi na zdobycie przewagi konkurencyjnej. Klient, który wie, że w przypadku zakupu towaru uszkodzonego, bądź nie spełniającego jego oczekiwań, firma zorganizuje sprawny jego odbiór czy wymianę, przywiązuje się do niej. Chętnie w przyszłości skorzysta z jej oferty ponownie i co ważne wśród znajomych będzie ją reklamował jako podmiot solidny i godny zaufania. Należy stwierdzić, że logistyka odwrotna odpowiadając za gospodarowanie odpadami wpływa bezpośrednio na ekologiczny wizerunek przedsiębiorstwa. Otrzymując wymierne efekty, firma wpływa jednocześnie na ochronę środowiska.

Sprawne zarządzanie odpadami wpływa również na wyrobienie u klientów marki przedsiębiorstwa proekologicznego, co także pozwala na uzyskanie lepszej pozycji konkurencyjnej. Gospodarka odpadami, którą za Ustawą o odpadach definiuje się jako „zbieranie, transport, odzysk i unieszkodliwianie odpadów, w tym również nadzór nad takimi działaniami oraz nad miejscami unieszkodliwiania odpadów[6]" staje się jedną z najważniejszych dziedzin zarządzania przedsiębiorstwem. Jest to wynikiem wzrastającej liczby towarów

sprzedawanych przez przedsiębiorstwa handlowe, skracającego się cyklu życia produktów, chęcią powtórnego wykorzystywania materiałów zużytych na wytworzenie określonego towaru, a także generowania przez społeczeństwa coraz większej ilości śmieci. Gospodarowanie odpadami wymaga od przedsiębiorstw planowania realizacji procesów z tym związanych. Konieczne jest przy tym utrzymywanie stałych kontaktów z klientami (zarówno indywidualnymi jak i podmiotami gospodarczymi), dostawcami (często producentami zainteresowanymi możliwością ponownego wykorzystania materiałów ze zużytego towaru), przedsiębiorstwami transportowymi oraz firmami przetwarzającymi odpady. Taka ilość uczestników procesów gospodarki odpadami wymaga stosowania odpowiednich narzędzi informatycznych. Narzędzi, które pozwolą na ewidencjonowanie, organizowanie przepływu i analizę informacji. Wszystkie zasoby informatyczne przedsiębiorstw współpracujących mają za zadanie wspieranie podejmowanych w nich procesów decyzyjnych.

3. Technologia informacyjna w procesach gospodarowania odpadami

Jak było już wspomniane, w związku z liczbą biorących udział w gospodarce odpadami uczestników, konieczne jest stosowanie elementów technologii informacyjnej. Powszechnie wykorzystuje się w tym celu systemy informatyczne, elementy infrastruktury sieciowej oraz łącza telekomunikacyjne. Systemy informatyczne powinny pozwalać na gromadzenie, przechowywanie oraz analizowanie informacji związanych z zarządzaniem odpadami. Infrastruktura sieciowa i łącza telekomunikacyjne zapewniają możliwość wymiany danych i informacji pomiędzy komputerami, zarówno w obrębie jednej organizacji, jak i pomiędzy współpracującymi podmiotami. W przypadku przepływu informacji wewnątrz jednego systemu informatycznego nie występuje problem różnicy formatów danych, który stanowi z kolei jedną z ważniejszych przeszkód występujących w trakcie współpracy systemów różnych firm. Problem ten musi zostać rozwiązany, jeśli tylko współpracujące podmioty chcą przy użyciu technologii informacyjnej zoptymalizować gospodarkę odpadami. Równie ważnym zagadnieniem jest zapewnienie swobodnego dostępu do potrzebnych informacji. Bez tego warunku nawet najlepszy system informatyczny nie będzie w stanie wspierać procesów podejmowania decyzji w przedsiębiorstwach.

Odpowiednio stosowane narzędzia informatyczne pozwalają znacznie skrócić czas obsługi zlecenia związanego z zagospodarowaniem odpadów. Przedsiębiorstwo handlowe otrzymując informacje o konieczności odebrania od klienta odpadów, w pierwszym kroku próbuje tak zorganizować obsługę zlecenia, aby nie angażować do tego własnych powierzchni magazynowych. W tym celu drogą elektroniczną kontaktuje się z podmiotem, który może być

zainteresowany pozyskaniem odpadów. W przypadku otrzymania odpowiedzi pozytywnej następuje ustalenie szczegółów obsługi złożonego przez klienta zlecenia. Kolejną zaletą jest również możliwość optymalizacji wykorzystywania środków transportu. Przedsiębiorstwa współpracujące dopasowują trasy i obciążenia poszczególnych środków transportów tak, aby były one możliwie maksymalnie angażowane. A to z kolei również przyczynia się do ochrony środowiska. Bowiem im mniejsza liczba pojazdów, tym mniejsza jest ilość wydobywających się spalin.

Infrastruktura informatyczna podmiotów gospodarczych aktywnie i wydatnie wspomaga sugerowany przedsiębiorstwom przez Ustawę sposób postępowania w przypadku zarządzania odpadami. Według Konrada Niziołka „prowadzenie prawidłowej gospodarki odpadami polega na:

- projektowaniu działań na poziomie przedsiębiorstwa,
- prowadzeniu odpowiednich działań na poziomie przedsiębiorstwa, w skład których wchodzi:
- zapobieganie powstawania odpadów i ograniczenie ich ilości, a także ograniczenie negatywnego oddziaływania na środowisko przy wytwarzaniu produktów, podczas i po zakończeniu ich użytkowania,
- zapewnienie odzysku zgodnie z zasadami ochrony środowiska,
- unieszkodliwienie odpadów zgodnie z zasadami ochrony środowiska,
- zbieranie odpadów w sposób selektywny[4]".

Analizując powyższe twierdzimy, że systemy informatyczne doskonale spełniają swą rolę we wszelkich procesach związanych z projektowaniem dowolnych działań. Również w odniesieniu do czynników wymienionych w punkcie 2 systemy te umożliwiają nadzorowanie ilości składowanych odpadów, a także wspomagają procesy dzielenia ich i rozsyłania do odbiorców.

Nie można również pominąć systemów informatycznych nadzorujących procesy utylizacji odpadów, a zainstalowanych w wyspecjalizowanych w tej dziedzinie przedsiębiorstwach.

4. Informatyczne wspomaganie zarządzania w sferze obsługi posprzedażnej - założenia systemu informatycznego

W obszarze obsługi posprzedażnej przeprowadza się procesy związane z zarządzaniem w sferze gospodarowania reklamacjami oraz odpadami. System wspomaga przyjmowanie zleceń, planowanie, a także finalną realizację. Uzyskuje się w ten sposób optymalizację przepływu informacji oraz czynności związanych z przyjmowaniem zlecenia i z jego realizacją. Przewidziana jest organizacja obsługi posprzedażnej w trybie automatycznym, który może być

bezpośredni lub pośredni. Bezpośredni polega na tym, że przedsiębiorstwo handlowe nie wykorzystuje własnych powierzchni magazynowych, a materiały będące podmiotem realizowanej obsługi są przekazywane bezpośrednio do dostawcy – reklamacje, bądź firmy przetwarzającej odpady. Natomiast pośredni polega na pozyskaniu materiałów, ich składowaniu i późniejszym przekazaniu firmie przetwarzającej odpady lub do dostawcy w formie reklamacji. W celu zwiększenia efektywności zarządzania system ma możliwość obsługiwania kontaktów z firmami przetwarzającymi odpady. Przesyła on i wykorzystuje również dane i informacje z oraz do pozostałych obszarów przedsiębiorstwa. Rys.1 przedstawia graficzną prezentację elementu systemu informatycznego wspomagającego zarządzanie w obszarze gospodarowania odpadami (obsługa posprzedażna). Wyszczególniono tu realizowane procesy, a także zależności informacyjne między nimi[3]. Uwzględniono również czynności realizowane w pozostałych obszarach funkcjonalnych przedsiębiorstwa, które mają wpływ na skuteczność gospodarowania odpadami.

Wśród procesów z zakresu obsługi posprzedażnej, w realizacji których przedstawiany element systemu informatycznego bierze czynny udział, wymienić można:

- przyjęcie zgłoszenia,
- identyfikację produktów oraz weryfikację zgłoszenia,
- organizację procesów posprzedażnych,
- ustalanie potrzeb organizowania procesów posprzedażnych,
- obsługę firm przetwarzających odpady,
- obsługę dyspozycji procesów posprzedażnych.

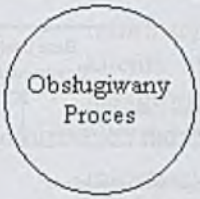
Aby przedstawiony na Rys. 1 element systemu informatycznego zarządzania mógł prawidłowo wypełniać stawiane mu zadania, musi korzystać z zasobów informacyjnych pozostałych obszarów funkcjonalnych przedsiębiorstwa. Z kolei on sam także generuje wiele informacji im potrzebnych. Opisywany element systemu posiada połączenia informacyjne z procesami:

- nadzoru transportu – wymiany informacji mającej na celu zorganizowanie czynności transportowych dla odpadów, bądź towarów reklamowanych,
- planowania procesów transportowych – wymiana informacji, służąca zaplanowaniu przez przedsiębiorstwo czynności transportowych dla odpadów bądź towarów reklamowanych,
- nadzoru zamówień i przyjęć magazynowych – informacje przesyłane z magazynu, a dotyczące wystąpienia w dostawie towarów uszkodzonych, które należy w ramach reklamacji odesłać do dostawcy,
- przygotowania kompletowania zleceń – z obszaru obsługi posprzedażnej do magazynu przesyłane są informacje na temat

odpadów i/bądź towarów reklamowanych, które należy wydać, w celu ich przetransportowania,

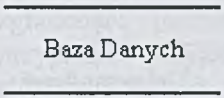
- kompletowania zlecenia – z magazynu do obsługi posprzedażnej zostają przesłane informacje o zrealizowaniu czynności wydania odpowiednich odpadów i/bądź towarów reklamowanych.

Przy opracowaniu korzystano z ogólnych zasad metodyki Yourdona – analizy i projektowania systemów informatycznych[5]. Do budowy modelu wykorzystano zbiór elementów graficznych, które oznaczają odpowiednio:



- proces obsługiwany przez system informatyczny

- obsługiwany proces wraz z symbolem oznaczającym jego przynależność do określonego obszaru zarządzania logistycznego



- wykorzystywana w systemie informatycznym baza danych

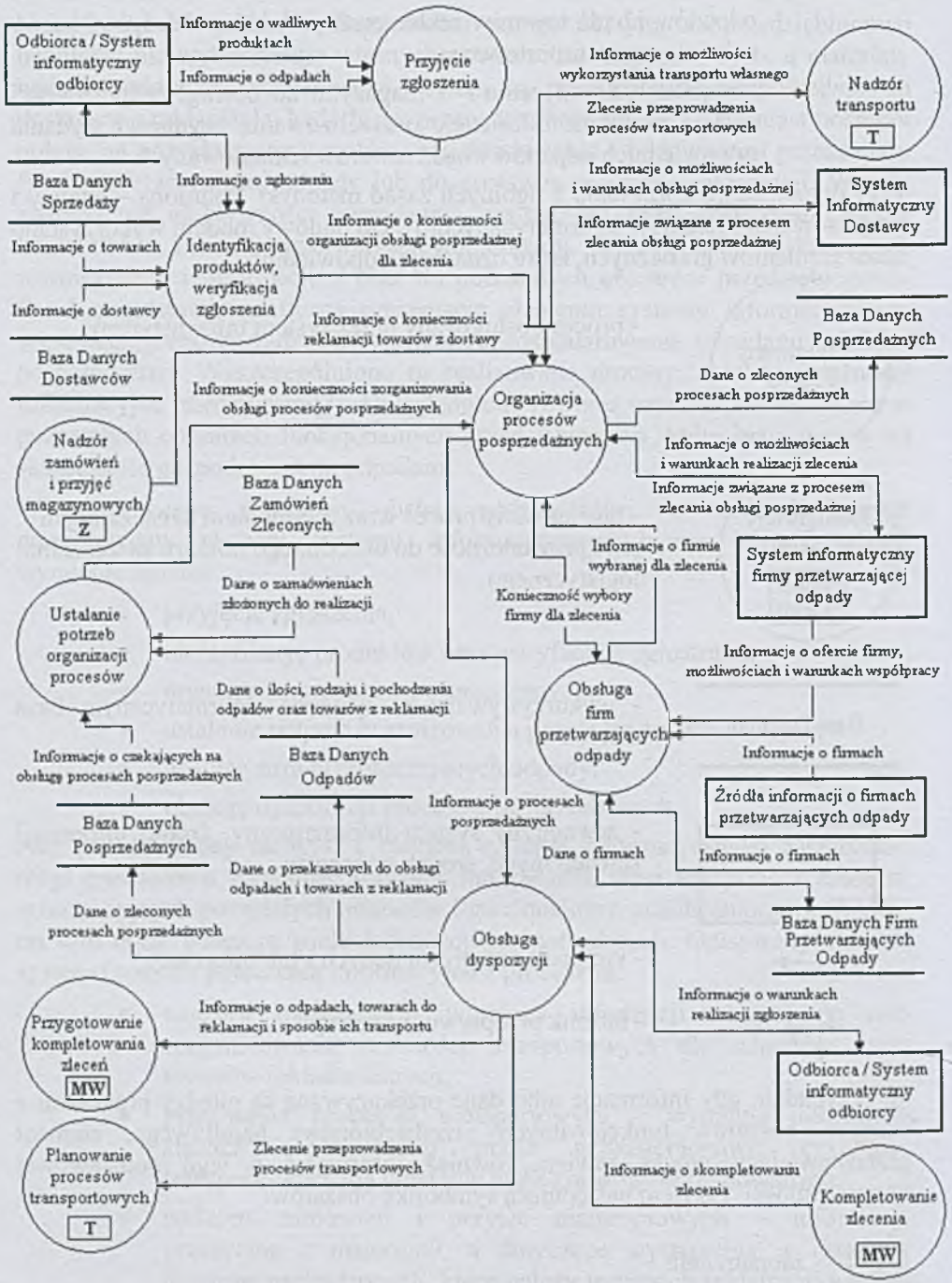
- zewnętrzny system informatyczny, źródło informacji zewnętrznych, środek transportu

- kierunek przepływu danych i/lub informacji

- łącznik przepływu informacji

W przypadku, gdy informacje albo dane przekazywane są między procesami z różnych obszarów funkcjonalnych przedsiębiorstwa handlowego, element przedstawiający proces zawiera również informację o jego podstawowej przynależności. Przyjęto następującą symbolikę obszarów:

Z - zaopatrzenie
MW - magazyn wydawanie towarów
T - transport



Rys. 1. Obsługa posprzedażna

Przedstawione na rysunku 1 procesy odpowiedzialne są za realizację następujących czynności[3]:

1. Przyjęcie zgłoszenia – na tym etapie, do systemu informatycznego zostają przekazane informacje o uszkodzeniach, które wystąpiły w towarach objętych terminem reklamacji, a także pojawieniu się odpadów, przetworzeniem których, może, bądź powinno zająć się przedsiębiorstwo handlowe.

Informacje i dane wchodzące:

- informacje o wadliwych produktach – dostarczane z systemu informatycznego klienta, związane są z wystąpieniem wadliwych towarów – zakupionych wcześniej z przedsiębiorstwa handlowego,
- informacje o odpadach – dostarczane z systemu informatycznego klienta, związane są z koniecznością pozyskania przez przedsiębiorstwo handlowe odpadów.

Informacje i dane wychodzące:

- informacje o zgłoszeniu – są przekazywane do procesu identyfikacja produktów, weryfikacja zgłoszenia, w obszarze obsługa posprzedażna.

2. Identyfikacja produktów oraz weryfikacja zgłoszenia – następuje tu identyfikacja materiałów, pod względem ich pochodzenia – identyfikuje się firmy, które dostarczyły je do przedsiębiorstwa handlowego. Informacje o uszkodzeniach pozwalają na zweryfikowanie zgłoszenia, pod kątem zasadności organizacji reklamacji. W razie potrzeby przeprowadzana jest dodatkowa wymiana informacji pomiędzy przedsiębiorstwem handlowym, a firmą zgłaszającą potrzebę zorganizowania procesów obsługi posprzedażnej. W przypadku przyjęcia zlecenia, przedsiębiorstwo handlowe pobiera drogą elektroniczną od zleceniodawcy wymagane dokumenty. Informacje i dane wchodzące:

- informacje o zgłoszeniu, kontakt z klientem – są przekazywane z procesu przyjęcie zgłoszenia, z obszaru obsługa posprzedażna. Zawierają informacje o zgłoszonym przez klienta wystąpieniu wadliwych produktów oraz konieczności zagospodarowania odpadów. Precyzują rodzaj towarów, datę i numer ich dostawy. W przypadku odpadów – z jakimi towarami są związane, kiedy te towary zostały zakupione, numer dostawy. Wykorzystuje się je do identyfikacji podmiotów zlecenia,
- informacje o zgłoszeniu, przekazywane elektronicznie – przekazywane z systemu informatycznego klienta, związane są z wymianą dodatkowych informacji o zgłoszeniu – wymaganych w celu zorganizowania obsługi,
- informacje o towarach – ich źródłem są dane pozyskiwane z bazy danych sprzedaży. Służą do identyfikacji towarów oraz dostawy, z którym są związane. Pozwalają również na weryfikację zgłoszenia,
- informacje o dostawcy – ich źródłem są dane pozyskiwane z bazy danych dostawców. Służą do identyfikacji dostawcy towarów, uzgodnionych z nim warunków reklamacji i zarządzania odpadami.

Informacje i dane wychodzące:

- informacje o zgłoszeniu – przekazywane do systemu informatycznego klienta, związane są z wymianą dodatkowych informacji o zgłoszeniu – wymaganych w celu zorganizowania obsługi,
- informacje o konieczności organizacji obsługi posprzedażnej dla zlecenia – są przekazywane do procesu organizacja procesów posprzedażnych, w obszarze obsługa posprzedażna.

3. Organizacja procesów posprzedażnych – wraz z otrzymaniem informacji o konieczności zorganizowania obsługi dla zlecenia, system informatyczny sprawdza możliwość przeprowadzenia procesów posprzedażnych. W tym celu, w zależności od rodzaju materiałów – towary do reklamacji, odpady, przedsiębiorstwo handlowe drogą elektroniczną kontaktuje się z dostawcą, bądź firmą przetwarzającą odpady i ustala warunki realizacji zlecenia. W przypadku kiedy zgłoszenie nie dotyczy materiałów składowanych na magazynie przedsiębiorstwa handlowego, poszukiwane jest rozwiązanie pozwalające na bezpośrednie ich przekazanie do dostawcy/firmy przetwarzającej odpady. Każdorazowo, w razie potrzeby, zleca się dobranie firmy przetwarzającej odpady. Sprawdzana jest również, możliwość zaangażowania do obsługi zlecenia, własnych środków transportu, wykonujących wcześniej przypisane im zadania.

Informacje i dane wchodzące:

- informacje o konieczności organizacji obsługi posprzedażnej dla zlecenia – są przekazywane z procesu identyfikacja produktów, weryfikacja zgłoszenia, z obszaru obsługa posprzedażna. Zawierają informacje o kliencie, który zgłosił potrzebę zorganizowania obsługi posprzedażnej, a także towarach i odpadach, których ma ona dotyczyć. Wykorzystuje się je do planowania obsługi,
- informacje o możliwości wykorzystania transportu własnego – są przesyłane z procesu nadzór transportu, z obszaru transport. Wskazują na możliwość zaangażowania transportu własnego. Wykorzystywane są do planowania i organizowania procesów posprzedażnych,
- informacje o możliwościach obsługi posprzedażnej – są przesyłane z systemu informatycznego dostawcy,
- ustalanie warunków obsługi posprzedażnej – są przesyłane z systemu informatycznego dostawcy,
- informacje o możliwościach obsługi posprzedażnej – są przesyłane z systemu informatycznego firmy przetwarzającej odpady,
- ustalanie warunków obsługi posprzedażnej – są przesyłane z systemu informatycznego firmy przetwarzającej odpady,
- informacje o firmie wybranej dla zlecenia – są przesyłane z procesu obsługa firm przetwarzających odpady, z obszaru obsługa posprzedażna. Zawierają informacje o firmach przetwarzających

odpady, które są w stanie sprostać wymaganiom organizowanego procesu. Wykorzystywane są do wyboru firmy do obsługi zlecenia,

- informacje o konieczności zorganizowania obsługi procesów posprzedażnych – są przesyłane z procesu ustalanie potrzeb organizacji procesów, z obszaru obsługa posprzedażna. Informują o konieczności zorganizowania procesów obsługi posprzedażnej, dla towarów i odpadów zgromadzonych na magazynie, bądź dla zgłoszonych zleceń z wydłużonym terminem realizacji. Informują również o możliwości zaangażowania środków transportu dostawcy – np. do pobrania towarów do reklamacji. Wykorzystywane są do planowania i organizowania procesów posprzedażnych,
- informacje o konieczności reklamacji towarów z dostawy – są przekazywane z procesu nadzór zamówień i przyjęć magazynowych, z obszaru zaopatrzenie. Informują o wystąpieniu wadliwych towarów w przyjętej dostawie. Wskazują ich pochodzenie i ilość. Wymuszają zainicjowanie organizacji procesów posprzedażnych – reklamacje.

Informacje i dane wychodzące:

- zlecenie przeprowadzenia procesów transportowych – są przekazywane do procesu nadzór transportu, w obszarze transport,
- zlecenie obsługi posprzedażnej – są przekazywane do systemu informatycznego dostawcy,
- dane o zleconych procesach posprzedażnych – zapisywane w bazie danych posprzedażnych opisują zlecone do realizacji procesy obsługi,
- informacje związane z procesem zlecania obsługi posprzedażnej – są przekazywane do systemu informatycznego firmy przetwarzającej odpady,
- konieczność wyboru firmy dla zlecenia – są przekazywane do procesu obsługa firm przetwarzających odpady, w obszarze obsługa posprzedażna,
- informacje o procesach posprzedażnych – są przekazywane do procesu obsługa dyspozycji, w obszarze obsługa posprzedażna.

4. Ustalanie potrzeb organizowania procesów posprzedażnych – w tym procesie kontroluje się ilość znajdujących się na magazynie materiałów przeznaczonych do obsługi posprzedażnej. Monitoruje się także utworzone w dziale zamówień zlecenia - pod kątem możliwości wykorzystania środków transportu dostawcy, do pobrania towarów do reklamacji. Kontroli podlegają również czekające na realizację procesy obsługi posprzedażnej. W przypadku kiedy ilość towarów do reklamacji, bądź odpadów, osiągnie pewną wartość graniczną, lub istnieje

możliwość wykorzystania środków transportu dostawcy, zgłaszana jest potrzeba organizacji procesów posprzedażnych. Informacje i dane wchodzące:

- dane o zamówieniach złożonych do realizacji – są pobierane z bazy danych zamówień zleconych. Informują o dostawcach, u których złożono zamówienie. Wykorzystuje się je do sprawdzenia możliwości zaangażowania środków transportu dostawcy, do pobrania towarów z reklamacji, bądź odpadów,
- dane o ilości, rodzaju i pochodzeniu odpadów oraz towarów z reklamacji – są pobierane z bazy danych odpadów. Informują o rodzaju i ilości towarów, w stosunku do których należy przeprowadzić procesy reklamacji, a także ich pochodzeniu – dostawcy. Informują również o zgromadzonych na magazynie odpadach. Wykorzystuje się je do zlecania organizacji procesów posprzedażnych,
- informacje o czekających na obsługę procesach posprzedażnych – pozyskiwane z bazy danych posprzedażnych dane, informują o zgłoszonych przez klientów zleceniach zorganizowania przez przedsiębiorstwo handlowe obsługi posprzedażnej - z wydłużonym terminem realizacji. Wykorzystuje się je do zlecania organizacji procesów posprzedażnych.

Informacje i dane wychodzące:

- informacje o konieczności zorganizowania obsługi procesów posprzedażnych – są przekazywane do procesu organizacja procesów posprzedażnych, w obszarze obsługa posprzedażna.

5. Obsługa firm przetwarzających odpady – poszukuje się tu firm zajmujących się przetwarzaniem odpadów, które mogą być przydatne, w związku z posiadaną przez przedsiębiorstwo handlowe ofertą asortymentową. Dane o firmach zapisywane są w bazie danych. Do innych realizowanych tu czynności, zalicza się wybór firmy dla zlecenia. Informacje i dane wchodzące:

- konieczność wyboru firmy dla zlecenia – z procesu organizacja procesów posprzedażnych, z obszaru obsługa posprzedażna przesyłane są informacje, których celem jest zainicjowanie czynności zmierzających do wyboru firmy dla przygotowywanego zlecenia,
- informacje o firmach – dane pobierane z bazy danych firm przetwarzających odpady pozwalają na pozyskanie informacji o działających na rynku firmach przetwarzających odpady. Wykorzystuje się je do wybrania firm dla organizowanej obsługi,
- informacje o ofercie firmy, możliwościach i warunkach współpracy – są pozyskiwane z systemu informatycznego firmy przetwarzającej odpady,

- informacje o firmach – są pozyskiwane z zewnętrznych źródeł informacji.
- Informacje i dane wychodzące:
- informacje o firmie wybranej dla zlecenia – są przekazywane do procesu organizacja procesów posprzedażnych, w obszarze obsługa posprzedażna,
- dane o firmach – są zapisywane w bazie danych firm przetwarzających odpady. Opisują firmy, których działalność polega na przetwarzaniu odpadów.

6. Obsługa dyspozycji procesów posprzedażnych – dokonuje się tu między innymi wymiany informacji z przedsiębiorstwem zgłaszającym potrzebę zainicjowania procesów obsługi posprzedażnej. Jeżeli zlecenie zostało przekazane do realizacji, informacje dotyczą sposobu jego przeprowadzenia, w przypadku braku możliwości automatycznej organizacji obsługi, ustalany jest akceptowany przez zleceniodawcę termin jej wykonania. Oprócz zarządzania przepływem informacji, organizuje się obsługę procesów posprzedażnych, w zakresie czynności związanych z wydawaniem towarów - do reklamacji i odpadów z magazynu, a także planowaniem procesów transportowych.

Informacje i dane wchodzące:

- informacje o procesach posprzedażnych – są dostarczane z organizacji procesów posprzedażnych, z obszaru obsługa posprzedażna. Informacje dotyczą czynności jakie muszą zostać podjęte w przedsiębiorstwie handlowym, w celu obsługi zlecenia. Wykorzystuje się je do organizowania procesów posprzedażnych,
- informacje o warunkach obsługi zgłoszenia – są pozyskiwane z systemu informatycznego klienta,
- informacje o skompletowaniu zlecenia – są dostarczone z procesu kompletowanie zlecenia, z obszaru magazynowanie – kompletowanie zamówień. Celem wykorzystania informacji, jest potwierdzenie zakończenia czynności wydawania towarów, bądź odpadów z magazynu i rozpoczęcia procesów dostawczych – uaktualnienie danych w bazie danych posprzedażnych,
- dane o zleconych procesach posprzedażnych – pozyskiwane z bazy danych posprzedażnych, informują o ustaleniach dotyczących przeprowadzenia obsługi posprzedażnej. Wykorzystuje się je do organizacji procesów.

Informacje i dane wychodzące:

- informacje o warunkach obsługi zlecenia – są przekazywane do systemu informatycznego klienta,
- zlecenie przeprowadzenia procesów transportowych – są przesyłane do planowania procesów transportowych, w obszarze transport,

- informacje o odpadach, towarach do reklamacji i sposobie ich transportu – są przekazywane do procesu przygotowanie kompletowania zleceń, w obszarze magazynowanie – kompletowanie zamówień,
- dane o zleconych procesach posprzedażnych – są zapisywane w bazie danych posprzedażnych – potwierdzenie pobrania towarów, bądź odpadów z magazynu i rozpoczęcia realizacji zlecenia.

5. Podsumowanie

Gospodarka odpadami odgrywa coraz większą rolę w zarządzaniu przedsiębiorstwami handlowymi. Dzieje się tak, ponieważ pojawia się coraz więcej nowych towarów, a także dlatego, że przedsiębiorstwa w celach oszczędnościowych dążą do wielokrotnego wykorzystywania materiałów, z których zrobiony został określony towar. Duże znaczenie odgrywa przy tym chęć uzyskania miana przedsiębiorstwa proekologicznego, tzn. takiego, które w swych działaniach uwzględnia problematykę ochrony środowiska i wymogi rozwoju zrównoważonego. Jest to jeden ze relewantnych sposobów polepszania pozycji konkurencyjnej firmy na rynku. Ze względu na podobieństwo oferowanych towarów przez różne podmioty gospodarcze, zarówno ze względu na cenę, własności użytkowe jak i przeznaczenie, decydenci w taki właśnie sposób próbują pozyskiwać nowych klientów i przywiązywać do siebie tych, którzy już dokonali u nich zakupu. Gospodarowanie odpadami jest również jednym z wymogów stawianych przedsiębiorstwom handlowym przez producentów, a wynikających z prawa o ochronie środowiska i obowiązku realizacji rozwoju zrównoważonego.

W związku z tym, że w procesie obrotu odpadami bierze udział wielu uczestników, konieczne jest stosowanie elementów technologii informacyjnej. Pozwalają one na skrócenie czasu koniecznego do przepływu informacji. Tym samym przyczyniają się one wydatnie do optymalizacji gospodarki odpadami. Istotne znaczenie ma dopasowanie systemów informatycznych różnych firm pod względem formatu obsługiwanych danych oraz zapewnienie za pomocą odpowiednich łącz telekomunikacyjnych możliwości swobodnej ich współpracy w płaszczyźnie elektronicznej przy użyciu infrastruktury sieciowej. Zaprezentowano tu model jednego z elementów systemu informatycznego, którego zadaniem jest wspomaganie zarządzania obsługi posprzedażnej (gospodarki odpadami) w małych i średnich przedsiębiorstwach handlowych. Dzięki niemu można uzyskać zamierzony efekt pod warunkiem, że ten model współpracuje z modułami zainstalowanymi w pozostałych obszarach funkcjonalnych danego przedsiębiorstwa handlowego.

LITERATURA

1. E. den Boer, J. den Boer, J. Jager, I. Maćkow, M. Sebastian, R. Szpadt, Planning of munnicipal waste management systems using life cykle analysis, W: VI-th International Waste Forum „Efficiency of Waste Management", Poznań 2005
2. Grabara J. K., Starostka-Patyk M., Współpraca w logistyce odwrotnej, W: Informatyka w zarządzaniu logistyką, pod red. Janusza K. Grabary, Polskie Towarzystwo Informatyczne – Oddział Górnośląski, Katowice 2006, s. 144
3. Lis T., Systemy informatyczne w zarządzaniu logistycznym przedsiębiorstwami handlowymi, praca doktorska, Politechnika Częstochowska, Wydział Zarządzania, Częstochowa 2006
4. Niziołek K., Gospodarowanie odpadami w systemach zarządzania środowiskowego przedsiębiorstw produkcyjnych, W: Komputerowo Zintegrowane Zarządzanie, T. II, pod red. Ryszarda Knosali, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole 2007, s. 98
5. Roszkowski J., Analiza i projektowanie strukturalne, Wydawnictwo Helion, Gliwice, 1998 r.
6. Ustawa o odpadach [2, art. 3, p. 3, u. 1]

Część 4.
Pozostałe zagadnienia

Rozdział 31

Analiza pojęć „nowej, innowacyjnej, nowoczesnej" technologii

Leszek Grocholski
Instytut Informatyki Uniwersytetu Wrocławskiego
Leszek.Grocholski@ii.uni.wroc.pl

Andrzej Niemiec
Prim Sp. z o.o.
ani@prim.com.pl

Streszczenie

W rozdziale omówiono praktyczne aspekty - możliwości uzyskania dofinansowania czy też ulgi jakie może przynieść odpowiednio inwestycja w „nową, innowacyjną" czy też „nowoczesną" technologię. W/w pojęcia, pozornie bardzo zbliżone do siebie, różnią się z punktu widzenia aktów prawnych w których są użyte. Opinie dotyczące tych pojęć może wydawać m in. uczelnia wyższa i w ograniczonym zakresie Polskie Towarzystwo Informatyczne. W opiniach należy udowodnić zgodność przedmiotu opinii z odpowiednimi „legalnymi" definicjami. Przedstawiono korzyści z posiadania opinii i ideę oceny „nowej technologii" w sensie ustawy z dnia 25 lutego 1992r., o podatku dochodowym od osób prawnych. W rozdziale omówiono też pojęcie „innowacyjności technologii" w sensie opinii o innowacyjności. Opinia taka jest wymagana przez Polską Agencję Rozwoju Przedsiębiorczości (PARP) i regionalne instytucje pośredniczące w niektórych konkursach na dofinansowania z funduszów regionalnych i Unii Europejskiej. Instytucje te wymagają analizy „innowacyjności technologii" wykorzystanej w projekcie. W rozdziale również, zgodny z delicjami innowacyjności przyjętymi przez PARP i Unię Europejską sposób oceny „innowacyjnej technologii" i projektu. Na zakończenie opisano, zgodną z rozporządzeniem Ministra

Pracy i Polityki Społeczne z dnia 19 grudnia 2007r. w sprawie zakładowego funduszu rehabilitacji osób niepełnosprawnych, możliwość uzyskania przez zakład pracy chronionej odpisu z zakładowego funduszu rehabilitacji osób niepełnosprawnych „na wprowadzenie nowoczesnej technologii i prototypowych wzorów".

1. Wstęp

W rozdziale przedstawiono poparta doświadczeniem analizę pojęć „nowej, innowacyjnej i nowoczesnej" technologii w odniesieniu do technologii informacyjno komunikacyjnej. Analiza została przeprowadzona na postawie szeregu opinii sporządzanych przez autorów, którzy są:

- pracownikami wyższych uczelni,
- rzeczoznawcami Polskiego Towarzystwa Informatycznego,
- biegłymi sądowymi.

Opinie sporządzane były w celu uzyskania przez inwestora odpowiedniej korzyści finansowej. Korzyść ta – odliczenie od podstawy opodatkowania, zwrot części kosztów, finansowanie z funduszu może być uzyskana po przedstawieniu odpowiedniej opinii. Opinie takie mogą być wydawane wyłącznie przez uprawnione instytucje.

Dla osoby posiadającej wykształcenie wyższe techniczne i matematyczne logiczne jest przyjęcie, ze w opinii należy pokazać zgodność z przedmiotu analizy z odpowiednimi definicjami. Problem polega na tym, że rzadko które definicje są dla ścisłego umysłu precyzyjne.

Stosowane w Polsce pojęcia „technologia informacyjna" (ang. IT – Information Technology) i nowsze „technologia informacyjno komunikacyjna" (ang. ICT – Information and Communiction Technology) stanowią kalkę pojęć angielskich. Jednak tłumaczenie angielskiego słowa technology jako technologia wywoływało w przeszłości dyskusje. Stosowanie w polskiej informatyce pojęcia „technologia informacyjna" jeszcze parę lat temu było kontrowersyjne. Przeciwnicy tego pojęcia twierdzili, że słowo technologia może być używane jedynie w kontekście sposobu, metody przetwarzania materii a w informatyce nie przetwarza się materii. Zwolennicy pojęcia powołują się na to, że informatyka dotyczy przetwarza informacji, która posiada pewne cechy materii (np. wartość, konieczność przechowywania, przesyłania, przekształcania). Obecnie powszechnie używa się pojęcia technologii informacyjnej oraz technologii informacyjno komunikacyjnej. Technologie informacyjne (a nie nt. technika informatyczna, czy przetwarzanie danych) są nawet przedmiotem nauczanym w gimnazjum i liceum oraz znajdują się w siatkach przedmiotów na studiach wyższych – kierunkach nieinformatycznych. W takiej sytuacji oczywista jest powszechne funkcjonowanie definicji z podręcznika szkoły średniej, np.[1] : *„Technologia informacyjna (ang. Information Technology) jest*

zespołem środków (tj. urządzeń), narzędzi (tj. programów) oraz innych technologii, służących wszechstronnemu posługiwaniu się informacją i łączeniu zastosowań z wieloma innymi technikami pokrewnymi."

Ponieważ informatyka jest coraz bardziej związana z telekomunikacją pojęcie „technologii informacyjnej" jest wypierane przez pojęcie „technologia informacyjno komunikacyjna". Angielski termin ICT - Information and Communiction Technology jest powszechnie używany w Unii Europejskiej więc jego polski odpowiednik jest powszechnie używany przez rząd polski i jego organy. Definicja tej technologii jest podana m. in. w znajdującym się na stronie WWW Polskiej Agencji Rozwoju Przedsiębiorczości słowniku [2]:

„*TECHNOLOGIE INFORMACYJNE I KOMUNIKACYJNE – TIK [ang. Information and Communication Technology – ICT] - to najogólniej rzecz ujmując narzędzia pozwalające na komunikację między ludźmi. Technologie informacyjno-komunikacyjne, nazywane też technologiami informacyjnymi (IT), są technologiami związanymi ze zbieraniem, przechowywaniem, przetwarzaniem, przesyłaniem, rozdzielaniem i prezentacją informacji (tj. tekstów, obrazów, dźwięku). Obejmują one w szczególności technologie komputerowe (sprzęt i oprogramowanie) i technologie komunikacyjne. .....*"

Jak widać z przytoczonych definicji:

technologia informacyjna to w zasadzie to samo co informacyjno komunikacyjna,

obejmuje ona swoim zakresem zarówno oprogramowanie jak i sprzęt.

2. ICT jako nowa technologia

Państwo chcąc stymulować wdrażanie nowych technologii wprowadziło ulgę podatkową z nią związaną - możliwość odliczenia bezpośrednio od podstawy opodatkowania CIT do 50% kosztów zakupu nowej technologii. Odliczenia co do zasady dokonuje się w zeznaniu za rok podatkowy, w którym technologię wprowadzono do ewidencji środków trwałych/ wartości niematerialnych i prawnych lub w następnych 3 latach podatkowych (w przypadku osiągnięcia straty lub zbyt małego dochodu).

Artykuł 18 ust. 2 ustawy z dnia 25 lutego 1992 r., O podatku dochodowym od osób prawnych z późniejszymi zmianami (Dz. U. z 2008 r. nr 169, poz. 1049) zawiera definicję legalną pojęcia „nowe technologie": „*Za nowe technologie, uważa się, wiedzę technologiczną w postaci wartości niematerialnych i prawnych, która umożliwia wytwarzanie nowych lub udoskonalonych wyrobów i usług i która nie jest stosowana na świecie przez okres dłuższy niż ostatnich 5 lat, co potwierdza opinia niezależnej od podatnika jednostki naukowej w rozumieniu ustawy z dnia 8 października o 2004 r. o zasadach finansowania nauki.*"

Definicja ta podaje warunki jakie powinny spełniać zakupione oprogramowanie i sprzęt aby mogły być uznane jako „nowe" technologie:

a. Wiedza technologiczna w postaci wartości niematerialnych i prawnych.
b. Umożliwienie wytwarzania nowych lub udoskonalonych wyrobów i usług.
c. Stosowanie w świecie nie dłużej niż ostatnie 5 lat.
d. Potwierdzenie w postaci opinii niezależnej jednostki naukowej.

Jakie zadania stoją przed uczelnią wyższą, która ma potwierdzić spełnienie przez sprzęt i oprogramowanie tych warunków?

Ad a. Oprogramowanie i sprzęt stanowią niewątpliwe kwintesencje posiadanej przez wytwórcę wiedzy i doświadczenia w określonej dziedzinie, która jest niedostępna do samodzielnego zdobycia przez nabywcę. Wiedza ta reprezentowana przez budowę i sposób działania oprogramowania i sprzętu. Nabywca uzyskuje prawo do korzystania z tej wiedzy. Zakup nowej technologii powinien dotyczyć wartości niematerialnych i być potwierdzony patentem, czy licencją. Ponieważ oprogramowanie jest z reguły sprzedawane jako licencja to warunek ten jest łatwy do spełnienia, gorzej ze sprzętem, który klasyfikowany prawie zawsze jako środek trwały

Ad. b. Sprzęt i oprogramowanie ma służyć do wytwarzania nowych lub udoskonalonych wyrobów i usług. Jeżeli zakup umożliwia wytwarzanie nowych lub zmodyfikowanych wyrobów to należy je opisać. Podobnie w przypadku nowych usług. A jak należy rozumieć modyfikację usług? Po pierwsze to należy zauważyć, że usługi mogą być zarówno zewnętrzne jak i wewnętrzne. Np. jeżeli firma posiada dział księgowości to księgowanie jest usługą wewnętrzna. Modyfikacji może podlegać usługa księgowania. Po drugie modyfikacja może dotyczyć: sposobu wprowadzania i udostępniania danych, zwiększenia szybkości, zmniejszeniu możliwości popełniania błędów polegającej na integracji z innym oprogramowaniem.

Ad. c. Stosowanie w świecie nie dłużej niż 5 lat. Patent posiada datę rejestracji a licencja podaje tylko datę zakupu. W tym przypadku można się posłużyć datą wprowadzenia oprogramowania na rynek, która z reguły znajduje się w dostępnych opisach technicznych. Ta metoda może dotyczyć rozwiązań, co do których możliwe jest zdobycie informacji o ich stosowaniu w świecie nie dłużej niż 5 lat. Ale wspomaganie przez ICT z reguły dotyczy funkcjonalności znanych w świecie dłużej niż 5 lat np. zasady księgowania są znane od średniowiecza a dostęp przez Internet już kilkanaście lat. W tym przypadku można się odwołać do daty wprowadzenia na technologii wytworzenia danego oprogramowania czy sprzętu, która jest zwykle podawana przez producenta w opisach technicznych. Technologia ta zawsze determinuje nowe własności produktu: możliwość działania w danych standardach, współpraca z innymi komponentami, bardziej efektywne wykorzystanie zasobów czy szybkość działania. Jest oczywiste, że jeżeli dana technologia jest stosowana w świecie

nie dłużej niż 5 lat, to dotyczy to również wytworzonego w tej technologii produktu.

Reasumując stosunkowo łatwo jest uzasadnić zakup oprogramowania jako "nowej technologii" gdyż:

- związana jest z nią licencja stanowiąca własność niematerialna i prawną,
- jeżeli nie służy ono do wywarzania nowych produktów czy usług to z reguły usprawnia świadczenie usług zewnętrznych czy wewnętrznych,
- jeżeli nie jest oprogramowanie systemowym i narzędziowym wprowadzonym na rynek w ostatnich 5 latach to z reguły wytworzono je w technologii znanej i stosowanej nie dłużej niż 5 lat.

Jeżeli sprzęt zakupiono jako własność niematerialną i prawną to też można go – stosując powyższy schemat uznać go jako „nową" technologię. W przypadku zakupu sprzętu jako środek trwały jest to niemożliwe.

3. Kredyt na zakup nowej technologii

Kredyt technologiczny jest to kredyt udzielany na realizacje inwestycji technologicznej związanej z wdrożeniem nowej technologii, częściowo spłacany w formie premii technologicznej. Nie może być udzielany na zakup maszyn i urządzeń, w których została wdrożona technologia, będąca przedmiotem inwestycji technologicznej. W zależności od statusu przedsiębiorcy i województwa, w którym realizuje inwestycje, wnioskodawca może liczyć na premię technologiczną w wysokości odpowiadającej od 40 do 70 % kosztów kwalifikowanych, sfinansowanych kredytem technologicznym. Kredyt technologiczny został wprowadzony ustawą z dnia 30 maja 2008 r. O niektórych formach wspierania działalności innowacyjnej. Poniżej przestawiono pochodzące z ustawy, podstawowe definicje dot. inwestycji technologicznej, która jest bezpośrednio związana z pojęciem „nowej technologii".

„Działalność innowacyjna – działalność polegającą na opracowaniu nowej technologii i uruchomieniu na jej podstawie wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług.
Kredyt technologiczny – kredyt udzielany przedsiębiorcy przez bank kredytujący na realizację inwestycji technologicznej, który jest częściowo spłacany ze środków Funduszu Kredytu Technologicznego w formie premii technologicznej, do wysokości i na warunkach określonych w ustawie;
Inwestycja technologiczna – inwestycja polegającą na:
a) zakupie nowej technologii, jej wdrożeniu oraz uruchomieniu na jej podstawie wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług albo
b) wdrożeniu własnej nowej technologii oraz uruchomieniu na jej podstawie wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług.

Nowa technologia – technologia w postaci prawa własności przemysłowej lub usługi badawczo-rozwojowej (w rozumieniu Polskiej Klasyfikacji Wyrobów i Usług), która umożliwia wytwarzanie nowych lub znacząco ulepszonych towarów, procesów lub usług i nie jest stosowana na świecie dłużej niż 5 lat.
Podstawowym aktem prawnym dot. prawo własności przemysłowej jest ustawa z dnia 30 czerwca 2000 r. Prawo własności przemysłowej.
Zgodnie z polskim prawem do przedmiotów własności przemysłowej zaliczamy:
- projekty wynalazcze, w tym: wynalazki, wzory użytkowe, wzory przemysłowe,
topografia układu scalonego, projekty racjonalizatorskie,
- znaki towarowe,
- oznaczenia geograficzne.

W podstawowych aktach prawnych – ustawach i rozporządzeniach pojęcie usługi badawczo-rozwojowej nie jest zdefiniowane. Powszechnie, stosowane jest pojęcie „prace badawczo rozwojowe" z języka angielskiego – research and development. Często stosuje się pojecie prac naukowo badawczych. Usługi naukowo-badawcze występują w Polskiej Klasyfikacji Wyrobów i Usług (PKWiU).

Usługi naukowo-badawcze, zwane także badawczo-rozwojowymi, zaliczane są do grupowania statystycznego PKWiU 73. Obejmują one: badania, prace eksperymentalne i inne usługi badawczo-rozwojowe we wszystkich dziedzinach nauk.

Usługa badawczo-rozwojowa zdefiniowana jest w programie bon na innowacje [8]. *Przez usługę badawczo-rozwojową rozumie się usługę polegającą na opracowaniu nowych lub udoskonalenie już istniejących wyrobów i technologii. konieczne jest, by usługa badawczo-rozwojowa była zgodna z pozycją 73 PKWiU.*

Kredyt technologiczny nie może być udzielany na zakup, leasing lub wynajem środka trwałego, w którym została wdrożona nowa technologia będąca przedmiotem inwestycji technologicznej finansowanej za pomocą kredytu technologicznego.
„ Do wniosku o przyznanie premii technologicznej przedsiębiorca dołącza:
1) opinię sporządzoną na wniosek przedsiębiorcy przez jednostkę naukową lub centrum badawczo-rozwojowe, które nie są powiązane z przedsiębiorcą lub stowarzyszenie naukowo-techniczne o zasięgu ogólnopolskim, a ich zakres działania jest związany z inwestycją technologiczną, na którą ma być udzielony kredyt technologiczny, stwierdzającą, że technologia, która będzie wdrażana w wyniku realizacji inwestycji technologicznej finansowanej kredytem technologicznym, jest nową technologią;
2) informacje opracowane w porozumieniu z przedsiębiorcą przez podmiot, który sporządził opinię na podstawie pkt 1, zawierające:

a) charakterystykę technologii oraz przedstawienie właściwości świadczących o możliwości jej wdrożenia do wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług,
b) opis sposobu wdrożenia technologii do wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług mających być wynikiem realizacji inwestycji technologicznej finansowanej kredytem technologicznym,
c) wykaz i uzasadnienie zastosowania środków trwałych i wartości niematerialnych
i prawnych niezbędnych do wdrożenia technologii do wytwarzania nowych lub znacząco ulepszonych towarów, procesów lub usług mających być wynikiem realizacji inwestycji technologicznej finansowanej kredytem technologicznym,
d) opis towarów, procesów lub usług mających być wynikiem inwestycji technologicznej finansowanej kredytem technologicznym."

Jak widać pojęcie „nowej technologii" określone w ustawą z dnia 30 maja 2008 r. O niektórych formach wspierania działalności innowacyjnej, na znacznie węższe znaczenie niż to samo pojecie zdefiniowane w ustawie w dnia 25 lutego 1992 r., O podatku dochodowym od osób prawnych, z późniejszymi zmianami. Ustawa z dnia 30 maja 2008 r. O niektórych formach wspierania działalności innowacyjnej dokładniej:

definiuje co jest „nową technologią",
określa podmiot wydający opinie o „nowej technologii",
omawia wymaganą zawartość opinii,
oraz wymienia jakie informacji powinny być dołączone do opinii.

W związku z powyższym uzasadnienie, że planowane do zakupu oprogramowanie czy sprzęt stanowi „nową technologię" w sensie ustawy z dnia 30 maja 2008 r. O niektórych formach wspierania działalności innowacyjnej, jest trudniejsze niż to samo uzasadnienie spełniające wymogi ustawy z dnia 25 lutego 1992 r. O podatku dochodowym od osób prawnych, z późniejszymi zmianami.

4. Innowacyjna technologia

Bardzo modne od paru lat, za sprawą Unii Europejskiej, pojęcie innowacyjności nie posiada jednej, ogólnie stosowanej definicji. Wg Narodowych Strategicznych Ram Odniesienia na lata 2007-2013 - Narodowa Strategia Spójności [3]: *„Poprzez innowacyjność należy rozumieć zdolność i motywację przedsiębiorstw do ustawicznego poszukiwania i wykorzystania w praktyce wyników prac badawczych i rozwojowych, nowych koncepcji, pomysłów i wynalazków. Innowacyjność oznacza również doskonalenie i rozwój istniejących technologii produkcyjnych, eksploatacyjnych i dotyczących sfery usług,*

wprowadzania nowych rozwiązań w organizacji i zarządzaniu, doskonalenie i rozwój infrastruktury zwłaszcza dotyczącej gromadzeniu, przetwarzaniu i udostępnianiu informacji."

Wg Programu Operacyjnego Innowacyjna Gospodarka [4]:„*Przez innowacyjność rozumie się wdrożenie nowości do praktyki. Uszczegóławiając tą definicje, przez innowacyjność rozumie się wprowadzenie do praktyki nowego lub znacząco ulepszonego rozwiązania w odniesieniu do produktu (towaru lub usługi), procesu, marketingu lub organizacji*" (definicja na podstawie standardu Unii Europejskiej - podręcznika Oslo Manual [7]) „.... *Innowacyjność nie ma charakteru obiektywnego lecz relatywny w odniesieniu do konkretnego przedsiębiorstwa, które wdraża. ... W Programie Operacyjnym Innowacyjna Gospodarka największy nacisk zostanie położony na innowacje produktowe, a także procesową i marketingową związaną z elektroniczną gospodarka oparta na wiedzy w niej bowiem tkwi największy potencjał rozwoju.*"

W celu stymulowania innowacyjności refunduje się część kosztów związanych z wdrażaniem innowacyjności. Aby uzyskać dofinansowanie w ramach niektórych działań wymagane jest przedstawienie opinii o innowacyjności. Opinie takie mogą wydawać m. in. uczelnie wyższe i organizacje branżowe.

Opinia innowacyjności projektu nie ma standardowej postaci w skali kraju. Niektóre instytucje pośredniczące nie określają zawartości takiej opinii. Pisząc opinie dla której nie jest wymagany formularz logiczne jest że w opinii należy zgadać, czy projekt dotyczy innowacji produktowej, procesowej, marketingowej lub organizacyjne. Istnieją jednak organizacje np. PARP – Polska Agencja Rozwoju Przedsiębiorczości, które przewidziały dla opinii określony, ciągle doskonalony formularz [5]. Zawartość opinii o innowacyjności określa formularz takiej opinii. Zawiera on rubryki dotyczące wyników analizy innowacyjności produktu/usługi i/lub innowacyjności technologii. Należy zauważyć, że mimo tego, że literatura ani akty prawne nie podają definicji innowacyjności technologii formularz posługuje sie tym pojęciem. Najważniejsza zawarta formularzu informacja to określenie czy projekt dotyczy technologii: znanej i stosowanej nie dłużej niż 3 lata czy też nieznanej i niestosowanej dotychczas. Nie wystarczy podanie jedynie tej informacji. Powinna ona być uzasadniona. Uzasadnienie to - w szczególności musi wskazywać :

„ - *informację, które z elementów linii technologicznej produkcji lub procesu realizacji usługi są innowacyjne, a które pełnią rolę uzupełniającą;*

- *analizę informującą, czy wdrażana/nabywana technologia jest innowacyjna względem oferty już istniejącej na rynku i na czym ta innowacyjność polega;*

- *spis podstaw/źródeł danych, na podstawie których określono stopień innowacyjności technologii, z podaniem tytułów raportów, roczników statystycznych i dat ich wydania. adresów stron internetowych, roczników publikacji itp. oraz wskazaniem miejsca ich dostępności w celu zweryfikowania z informacjami przedstawionymi w opinii, z zastrzeżeniem, że podstawą opinii nie*

mogą być jedynie ogólne teksty reklamowo-opisowe dotyczące wdrażanej/nabywanej technologii."

Jak wynika z definicji ICT – technologii informacyjno komunikacyjnej, na technologie tą składa się sprzęt i oprogramowanie. Służą one z reguły do automatyzacji przesyłania informacji, procesu technologicznego czy wspomagania procesu np. świadczenia usług. Powołując się na odpowiednie źródła informacji można sporządzić opinie o innowacyjności planowanego do zakupu oprogramowania czy sprzętu. Należy jednak zauważyć, że w przypadku ostatnio wprowadzonych rozwiązań nie zawsze istnieją informacje inne niż pochodzące od producenta. Źródłem na które powołuje się opinia może wówczas być opublikowana przez producenta dokumentacja techniczna, która z definicji nie jest „tekstem reklamowo- opisowym".

5. Nowoczesna technologia

Państwo stymuluje również wdrażanie nowoczesnej technologii w zakładach pracy chronionej. Na podstawie przepisu § 2 ust. 1 pkt 2 Rozporządzenia Ministra Pracy i Polityki Społecznej z dnia 19 grudnia 2007 r. w sprawie zakładowego funduszu rehabilitacji osób niepełnosprawnych (Dz. U. Nr 245, poz. 1810), *„ze środków zakładowego funduszu rehabilitacji podlegają finansowaniu wydatki na finansowanie części kosztów wprowadzania nowoczesnych technologii i prototypowych wzorów oraz programów organizacyjnych proporcjonalnie do liczby zatrudnionych osób niepełnosprawnych w przeliczeniu na pełen wymiar czasu pracy."*

Urzędy Skarbowe wymagają czasami uzasadnienia, że wydatki ze środków zakładowego funduszu osób niepełnosprawnych dotyczyły nowoczesnej technologii i prototypowych wzorów. Podatnicy zwracają się o napisanie takiej opinii do instytucji uprawnionych do wydawania opinii o innowacyjności.

Należy zauważyć, że Minister an Rząd RP nie określa jak należy rozumieć poszczególne rodzaje wydatków z w/w rozporządzenia tzn. m in. nowoczesnych technologii i prototypowych wzorów. W takich sytuacja zalecane jest stosowanie wykładni językowej.

Zgodnie ze Słownikiem Języka Polskiego, PWN 1999:

„- wprowadzić oznacza zacząć coś stosować, oddać do użytku, na usługi czyjeś lub
czegoś, wprowadzić coś w czyn, w życie, zacząć coś realizować, stosować,
- nowoczesny - uwzględniający najnowsze osiągnięcia w jakiejś dziedzinie, postępowy,
- technologia to przetwarzanie w sposób celowy i ekonomiczny dóbr naturalnych w dobra ekonomiczne; wiedza o tym procesie,
- prototyp - pierwotny, najwcześniejszy wzór czegoś, według którego coś się tworzy,
który naśladuje, pierwszy wykonany według dokumentacji model maszyny lub

urządzenia, stanowiący po odpowiednich próbach i badaniach podstawę do dalszej
seryjnej produkcji."

Stosując do fragmentu rozporządzenia powyższą wykładnię językową włącznie z przytoczoną powyżej definicją technologii informacyjno komunikacyjnej ze słownika PARP możemy uzasadnić zaliczenie zakupu jako nowoczesnej technologii czy prototypowych wzorów. A zatem ze środków zakładowego funduszu rehabilitacji osób niepełnosprawnych sfinansowaniu podlegać mogą wydatki na zakup nowoczesnej technologii rozumianej jako: oddane do użytku przez dany podmiot, uwzględniające najnowsze osiągnięcia w jakiejś dziedzinie maszyny, urządzenia, oprogramowanie, które realizują (lub wspomagają) proces przetwarzania w sposób celowy i ekonomiczny dóbr naturalnych (w tym danych) w dobra ekonomiczne (np. informacje), oraz wydatki związane z nabycie wiedzy o tym procesie (np. szkolenie).

Ponieważ w przypadku ICT z reguły nabywany sprzęt i oprogramowanie uwzględniają najnowsze osiągnięcia. Producenci z reguły w opisach funkcjonalnych i technicznych produktu wraz datą wprowadzenia na rynek wymieniają wprowadzone nowości. Dlatego stosunkowo prosto jest uzasadnić, że sprzęt elektroniczny czy oprogramowanie stanowią „nowoczesną technologię" w sensie omawianego rozporządzenia. W razie wątpliwości można powałać się na orzecznictwo urzędów skarbowych. Przykładowo w postanowieniu [6] Naczelnik Małopolskiego Urzędu Skarbowego w Krakowie uznał, ż zakup komputerów i wprowadzenie nowoczesnej sieci komputerowej można sfinansować ze środków zakładowego funduszu osób niepełnosprawnych.

6. Podsumowanie

Pojęcia nowej, innowacyjnej, nowoczesnej technologii w języku potocznym są używane jako pojęcia pierwotne. Ustawodawca związał z nimi określone korzyści finansowe.

Dla nowej technologii ustawodawca podał w ustawach jej definicje. Definicje te różnią w zależności od ustawy. Dlatego sporządzając opinię dot. nowej technologii należy zwrócić uwagę jakiego aktu prawnego ona dotyczy.

Pojecie innowacyjnej technologii nie zostało zdefiniowane w dokumentach opisujących odpowiednie działania programów operacyjnych i regionalnych, które wymagają opinii o innowacyjności. Wskazówki co powinna zawierać opinia o innowacyjności technologii zawarte są formularzu opinii o innowacyjności. Formularz taki ulega zmianom się i dlatego sporządzając opinię należy zwrócić uwagę czy i na jakim formularzu powinna być wykonana.

Nowoczesna technologia nie została zdefiniowana przez ustawodawcę. Dlatego rozstrzygając o nowoczesności technologii można posługiwać się wykładnią

językową. Można również powoływać się na orzecznictwo urzędów skarbowych.

Przy rozstrzyganiu tego, czy oprogramowanie, sprzęt i związane z nimi usługi są nową, innowacyjna, nowoczesną technologia należy zwrócić uwagę na relatywizm tych pojęć. W zależności od potrzeby definicji pojęć i formularza mogą się różnić. Sporządzając opinie należy mieć na uwadze przez akty prawne, które uzyskanie korzyści finansowej uzależniają od przedstawienia określonej opinii oraz to czy nie jest wymagane przedstawienie opinii o określonym formularzu.

LITERATURA

1. Gurbiel. E. Hnadt-Olejniczak G., Kołczyk E., Krupicka H., Sysło M. M.: Technologia informacyjna Podręcznik dla liceum ogólnokształcącego, liceum profilowanego i technikum, WSiP, Warszawa 2008.
2. Matusiak K. B. red.: Innowacje i transfer technologii słownik pojęć. wydanie 2, Polska Agencja Rozwoju Przedsiębiorczości 2008 : http://www.pi.gov.pl/slowniki .
3. Narodowa Strategia Ram Odniesienia na lata 2007-2013 - Narodowa Strategia Spójności: http://www.funduszeeuropejskie.gov.pl/WstepDoFunduszyEuropejskich/Documents/NSRO_maj2007.pdf .
4. Program Operacyjny Innowacyjna Gospodarka: http://www.poig.gov.pl/Dokumenty/Lists/Dokumenty%20programowe/Attachments/89/POIG_01102008.pdf .
5. Przykładowy formularz opinii o innowacyjności: http://www.parp.gov.pl/index/index/608 .
6. Postanowienie małopolskiego Urzędu Skarbowego w Krakowie:
7. http://interpretacje.tmxp.pl/Podatek-dochodowy-od-osob-fizycznych/84505.html .
8. Podręcznik Oslo Manual: http://www.nauka.gov.pl/mn/index.jsp?place=Lead07&news_cat_id=1439&news_id=7333&layout=2&page=text .
9. Program Ban na Innowacje: http://www.parp.gov.pl/index/index/1089 .

Rozdział 32

Technologia informacyjno-komunikacyjna jako czynnik ewolucji organizacji gospodarczych w erze informacji i wiedzy

Damian Dziembek
Politechnika Częstochowska
dziembor@zim.pcz.czest.pl

Streszczenie

W rozdziale przedstawiono zagadnienia technologii informacyjno-komunikacyjnej jako czynnika stymulującego ewolucję organizacji gospodarczych.

1. Wstęp

Organizacje gospodarcze stanowiąc systemy otwarte funkcjonujące w turbulentnym otoczeniu muszą stale adoptować się do nowych warunków środowiska, a tym samym wprowadzać szereg różnorodnych zmian umożliwiających ich dalsze funkcjonowanie i rozwój. Obecnie dynamiczne i nieprzewidywalne zmiany zachodzące zarówno w otoczeniu konkurencyjnym, jak i globalnym wymuszają proces permanentnych przeobrażeń ogółu współczesnych organizacji gospodarczych. Zmiany otoczenia powinny być bieżąco analizowane, a ich wpływ powinien znaleźć odbicie w strategii biznesowej, istniejących procesach oraz podsystemach organizacji gospodarczej. Tym samym zmianie musi ulegać kształt, struktura oraz sposób funkcjonowania organizacji gospodarczych.

Do głównych sił i procesów obserwowanych w otoczeniu wpływających na działalność współczesnych organizacji gospodarczych można zaliczyć m.in.: globalizację, wzrost konkurencji, krótki cykl życia produktu, wzrastająca rola odbiorców, zmiany w wykształceniu i umiejętnościach pracowników, migracja ludności oraz coraz większe znaczenie norm etycznych i ekologicznych. Znaczne przeobrażenia współczesnych organizacji gospodarczych zachodzą również pod wpływem postępu w technologii informacyjno-komunikacyjnej.

2. Rola technologii informacyjno-komunikacyjnej w powstaniu Ery Informacji i Wiedzy

Technologie informacyjno-komunikacyjne (TIK) obecnie swym oddziaływaniem obejmują nie tylko sferę gospodarczą, ale również wpływają na pozostałe dziedziny działalności ludzkiej (np. obszar społeczny, polityczny, kulturowy, itp.). Przed przystąpieniem do określenia wpływu TIK na ewolucję organizacji gospodarczych celowe wydaje się być bliższe scharakteryzowanie Ery Informacji i Wiedzy, której powstanie było w znacznej mierze uzależnione do postępu w technologiach informacyjno-komunikacyjnych.

Współcześnie wielu badaczy jest zgodna, że głębokość i natężenie współczesnych przemian zachodzących w organizacjach gospodarczych (OG) ma skalę, która pozwala na formułowanie tezy o pojawieniu się nowych zasad i reguł funkcjonowania obowiązujących przy prowadzeniu działalności gospodarczej. Jednym z największych propagatorów głębokich zmian w funkcjonowaniu organizacji gospodarczych (jak również całego systemu społeczno-gospodarczego) jest A. Toffler twórca koncepcji tzw. trzeciej fali, która objawia się kwestionowaniem dotychczasowych struktur, zachowań i relacji obowiązujących w dotychczasowym świecie społeczno-gospodarczym por. [19, 20]. Według tego ujęcia - tradycyjny przemysł (np. włókienniczy, żelaza, stali, samochodowy, itp.) odgrywający pierwszoplanową rolę w okresie określanym jako Era Przemysłowa, traci na swym znaczeniu w obliczu nowych dziedzin przemysłu (związanych z głównie z technologiami informacyjno-komunikacyjnymi), charakterystycznych dla nowej epoki definiowanej jako Era Informacyjna. W tej nowej epoce (określanej przez filozofów, socjologów i ekonomistów także jako Gospodarka Informacyjna, Rewolucja Informacyjna, Era Świadomości Cyfrowej, Nowa Ekonomia, Gospodarka Elektroniczna, Gospodarka Oparta Na Wiedzy, Wiek Technotroniczny), gospodarka bazuje przede wszystkim na tych gałęziach przemysłu, które w aktywny sposób wykorzystują dane, informacje i wynikająca z nich wiedzę [1, 17, 12]. Dostęp do informacji i wiedzy, możliwość ich gromadzenia, przesyłania, powielania i udostępniania stają się we współczesnej gospodarce ważnym elementem wpływającym na uzyskiwanie przewagi konkurencyjnej.

Era Informacyjna (która ze względu na rolę i znaczenie wiedzy może być określana także jako Era Informacji i Wiedzy) zrodziła się przede wszystkim za sprawą obserwowanego na przestrzeni ostatnich lat dynamicznego postępu w obszarze technologii informacyjno-komunikacyjnych. Głównymi wyznacznikami Ery Informacji i Wiedzy są zjawiska zachodzące w otaczającej nas rzeczywistości, do których zaliczyć można [3]:

- informacja i wiedza staje się strategicznym zasobem (pełniącym rolę kolejnego czynnika wytwórczego), warunkującym strukturę i funkcjonowanie zarówno organizacji gospodarczych jak i całej gospodarki globalnej (światowej),

- rozwój TIK umożliwia tworzenie infrastruktury służącej dla gromadzenia, przetwarzania, przesyłania i udostępniania danych, informacji i wiedzy,
- postęp w dziedzinie TIK reorganizuje światowy system gospodarczo-finansowy umożliwiając monitorowanie międzynarodowych zależności ekonomicznych, społecznych i politycznych,
- rozwój TIK przyczynia się do zaniku tradycyjnych granic organizacyjnych (borderless organization) i powstania nowych form działalności (np. virtual organization, real time organization) o globalnym zasięgu działania,
- postęp w technologii informacyjno-komunikacyjnej przyczynia się do procesów integracyjnych gospodarek regionalnych i krajowych z gospodarką światową, sprzyjając globalizacji.

Doniosłe znaczenie Ery Informacji i Wiedzy przeistaczającej dotychczasowy obraz rzeczywistości gospodarczej, społecznej i politycznej podkreśla także D. Tapscott. Uważa on, iż wspólna kooperacja różnorodnych podmiotów (tj. ludzi, organizacji gospodarczych, instytucji, itp.) realizowana za pośrednictwem technologii informacyjno-komunikacyjnych umożliwia zespolenie zarówno informacji i wiedzy, jak również inteligencji i twórczości wpływając na rozwój społeczny i gospodarczy. D. Tapscott formułuje dwanaście podstawowych reguł obowiązujących w Erze Informacji i Wiedzy (określanej przez niego jako Gospodarka Cyfrowa), tj. [18]:

- Współczesna gospodarka to gospodarka informacji i wiedzy - podkreślająca rolę i znaczenie wiedzy, doświadczenia, umiejętności i innych wartości intelektualnych oraz inwestycji w zasoby ludzkie. W nowej gospodarce zachodzi proces stopniowego odchodzenie od pracy fizycznej na rzecz pracy umysłowej, a strategiczne znaczenie przypisywane jest pracom badawczo-rozwojowym (np. tworzenie i rozwój inteligentnych produktów – smart product);
- Nowy ład ekonomiczny ma charakter cyfrowy - w nowej ekonomii dominuje technologia cyfrowa, sprawiająca że większość gromadzonych danych przyjmuje postać elektroniczną. Cyfrowy zapis umożliwia szybki dostęp, przetwarzanie oraz archiwizację zasobów informacyjnych, jakie bez względu na czasokres, lokalizacje geograficzną i formę zostały zapisane przez człowieka;
- Rozwój rzeczywistości wirtualnej - transformacje analogowo-cyfrowe przekształcają byty realne w formy wirtualne (np. virtual reality, virtual shops, virtual product, virtual organization), co powoduje zmianę charakteru relacji partnerów gospodarczych oraz redefinicję zasad prowadzenia działalności gospodarczej;
- Molekularyzacja organizacji i gospodarki - nowa gospodarka to „gospodarka cząstek" cechującą się molekularną budową

poszczególnych jej elementów, w której w wyniku zmian otoczenia następuje proces dezintegracji zhierarchizowanych i zbiurokratyzowanych organizacji gospodarczych, zastępowanych przez mniejsze i dynamiczne podmioty, często wzajemnie ze sobą powiązanych;

- Rozwój organizacji sieciowych - gospodarka informacji i wiedzy jest gospodarką sieciową, bazującą na integracji i współdziałaniu organizacji gospodarczych ze swymi dostawcami, odbiorcami a nawet konkurentami, w której następuje szybki rozwój organizacji sieciowych, opartych na dynamicznych podmiotach gospodarczych wyposażonych w zasoby pozwalające niwelować granice fizyczne, czasowe, organizacyjne czy technologiczne, co pozwala budować systemy wzajemnych relacji;
- Eliminacja pośredników - następuje redefinicja zasad funkcjonowania organizacji gospodarczych, przejawiająca się w eliminacji funkcji i ogniw pośrednich (pośredników gospodarczych), zmuszonych do poszukiwania, zdefiniowania i zaoferowania swym kontrahentom nowych wartości rynkowych;
- Przenikanie się różnych obszarów aktywności gospodarczej – we współczesnej gospodarce następuje wzajemne przenikanie się sfer przetwarzania danych (computing), transmisji danych (communications), gromadzenia i udostępniania danych i wartości (content). Zespolenie wspomnianych sfer, stanowi ważny czynnik dla powstania i rozwoju gospodarki informacji i wiedzy;
- Innowacyjność i kreatywność stanowią podstawowy czynnik sukcesu – tworzenie nowych wartości w nowej gospodarce wymaga ludzkiej innowacyjności i kreatywności a istotną rolą współczesnych organizacji gospodarczych jest odpowiednio promować i nagradzać personel za jego twórcze rozwiązania;
- Postępujące zanikanie ścisłych podziałów między odbiorcami a dostawcami
  – w nowej gospodarce występuje silny współudział odbiorców w tworzeniu produktów przez dostawców, których pomysły, wiedza i umiejętności odgrywają istotną rolą na etapie projektowania i wytwarzania produktów;
- Funkcjonowanie w czasie rzeczywistym - szczególnego znaczenia w gospodarce cyfrowej nabiera funkcjonowanie w czasie rzeczywistym (real time economy), zapewniając organizacjom gospodarczym możliwość bieżącej realizacji usług, zleceń i procesów gospodarczych;
- Globalizacja gospodarki – nowa gospodarka jest gospodarką globalną, w której dane, informacja i wiedza przepływa bez ograniczeń ponad granicami krajów i kontynentów, umożliwiając i

zacieśniając współprace gospodarczą pomiędzy geograficznie rozproszonymi organizacjami;

- Wzrost niepokoju i zagrożenia - nowa gospodarka jest epoką społecznych niepokojów i zagrożeń następujących w dużej mierze na skutek rozdźwięku między ludźmi posiadającymi oraz nie dysponującymi dostępem do danych, informacji i wiedzy, czego efektem jest np. znaczna likwidacja niewykwalifikowanej siły roboczej, powstawanie nowych wysoko opłacalnych profesji, itp.

Jak podkreślają w swych pracach J. Huey i K. Kelly - w Erze Informacji i Wiedzy występują zarówno nowe siły napędowe dla rozwoju dotychczasowych (lub nowo tworzonych) organizacji gospodarczych, jak i nowe reguły i zasady ich odnoszące się do ich funkcjonowania i wewnętrznej struktury [6, 7, 8]. Dotychczasowe (tradycyjne) organizacje gospodarcze z chwilą nastania Ery Informacji i Wiedzy powinny sprawnie i skutecznie adoptować pojawiające się technologie informacyjno-komunikacyjne w celu przystosowania się do nowych warunków otoczenia.

Tab.1 Różnice pomiędzy Erą Przemysłową i Erą Informacji i Wiedzy w kontekście ewolucji współczesnych organizacji gospodarczych

Kryterium	Era Przemysłowa	Era Informacji i Wiedzy
Podstawowy zasób	Kapitał materialny	Kapitał niematerialny (informacja i wiedza, kreatywność, itp.)
Typ zachodzących zmian	Ewolucyjne	Rewolucyjne
Typ podejmowanych działań	Realne	Wirtualne
Sposób traktowania zmian zachodzących w otoczeniu	Zagrożenie	Szansa
Cel gospodarczy	Korzyść pojedynczej organizacji	Wspólny interes partnerów
Kultura organizacyjna	Oparta na hierarchii	Oparta na zaufaniu
Władza	Zależy od zajmowanego stanowiska w strukturze organizacyjnej	Zależy od posiadanych umiejętności, wiedzy i reputacji
Styl zarządzania	Nakazy i kontrola	Partycypacyjny
Struktura organizacyjna	Hierarchiczna (scentralizowana)	Sieciowa (wirtualna)
Ukierunkowanie	Orientacja na zadania	Orientacja na proces
Strategia	Nastawiona na konkurowanie	Nastawiona na kooperację
Cykl życia produktu	Długi	Krótki
Zakres oddziaływania	Krajowy	Globalny
Wartość rynkowa	Uzależniona od posiadanych aktywów finansowych i rzeczowych	Zależy od kapitału intelektualnego (aktywów niematerialnych)
Wykorzystywanie nowoczesnych technologii	Ważne	Konieczne
Postrzeganie zasobów ludzkich	Jako koszt	Jako inwestycja
Dominujący sektor	Przemysł ciężki	Usługi

Źródło: Opracowanie własne na podstawie: Boar: 1997, Strojny: 2000

L. Zacher wskazuje, iż Era Informacji i Wiedzy stwarza niezwykłe szanse dla organizacji gospodarczych i społeczeństwa. Poprzez radykalne zwiększenie możliwości przekazu informacji i wiedzy tworzy się całkowicie nowe warunki dla komunikowania się i współdziałania [25]. Wymaga to jednak głębokiej rekonstrukcji poglądów, sposobów myślenia oraz zasad funkcjonowania ogółu ludzi stanowiących zasoby własne organizacji gospodarczych oraz pozostających w jej bliższym lub dalszym otoczeniu. Zestawienie podstawowych różnic pomiędzy Erą Przemysłową a Erą Informacji i Wiedzy w kontekście zmian zachodzących w organizacjach gospodarczych, zestawiono w tabeli 1.

3. Obszary technologii informacyjno-komunikacyjnej i ich rola w przemianach współczesnych organizacji gospodarczych

Zasadniczą siłą sprawczą przekształcającą dotychczasowe reguły i zasady charakterystyczne dla Ery Przemysłowej do nowej epoki tj. Ery Informacji i Wiedzy - są rozwój i wzajemnie przenikanie się trzech podstawowych obszarów technologii informacyjno-komunikacyjnej tj.:

- sprzętu komputerowego (hardware) – np. wzrost mocy obliczeniowej mikroprocesorów, zwiększenie szybkości i pojemności pamięci operacyjnej i pamięci zewnętrznych, wzrost komunikacji pomiędzy elementami modularnymi komputera (mikroprocesorem, pamięcią oraz urządzeniami zewnętrznymi wejścia/wyjścia),
- oprogramowania (software) – dynamiczny rozwój aplikacji systemowych (systemów operacyjnych, języków programowania), aplikacji użytkowych (głównie systemów zintegrowanych typu MRP – Material Requirements Planning, MRP II - Manufacturing Resource Planning, ERP i ERP II – Enterprise Resource Planning, oraz systemów typu SCM – Supply Chain Management, CRM – Customer Relationship Management lub BI – Business Intelligence, systemów przepływu pracy - Workflow, portali korporacyjnych wiedzy – Enterprise Knowledge Portal) oraz innych metod i narzędzi programowych,
- baz danych (database) – np. postęp w zakresie relacyjnych i obiektowych baz danych oraz w obszarze lokalnych, scentralizowanych i rozproszonych systemów baz danych, rozwój architektury klient-serwer (GUI), hurtowni danych (data warehouse), powstawanie baz metod, baz modeli i baz wiedzy, rozwój w multimedialnych bazach danych, postęp w zabezpieczeniach baz danych, itp.,

- telekomunikacji (telecommunication) – rozwój technologii sieci teleinformatycznych (tj. w zakresie sieci lokalnych LAN – Local Area Network, sieci miejskich MAN – Metroplitan Area Network, sieci rozległych WAN – Wide Area Network – szczególną rolę odgrywa tu globalna sieć Internet oraz takie technologie jak np. GigabitEthernet, ATM, FDDI, EDI, itp.), rozwój technologii bezprzewodowych (np. WI-FI 802.11a/b/g), rozwój telefonii stacjonarnej (przejście do technologii cyfrowej ISDN) i komórkowej (np. GSM, UMTS), postęp w technologiach szerokopasmowych (xDSL),itp.

Do najważniejszych możliwości technologii informacyjno-komunikacyjnej z punktu widzenia funkcjonowania organizacji gospodarczych i wprowadzania w nich zmian należy zaliczyć [2, 21]:

- ułatwienie gromadzenia danych niezbędnych w funkcjonowaniu danego procesu,
- usprawnienie analizy zabranych danych i informacji (poprawa technik analitycznych i sposobów podejmowania decyzji),
- możliwość przekształcenia nieustrukturalizowanego procesu w rutynowo przebiegającą transakcję (automatyzacja procesów gospodarczych),
- ułatwienie wprowadzania zmian w kolejności przebiegu procesu oraz umożliwienie symultanicznego przebiegu niektórych jego elementów,
- umożliwienie łatwego dostępu do danych niezależnie od miejsca ich fizycznego przechowywania,
- umożliwienie łatwego dostępu do zasobów wiedzy oraz ich swobodnego transferu,
- umożliwienie eliminacji zbędnych pośredników z procesu,
- możliwość łatwego monitorowania przebiegu zarówno całego procesu, jak i jego poszczególnych elementów,
- możliwość zastępowania (lub zmniejszanie udziału) czynnika ludzkiego w realizowanych procesach.

Złożone i dynamicznie zmieniające się otoczenie powoduje konieczność restrukturyzacji wewnętrznej struktury i sposobów zarządzania organizacjami gospodarczymi. Jednym z powszechnie stosowanych narzędzi restrukturyzacji jest technologia informacyjno-komunikacyjna. Technologie informacyjno-komunikacyjne główną rolę odgrywają w usprawnieniu procesów komunikacyjno-informacyjnych zachodzących wewnątrz OG oraz w jego konkurencyjnym i globalnym otoczeniu. Jednakże skuteczność zastosowania narzędzi technologii informacyjno-komunikacyjnej jest uwarunkowana od ewolucyjnych przeobrażeń zachodzących wewnątrz i w otoczeniu OG oraz od przyjętej przez system zarządzania strategii działania i rozwoju danej organizacji

gospodarczej. Jest mało prawdopodobne by implementacja nowych TIK w ramach nie zmodernizowanych struktur organizacyjnych danej OG i braku właściwej strategii funkcjonowania spowodowała zwiększenie jej konkurencyjności, efektywne wykorzystywanie pojawiających się szans oraz minimalizacje zagrożeń płynących z otoczenia. Szczególnie ważne wydaje się być stwierdzenie J. Kisielnickiego podkreślającego, że wpływ TIK na nowe i efektywne formy organizacyjne współczesnych OG będzie widoczny wówczas, gdy będzie wspierany przez działania kreatywnego zespołu [9]. A zatem sama technologia informacyjna-komunikacyjna nie dostarcza przewagi konkurencyjnej, ale w istotny sposób może przyczynić się do jej osiągnięcia. Jak podkreśla M. Warner i M Witzel technologia informacyjno-komunikacyjna jest tylko artefaktem, nie mogąc ze swej istoty funkcjonować samodzielnie – co implikuje stwierdzenie że bez czynnika ludzkiego, staje się ona bezwładna i nie przedstawia żadnej wartości [24]. Współcześnie przewaga konkurencyjna może być osiągnięta w zasadzie wyłącznie poprzez działania ludzkie ukierunkowane na permanentną innowacyjność - obejmującą wprowadzanie nowych lub udoskonalenie dotychczasowych produktów, procesów, struktur i metod zarządzania w OG [4]. Można bez obaw stwierdzić, że pełne wykorzystanie szans oraz minimalizacja zagrożeń płynących z otoczenia, jak również tworzenie i wdrażanie innowacji byłoby bez TIK co najmniej trudne jeżeli w ogóle możliwe. Większość współczesnych innowacji bazuje bowiem na technologii informacyjno-komunikacyjnej, która zwykle stanowi fundament dla jej implementacji i dalszego wykorzystywania w praktyce gospodarczej. Jednak w obecnych czasach każda przewaga rynkowa uzyskana poprzez innowacje wykorzystujące technologię informacyjno-komunikacyjną jest nietrwała. Współcześnie innowacyjne rozwiązania są bowiem bardzo szybko powielane przez konkurentów lub stają się w ciągu krótkiego okresu czasu przestarzałe. Powoduje to konieczność stałego poszukiwania nowych prekursorskich rozwiązań wykorzystujących technologie informacyjno-komunikacyjne zapewniających wysoką efektywność oraz zwiększenie pozycji rynkowej organizacji gospodarczych. W wyniku przemian zachodzących w technologii informacyjno-komunikacyjnej powstaje nowy potencjał, nowe możliwości, organizacje nowego typu, nowe reguły przedsiębiorczości, jak również nowy porządek i ład gospodarczo-społeczny. W efekcie mogą zostać znacząco zmodyfikowane lub zreorganizowane dotychczasowe reguły działalności w zakresie struktury i funkcjonowania organizacji gospodarczych. Wspomniany D. Tapscott przedstawiając dziesięć podstawowych zmian technologii, wskazuje na skalę przemian i nową jakość i znaczenie TIK w Erze Informacji i Wiedzy [18]. Z jego wizją przeobrażeń zgadza się K. Kelly wskazując, iż w Erze Informacji Wiedzy - rola TIK ewoluuje z narzędzi stosowanych dla sterowania informacjami do rozwiązań zasadniczo ukierunkowanych na wymianę (i współdzielenie) informacji i wiedzy. Interaktywne systemy multimedialne bazujące na sieciach teleinformatycznych stwarzają jakościowo nowy potencjał, wspomagając proces zespolenia ludzkich myśli, wiedzy i inteligencji [8]. Stale

rozwijająca się TIK stopniowo zastępuje stare rozwiązania technologiczne tworząc nowe warunki dla prowadzenia działalności gospodarczej, a tym samym urzeczywistniając nową odmianę gospodarki bazującej na sieci tj. globalną gospodarkę cyfrową. Skalę zmian wywołanych TIK można porównać do przemian wywołanych wraz z pojawieniem się elektryczności czy silnika spalinowego. Wpływ TIK na przeobrażenia OG może dotyczyć praktycznie każdego obszaru jej działalności. Powszechność technologii informacyjno-komunikacyjnych, spadek cen ich zakupu i użytkowania oraz systematyczny wzrost szybkości przetwarzania przyczyniają się do wzrostu poziomu ich stosowania w różnorodnych obszarach społeczno-gospodarczych. Dodatkowo malejące koszty gromadzenia, przetwarzania, przesyłania i udostępniania danych (i wynikających z nich informacji), skutkują zwiększonym ich obiegiem w otoczeniu konkurencyjnym i globalnym. Obecnie dynamiczny rozwój i wzajemne przenikanie się takich obszarów technologii informacyjno-komunikacyjnej jak sprzęt komputerowy, oprogramowanie, bazy danych i telekomunikacja (ze szczególną rolą globalnej sieci Internet), kreują w Erze Informacji i Wiedzy nowe możliwości dla organizacji gospodarczych.

4. Zmiany w strukturze i funkcjonowaniu organizacji gospodarczych zachodzące pod wpływem technologii informacyjno-komunikacyjnej

Jak wspomniano wcześniej bazą dla wyłaniającej się Ery Informacji i Wiedzy jest informacja i wiedza oraz ich aktywne wykorzystanie we wszystkich procesach wytwórczych i usługowych. Główni prekursorzy tej filozofii gospodarowania (tj. D. Tapscott i K. Kelly) przyjmują, iż przyrost wartości dodanej jest efektem pracy umysłowej a nie wytwórczej. Wartość dodana przejawia się w postaci wiedzy zawartej w produktach i usługach, i ulega powiększeniu wraz ze wzrostem roli i udziału technologii informacyjno-komunikacyjnej oraz wyobraźni producentów, dostawców i odbiorców, które to rozpatrywane łącznie rewolucjonizują wszystkie dziedziny aktywności gospodarczej. Znaczny udział i stałe dodawanie do produktów wiedzy wpływa na ich innowacyjność stając się współcześnie kluczem dla osiągnięcia i utrzymania przewagi konkurencyjnej. Wiedza będzie przenikać każdy aspekt działalności OG a sieć stanie się nową infrastrukturą dla procesów zarządzania wiedzą zarówno wewnątrz organizacji gospodarczej, jak i pomiędzy OG a jej partnerami gospodarczymi i klientami. Globalna sieć umożliwi tworzenie wzajemnych powiązań ludzkiego intelektu, know-how, i pomysłowości, co w efekcie spowoduje nie tylko transformacje organizacji gospodarczych, ale także wpłynie na rozwój społeczno-gospodarczy odniesiony do całej gospodarki świata. Przemiany spowodowane technologią informacyjno-komunikacyjną, które zachodzą w organizacjach (jak i w całej cyfrowej gospodarce) będą odnosić się do powstania [18]:

- skutecznych jednostek (effective individuals),
- wysoko wydajnych zespołów (high performance teams),
- zintegrowanych organizacji (integrated organizations),
- rozszerzonych (wszechstronnych) organizacji (extended organizations).

Prezentację poszczególnych poziomów przemian określających kierunek dalszej ewolucji i doskonalenia organizacji gospodarczych zachodzącej w Erze Informacji i Wiedzy pod wpływem TIK zaprezentowano na rysunku 1. Poszczególnym obszarom – pożądanym efektom rozwoju gospodarczego przypisano rangi zilustrowane w ten sposób, że z każdego obszaru do następnego prowadzi jeden krok do góry, dając w efekcie model tworzenia wartości, wyznaczającego podstawę i sens prowadzenia działalności gospodarczej w sieci globalnej [13]. Transformacja OG zachodząca pod wpływem technologii informacyjno-komunikacyjnych przebiegała zarówno w sferze przekształceń wewnętrznych jak również obejmowała reorganizację relacji z otoczeniem. W zakresie przekształceń wewnętrznych TIK umożliwiła przed wszystkim automatyzację i optymalizację zadań i procedur zachodzących w OG oraz wdrażanie nowych koncepcji zarządzania (np. TQM czy BPR), stanowiąc czynnik zmieniający (a nawet burzący) strukturę organizacyjną tradycyjnie funkcjonujących organizacji gospodarczych. Dotychczasowe (tradycyjne) struktury organizacyjne nacechowane hierarchią, centralizacją i kontrolą oraz podlegające nadmiernemu usztywnieniu i uszczegółowieniu, zwykle utrudniają uwzględnianie szybko zmieniających się warunków otoczenia. Ponadto wspomniana struktura organizacyjna nie wspomaga szybkiego przepływu informacji zarówno pomiędzy podsystemem zarządzania i wytwarzania jak również pomiędzy OG a jej partnerami, dostawcami i odbiorcami.

Rys. 1. Nowe ICT a przemiany zachodzące w organizacjach gospodarczych

Źródło: [18]

Skala i zakres zmian organizacyjnych pod wpływem TIK może mieć charakter marginalny, aż do całkowicie zmieniających formę organizacyjną współczesnych OG. W literaturze wyróżnia się kilka poziomów przemian organizacji pod wpływem technologii informacyjno-komunikacyjnych [11]:

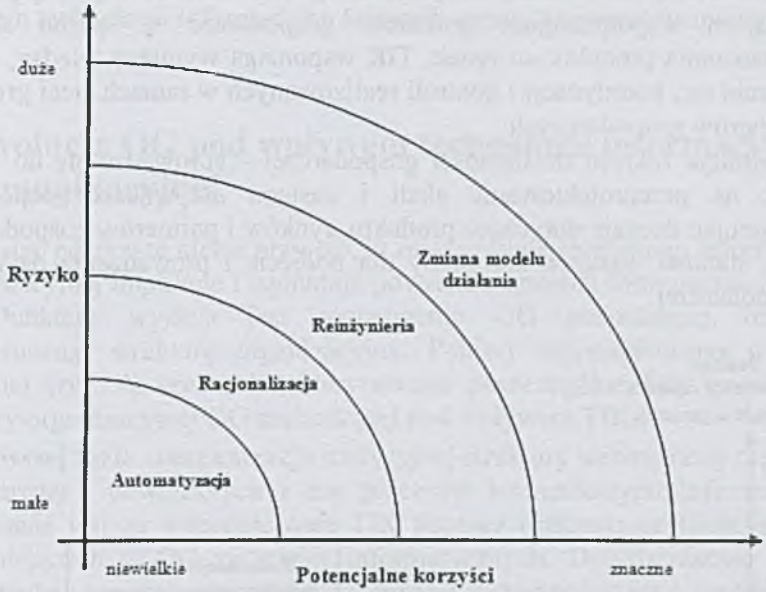
- automatyzacja i integracja wewnętrznych procedur – implementacja TIK służy bardziej wydajnej i efektywnej ekonomicznie realizacji zadań realizowanych przez personel OG; poprzez rozszerzanie zastosowań technologii informacyjno-komunikacyjnej na ogół obszarów funkcjonalnych, większość zadań i procedur realizowanych w OG podlega ujednoliceniu i optymalizacji,
- restrukturyzacja (racjonalizacja) – usunięcie wąskich gardeł i słabych punktów, uwidocznionych pod wpływem automatyzacji i integracji wewnętrznych procedur; tym samym następuje poprawa ich wydajności i efektywności,
- reinżynieria procesów biznesowych – polegająca na całkowitym przeprojektowaniu istniejących procesów biznesowych w oparciu o TIK, w celu zwiększenia efektywności funkcjonowania całej OG;
- zmiana modelu działania OG – zakładająca radykalne zmiany organizacyjne oraz reorganizację sposobu funkcjonowania OG; zmianie mogą podlegać dotychczas ustalone cele gospodarcze lub zostają wytyczone nowe innowacyjne plany, koncepcje i zamierzenia odnośnie dalszych metod organizacji i funkcjonowania OG,

Warto wspomnieć, że nie wszystkie inwestycje w TIK dostarczają korzyści. W wielu przypadkach TIK może wpływać na uzyskiwanie znacznych przychodów, lecz może również powodować generowanie dużych strat. Korzyści z implementacji technologii informacyjno-komunikacyjnej są szczególnie widoczne, jeżeli są scalone z wprowadzeniem znaczących zmian w organizacji i funkcjonowaniu OG. Implementacja technologii informacyjno-komunikacyjnej zmierzająca do uzyskania coraz wyższych przemian w organizacji i funkcjonowaniu OG, wiąże się z coraz wyższym ryzykiem, jednak wzrost ryzyka wpływa na możliwość generowania większych korzyści ekonomicznych. Prezentację wspomnianej zależności przedstawia rysunek 2.

Zakres zastosowania TIK może ewoluować stopniowo przechodząc z orientacji ukierunkowanej na automatyzację i integrację na doskonalenie i przeprojektowanie procesów gospodarczych aż do wspierania nowych strategii prowadzenia aktywności rynkowej (np. sieciowej współpracy gospodarczej). Poszczególne OG powinny gruntownie przeanalizować swoje wewnętrzne struktury, i zasady funkcjonowania a następnie zestawić je z istniejącym i przyszłym stanem otoczenia.

Przeprowadzona analiza winna umożliwić podjęcie decyzji, co do implementacji technologii informacyjno-komunikacyjnej w OG, która wpłynie na osiągnięcie takiego poziomu przemian organizacyjnych, przy którym relacja nakładów do

korzyści będzie optymalna. W miarę upływu czasu nowe warunki i czynniki pojawiające się w otoczeniu OG, mogą powodować jej dalsze przeobrażenia, a tym samym zwiększy się stopień nasycenia technologią informacyjno komunikacyjną.



Rys. 2. Poziomy przemian organizacyjnych w relacji do potencjalnych korzyści i ryzyka

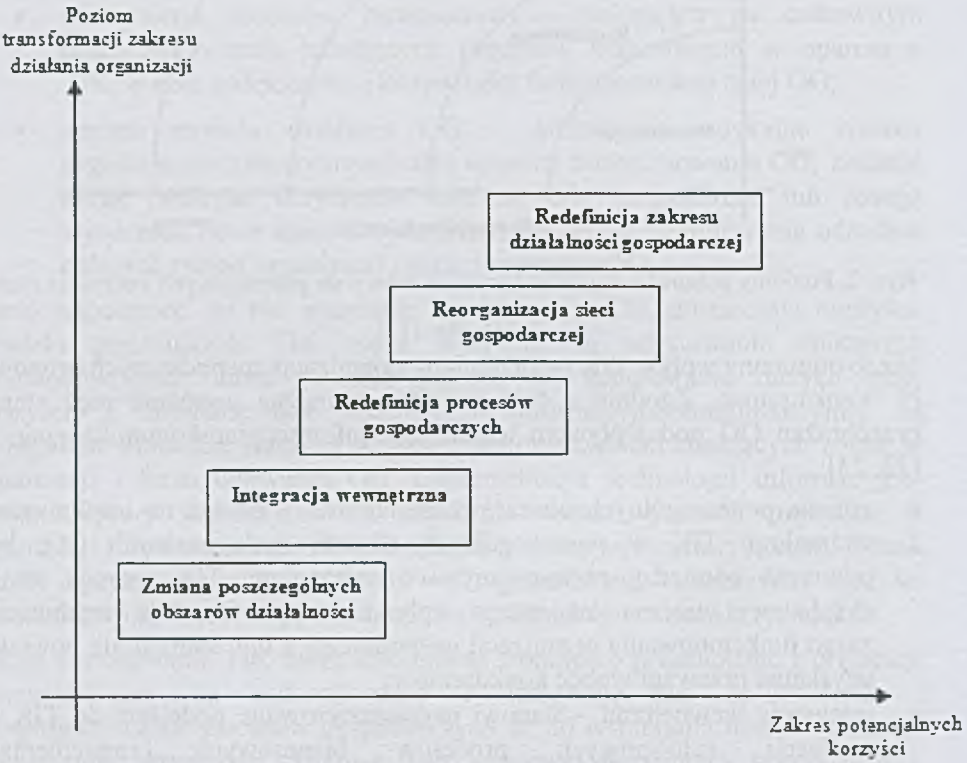
Źródło: [11]

Nieco odmienny wpływ TIK na przemiany organizacji gospodarczych proponuje N. Venkatraman. Zgodnie z tym podejściem można wyróżnić pięć etapów przeobrażeń OG pod wpływem technologii informacyjno-komunikacyjnej tj.: [22, 23]:

- zmiana poszczególnych obszarów działalności – polega na implementacji technologii TIK w poszczególnych sferach funkcjonalnych OG; brak powiązań pomiędzy zastosowanymi rozwiązaniami TIK sprawia, że ich eksploatacja nie ma znaczącego wpływu na transformację organizacji i zasad funkcjonowania organizacji gospodarczej a tym samym nie powoduje uzyskania przewagi wobec konkurentów;
- integracja wewnętrzna – stanowi usystematyzowane podejście do TIK dla realizacji całościowych procesów biznesowych; implementacja zintegrowanych rozwiązań TIK umożliwia wprowadzenie zmian w strukturze i funkcjonowaniu OG, co znajduje odbicie spójności organizacyjnej, poprawie wydajności, lepszej obsłudze odbiorcy, itp.,
- redefinicja procesów gospodarczych – obejmuje reorganizację kluczowych procesów wewnętrznych organizacji gospodarczej; połączenie TIK z przeprojektowaniem procesów gospodarczych znacząco transformuje

dotychczasowe struktury i sposób funkcjonowania OG, stanowiąc istotny czynnik umożliwiający osiągniecie przewagi rynkowej nad konkurencją;

- reorganizacja sieci gospodarczej – polega na ponownym zdefiniowaniu procesów zewnętrznych zachodzących pomiędzy OG a otoczeniem (dostawcami, odbiorcami, kooperantami), TIK stanowi wspólną platformę scalającą współpracujące podmioty gospodarcze w spójna sieć dla dostarczania produktu na rynek, TIK wspomaga wymianę wiedzy, procesy uczenia się, koordynacji i kontroli realizowanych w ramach sieci grupującej partnerów gospodarczych;
- redefinicja zakresu działalności gospodarczej – sprowadza się do wpływu TIK na przeprojektowanie skali i zasięgu aktywności gospodarczej, obejmując decyzje dotyczące produktu, rynków i partnerów gospodarczych, TIK stanowi warunek konieczny dla podjęcia i prowadzenia działalności gospodarczej.



Rys. 3. Poziomy przeobrażeń organizacji gospodarczych pod wpływem TIK

Źródło: Venkatraman: 2003, Kubiak, Korowicki: 1997

Poszczególne etapy ewolucji OG zachodzące pod wpływem technologii informacyjno-komunikacyjnej zależą do potencjału i gotowości do zmian OG

oraz branży i otoczenia w której funkcjonuje. Każdy poziom transformacji oferuje korzyści wynikającego z jego implementacji, przy czym im wyższy poziom tym skala możliwych pozytywnych efektów ulega powiększeniu. Wzrost korzyści jest także związany z zagrożeniami, jakie odnoszą się do poszczególnych etapów ewolucji. Prezentację poziomów transformacji pod wpływem technologii informacyjno komunikacyjnej zaprezentowano na rysunku 3.

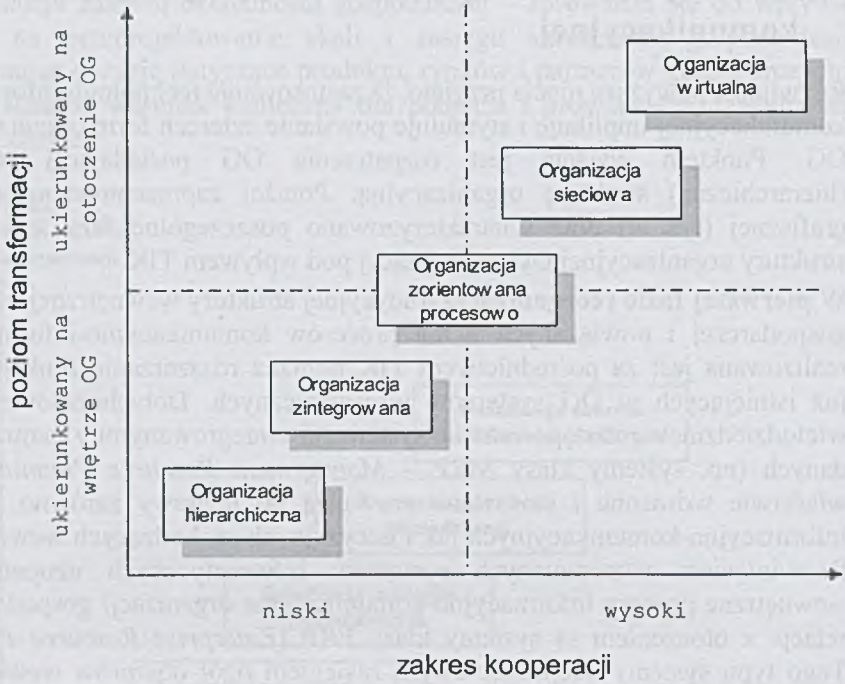
5. Ewolucja OG pod wpływem technologii informacyjno-komunikacyjnej

Rozwijając powyższe ujęcia przyjęto, iż zastosowanie technologii informacyjno-komunikacyjnej implikuje i stymuluje powstanie czterech form organizacyjnych OG. Punktem wyjścia jest rozpatrzenie OG posiadającej tradycyjną (hierarchiczną) strukturę organizacyjną. Poniżej zaprezentowano w formie graficznej (rys. 4) oraz scharakteryzowano poszczególne fazy transformacji struktury organizacyjnej OG zachodzącej pod wpływem TIK.

W **pierwszej fazie reorganizacja** tradycyjnej struktury wewnętrznej organizacji gospodarczej i powiązanych z nią procesów komunikacyjno-informacyjnych realizowana jest za pośrednictwem TIK poprzez rozszerzanie funkcjonalności już istniejących w OG systemów informatycznych. Dotychczasowe systemy wielodziedzinowe zastępowane są systemami zintegrowanymi z centralną baza danych (np. systemy klasy MRP – *Management Resource Planning*), które właściwie wdrożone i stosowane prowadzą do poprawy zarówno procesów informacyjno-komunikacyjnych jak i decyzyjnych zachodzących wewnątrz OG. Rozwinięciem wspomnianych systemów informatycznych uzupełniających wewnętrzne procesy informacyjno-komunikacyjne organizacji gospodarczych o relacje z otoczeniem są systemy klasy ERP (*Enterprise Resource Planning*). Tego typu systemy obejmując swym zasięgiem ogół obszarów wewnętrznych organizacji gospodarczych, utrwalają tradycyjne i scentralizowane struktury hierarchiczne funkcjonujące w OG według określonej specjalizacji dziedzinowej. Systemy klasy ERP pozwalają jednak na stosowanie różnorodnych form zarządzania ukierunkowanych i optymalizujących takie sfery działalności organizacji gospodarczych jak marketing, finanse czy logistyka, nie zapewniając jednakże pełnego wspomagania takich koncepcji zarządzania jak BPR czy VO.

Zastosowanie takich koncepcji jak BPR wymaga odejścia od orientacji funkcjonalnej na rzecz podejścia procesowego, łączącego ogół działań podejmowanych przez OG w spójny system realizowany w celu spełnienia wymogów otoczenia (głównie odbiorców). W orientacji procesowej następuje decentralizacja w obszarze wytwarzania i zarządzania a zasadniczą rolę w OG odgrywają samodzielne zespoły osób o dużych kompetencjach w realizacji powierzonych zadań i komunikujących się między sobą za pośrednictwem

sprawnego systemu informacyjnego. Powstaje nowy schemat struktury organizacyjnej OG wraz z odnoszącym się do niego wykazem łańcuchów realizowanych procesów oraz zestawem niezbędnych zasobów niezbędnych dla ich prawidłowego przeprowadzenia. W efekcie podejście procesowe powinno przynieść obniżkę kosztów projektowania, wytwarzania i dystrybucji produktów oferowanych przez daną organizacje gospodarczą przyczyniając się do pełniejszego zaspokojenia potrzeb zgłaszanych przez odbiorców.



Rys. 4. Fazy przemian organizacji gospodarczej pod wpływem TIK

Źródło: Opracowanie własne

Rozwój TIK umożliwił przejście OG do drugiego etapu restrukturyzacji umożliwiającego odchodzenie od orientacji funkcjonalnej w kierunku wspomnianego podejścia procesowego. Przyjęcie orientacji procesowej mające dostosować OG do zmieniającego się otoczenia zasadniczo przewartościowuje istniejące struktury organizacyjne. Jak wspomniano wcześniej, podejście procesowe powoduje że dotychczasowa wewnętrzna struktura organizacyjna (tj. struktury funkcjonalne i hierarchię, oddziały, itp.) zostaje zastąpiona przez grupę kompetentnych zespołów tworzonych przez personel z różnych działów OG, odpowiedzialnych za kompleksową, szybką i sprawną realizację zgłaszanych potrzeb konkurencyjnego otoczenia. Technologia informacyjno-komunikacyjna umożliwia reorganizację OG w kierunku orientacji procesowej poprzez

zastosowanie: najnowszych systemów informatycznych stanowiących rozwinięcie systemów klasy ERP, (np. ERP II, SCM, CRM, DEM), zintegrowanych baz i hurtowni danych, globalnych sieci teleinformatycznych.

Systemy klasy ERP II (uzupełnione innymi komponentami) powstały w celu stworzenia możliwości projektowania, optymalizacji oraz integracji procesów zachodzących w OG i jej otoczeniu oraz służą do implementacji i zastosowania orientacji procesowej w praktyce gospodarczej. Z założenia systemy tego typu ukierunkowane są na modelowanie struktur i ogółu procesów zachodzących we wnętrzu OG a poprzez oparcie swego funkcjonowania na globalnych sieciach teleinformatycznych, wspomagają istotnie większość relacji zachodzących pomiędzy OG a najbliższym otoczeniem reprezentowanym przez odbiorców, dostawców i partnerów gospodarczych.

Kolejnym, trzecim etapem redefinicji struktury wewnętrznej organizacji gospodarczej pod wpływem TIK jest możliwość głębszego rozwinięcia orientacji procesowej na kooperantów zewnętrznych. Projektowanie i integracja zewnętrzna procesów obejmująca kooperantów danej OG, następuje poprzez pełniejsze wykorzystanie najnowszych systemów klasy ERP II, SCM, CRM, DEM, rozproszonych lub scentralizowanych baz i hurtowni danych oraz globalnych sieci teleinformatycznych. W rezultacie możliwe jest przyśpieszenie przeobrażeń współczesnych organizacji gospodarczych związanych z postępująca dezagregacją dużych OG na mniejsze obiekty gospodarcze oraz wspieranie kooperacji w ramach różnorodnych powiązań zachodzących między organizacjami gospodarczymi (np. alianse strategiczne). Wspomniana forma współdziałania różnorodnych organizacji gospodarczych dla osiągania celów ekonomicznych powoduje wykreowanie tzw. sieciowych organizacji gospodarczych (SOG). Powstanie organizacji sieciowej ma zasadniczo miejsce wówczas, gdy kooperacja pomiędzy poszczególnymi OG ma charakter powtarzalny i realizowana jest w dłuższym okresie czasu. Całość procesów realizowanych w SOG rozdzielona jest na poszczególnych kooperantów w zależności od posiadanych umiejętności i doświadczenia. Zwykle w tego typu organizacji jeden z podmiotów kooperujących pełni nad pozostałymi funkcje koordynująco-sterujące (określany jako broker). Struktura organizacyjna SOG cechuje się względną stabilnością, istnieje jednak możliwość szybkiej rekonfiguracji zespołu kooperujących podmiotów. Sprawne przekonfigurowanie SOG uwypukla elastyczność tej formy organizacyjnej a pogłębiająca się specjalizacja poszczególnych kooperantów może powodować większą koncentrację podmiotów przy realizacji złożonego produktu na rzecz otoczenia.

Kooperacja partnerów gospodarczych z daną OG przy zastosowaniu wysokiej klasy systemów informatycznych (a w szczególności systemów interorganizacyjnych bazujących na globalnych sieciach teleinformatycznych) obejmująca pierwotnie dostarczanie poszczególnych towarów i usług, rozszerza się na wspólne doskonalenie ogółu procesów gospodarczych oraz współtworzenie nowych produktów. Wspomniana integracja zewnętrzna

ukierunkowana jest również na wspieranie kluczowych umiejętności partnerów gospodarczych współdziałających z dana organizacją gospodarczą w ramach outsourcingu oraz organizacji wirtualnej. W rezultacie OG tworzy system logistyczny złożony ze współdziałających i niezależnych partnerów gospodarczych obsługujących komplementarne sfery gospodarcze dla realizacji wspólnego celu. Postęp w technologii informacyjno-komunikacyjnej umożliwił powstanie przyjaznych narzędzi wspomagających współpracę różnych OG przy jednoczesnym zachowaniu niezależności poszczególnych kooperantów.

Podejmowanie kooperacji gospodarczej w formie outsourcingu lub organizacji wirtualnej ma na celu wyeliminować nieefektywne i niesprawne działania OG i zastąpić je działaniami efektywniejszymi, wykonywanymi przez zewnętrzne (obce) podmioty gospodarcze. Przekształcenie OG w organizację wirtualną uważa się za jedną z najnowszych form strukturalnych stosowanych we współczesnych organizacjach gospodarczych [10, 5].

Wirtualna organizacja gospodarcza (WOG) stanowi czwarty etap reorganizacji struktury wewnętrznej wskutek zastosowania technologii informacyjno-komunikacyjnej. WOG jest to forma kooperacji niezależnych i rozproszonych terytorialnie podmiotów gospodarczych, współdzielących kluczowe zdolności, zasoby koszty i ryzyko dla realizacji ustalonego celu, zintegrowanych w spójną jedność poprzez narzędzia i środki technologii informacyjno-komunikacyjnej. Współdziałające podmioty występują wobec odbiorców (i innych podmiotów) jako jednolita organizacja gospodarcza. Jak podkreśla P. Sieber i D. Swagerman wirtualna organizacja jako koncepcja strategiczna zwiększa konkurencyjność wchodzących w jej skład podmiotów gospodarczych, w dynamicznie zmieniającym się otoczeniu. [14]. Wirtualna OG stosując szerokie spektrum rozwiązań z obszaru TIK efektywnie przekracza geograficzne, kulturowe i organizacyjne granice, dostarcza produkty spełniające oczekiwania odbiorców i w umiejętny sposób dokonuje konwersji danych na informację, informacji na wiedzę a wiedzy na efektywne działania w zakresie adaptacji zarówno swej struktury organizacyjnej jak i prowadzenia działalności do dynamicznie zmieniającego się otoczenia.

Transformacja organizacji gospodarczej w organizację wirtualna stanowi efekt redefinicji dotychczasowej strategii funkcjonowania ukierunkowanej na aktywną adaptację do wymogów otoczenia. Wielopłaszczyznowe i wieloczynnikowe przekształcenia OG powodują utworzenie wirtualnej organizacji gospodarczej zdolnej do konkurowania na nowych rynkach, dostarczającego innowacyjnych produktów i nastawionego na zdobycie pozycji lidera w nowym cyfrowym środowisku utworzonym głównie dzięki połączeniu zasobów ludzkich, rzeczowych, finansowych i informacyjnych globalnymi sieciami teleinformatycznymi. Technologia informacyjno-komunikacyjna stworzyła możliwości dla sieciowego współdziałania podmiotów i uzyskiwania przewagi konkurencyjnej poprzez realizację zleceń klientów niezależnie od lokalizacji geograficznej i ograniczeń czasowych.

W odróżnieniu od tradycyjnych organizacji gospodarczych, WOG nie posiada fizycznej struktury organizacyjnej oraz zasobów stanowiących jej własność - powstaje więc zupełnie nowa jakość w obszarze organizacji i funkcjonowania współczesnych obiektów gospodarczych. Strukturę organizacyjną WOG tworzy zbiór dynamicznych i uzupełniających się podmiotów zintegrowanych i koordynowanych za pośrednictwem TIK dla realizacji określonego celu gospodarczego. Głównymi cechami struktury organizacyjnej WOG jest zmienność, tymczasowość i brak jednoznacznie identyfikowalnych granic. W literaturze podkreśla się brak fizycznej struktury organizacyjnej w WOG, określając ją poprzez takie sformułowania jak „dynamiczna sieć" lub „kolektyw organizacyjny" [16]. Podmioty tworzące WOG mogą równolegle wchodzić w skład innych organizacji gospodarczych a ich uczestnictwo w WOG ma dobrowolny charakter. Pod względem organizacyjnym WOG może być rozpatrywana jako połączenie wysokiej jakości kompetencji niezależnych podmiotów tworzących wspólnie jednolity obiekt gospodarczy. Kształt struktury organizacyjnej WOG (a tym samym ilość i typ tworzących ją podmiotów) wynika bezpośrednio z celu gospodarczego, dla którego została powołana (np. rodzaju zlecenia złożonego przez odbiorcę lub odbiorców).

Wspomniana wcześniej zmienność struktury wewnętrznej WOG może podlegać znaczącym ograniczeniom, co zdeterminowane jest głównie stopniem podobieństwa i powtarzalności realizowanych projektów oraz poziomem zadowolenia z dotychczasowej współpracy poszczególnych kooperantów. W związku z tym, struktura wewnętrzna WOG może charakteryzować się niską fluktuacją poszczególnych kooperantów, a nawet przyjmować względnie statyczny stan swej wewnętrznej struktury. Na tej podstawie można przyjąć, iż wewnętrzna struktura WOG może występować jako:

- statyczna,
- dynamiczna.

Statyczna struktura cechuje się wysokim poziomem stabilności (niezmienności) poszczególnych podmiotów wchodzących w skład wirtualnej organizacji gospodarczej. Oznacza to, iż kolejne projekty i zadania w ramach WOG realizowane są przez grupę tych samych współdziałających podmiotów. Zmienny skład kooperujących podmiotów jest charakterystyczny dla dynamicznej struktury WOG i wynika głównie z niskiej powtarzalności i podobieństwie realizowanych projektów, a także może być podyktowany brakiem zadowolenia i satysfakcji ze współpracy w ramach WOG.

Przedstawione różne formy organizacji i funkcjonowania OG mogą się urzeczywistnić w dużej mierze dzięki dynamicznemu rozwojowi technologii informacyjno-komunikacyjnej. Transformacja OG ukierunkowana jest na uwolnienie ich wewnętrznego potencjału ograniczonego sztywnymi strukturami funkcjonalnymi dla tworzenia wspólnych, wirtualnych i sieciowych organizacji dostarczających zbiorowo produkt spełniający oczekiwania odbiorcy. Nowe formy OG, dla których platformę współdziałania tworzy technologia

informacyjno-komunikacyjna, dysponują zarówno nieograniczonymi możliwościami tworzenia dowolnych zbiorowości, skomponowanych z producentów, pośredników i innych podmiotów kooperujących jak i kształtowania pomiędzy nimi właściwych relacji gospodarczych. Tak powstałe organizacje gospodarcze posiadają znaczny potencjał ekonomiczny dzięki skupieniu wielu podmiotów o różnych specjalizacjach i kompetencjach, tworząc nową jakość dla ostatecznego odbiorcy. W nowych formach OG, technologia informacyjno-komunikacyjna wpływa na stopniowe zacieranie się organizacyjnych i strukturalnych granic czyniąc ze zbiorowości rozproszonych terytorialnie podmiotów jeden pulsujący organizm gospodarczy.

6. Zakończenie

Globalny charakter Ery Informacji i Wiedzy, duży stopień nasycenia technologiami informacyjno-komunikacyjnymi oraz wzrost znaczenia informacji i wiedzy sprzyja powstawaniu nowych form strukturalnych i zmianie dotychczasowych zasad funkcjonowania organizacji gospodarczych. Skala przemian organizacji gospodarczych dokonująca się pod wpływem TIK uzależniona jest od specyfiki danej organizacji gospodarczej oraz stopnia zainteresowania kierownictwa zmianami dotychczasowej formy organizacyjnej oraz sposobem realizowanych procesów gospodarczych.

Wydaje się, że poziom nasycenia technologiami informacyjnymi w organizacjach gospodarczych wykazywał będzie tendencję rosnącą. Tym samym możliwa będzie praktyczna implementacja zaawansowanych form strukturalnych w organizacjach gospodarczych (np. organizacji sieciowych lub organizacji wirtualnych), których zarys został zaprezentowany w niniejszym rozdziale. Zapewne dalszy postęp w technologiach informacyjno-komunikacyjnych zapewniał będzie generowanie i rozwój nowych i bardziej innowacyjnych form OG, co powinno skutkować ich lepszym dostosowaniem do zmiennego otoczenia oraz umożliwiać uzyskanie zwiększonych korzyści gospodarczych.

LITERATURA

1. Boar B.H.: Strategic Thinking for Information Technology, John Wiley & Sons Inc., New York 1997
2. Davenport T.: Process Innovation: Reengineering Work through Information Technology, Harvard Business School Press, Boston 1993
3. Dziuba D.: "Przyjazne dla użytkownika" społeczeństwo informacyjne, w: Problemy społeczeństwa globalnej informacji, red. A .Szewczyk, USZ, Szczecin 2000
4. Freeman C, Asoete L.: The Economics of Industrial Innovation, The MIT Press, Cambridge 1999
5. Grudzewski W.M, Hejduk I.: Przedsiębiorstwo wirtualne, Difin, Warszawa 2002

6. Huey J.: Waking Up to the New Economy, Fortune, June 27, 1994
7. Kelly K.: New Rules for the New Economy - 10 Radical Strategies for a Connected World, New York 1998
8. Kelly K.: Nowe reguły nowej gospodarki. Dziesięć przełomowych startegii dla świata połączonego siecią, WIG-Press, Warszawa 2001
9. Kisielnicki J.: Technologia informacyjna jako szansa zaistnienia polskiej gospodarki na europejskim i globalnym rynku, w: Informatyka we współczesnym zarządzaniu (red) Kisielnicki J, Nowak J.S, Grabara J.K, WNT, Warszawa 2004
10. Kubiak B.F, Korowicki A.: Restrukturyzacja zarzadzania procesami gospodarczymi współczesnej organizacji z wykorzystaniem technologii informacji, w: Human Computer Interaction ' 97 (red.B. Kubiak, A.Korowicki), Gdańsk 1997
11. Laudon C.K, Laudon J.P.: Management Information Systems. Managing the Digital Firm, 8-th Edition, Pearson Prentice Hall, Upper Saddle River 2003
12. Nakamura L.J: Economics and the New Economy: The Invisible Hand Meets Creative Destruction, Federal Reserve Bank of Philadelphia Business Review, July-August 2000.
13. Pańkowska M, Sroka H (red).: Systemy informatyczne organizacji wirtualnych, Wydawnictwo AE Katowice, Katowice 2002
14. Sieber P, Swagerman D.: Operational Logic in Virtual Organizations - Example of Application Based on a Case Study, Electronic Journal of Organizational Virtualness, Virtual Organization Net, vol.1, No.1/2001
15. Strojny M.: Teoria i praktyka zarządzania wiedzą, EiOP, 10/2000
16. Sydow J.: Erfolg als Vertauensorganisation?, Office Management, 7-8, 1996
17. Tapscott D (ed.): Creating Value in the Network Economy, Harvard Business Review Book, 2000
18. Tapscott D.: Gospodarka cyfrowa, Wydawnictwo Business Press, Warszawa 1998
19. Toffler A.: Powershift: Knowledge, Wealth and Violence at the Edge of the 21 Century, Bantam Books, New York 1990
20. Toffler A.: Trzecia fala, PWN, Warszawa 1997
21. Unold J.: Systemy informacyjne marketingu, Wydawnictwo AE we Wrocławiu, Wrocław 2001
22. Venkatraman N.: IT-Enabled Business Transformation: From Automation to Business Scope Redefinition, Sloan Managament Review, 1994
23. Venkatraman N.: Other Voices: The Real Impact Of IT Is Just Beginning, Information Week, 23.06.2003
24. Warner M, Witzel M.: Zarządzanie organizacją wirtualną, Oficyna Ekonomiczna, Kraków 2005
25. Zacher L.W. (red): Problemy społeczeństwa informacyjnego, Wydawnictwo WSPiZ, Warszawa 1997

Rozdział 33

Ewolucja programów nauczania informatyki w polskim systemie edukacyjnym

Sławomir Iskierka, Janusz Krzemiński, Zbigniew Weżgowiec
Politechnika Częstochowska
wezgow@el.pcz.czest.pl

Streszczenie

W rozdziale podjęto próbę analizy zmian, jakim podlegały programy nauczania informatyki w polskim systemie edukacyjnym ze szczególnym uwzględnieniem zagadnień związanych z bezpiecznym wykorzystywaniem technologii informatycznych. Zwrócono uwagę na trudności wypracowania określonych minimów programowych dla poszczególnych etapów kształcenia w związku z dynamicznymi zmianami zachodzącymi w sektorze informatyki i teleinformatyki. Przeanalizowano podstawowe koncepcje metodologiczne i dydaktyczne związane z doborem treści programowych nauczania informatyki i technologii informacyjnej. Szczególną uwagę poświęcono sieciom komputerowym a zwłaszcza bezpiecznemu korzystaniu z sieci Internet. Zwrócono uwagę na konieczność kształcenia w ujęciu problemowym, tak, aby przekazywane treści programowe związane z nauczaniem informatyki miały charakter uniwersalny i w maksymalnym stopniu nie były związane z konkretną platformą sprzętową czy też programistyczną.

1. Wstęp

Pojęcia społeczeństwa informacyjnego i gospodarki opartej na wiedzy zaczęły funkcjonować w powszechnej świadomości w drugiej połowie dwudziestym wieku. Stało się to w wyniku dynamicznego rozwoju techniki opartej głównie na nowych technologiach wykorzystujących zdobycze informatyki i telekomunikacji. Technologie te, początkowo wykorzystywane praktycznie tylko w wybranych gałęziach gospodarki, szybko przeniknęły do wszystkich

dziedzin życia, scalając się w między czasie w jeden duży blok nazywany obecnie technologiami teleinformatycznymi.

Rozwój telekomunikacji i informatyki, jaki dokonał się w dwudziestym wieku jego skala, zakres i tempo są bezprecedensowe w dziejach. Powstała w związku z tym ogromna ilość wiedzy, której przekazanie uczniom i studentom w ramach przedmiotu informatyka stanowi istotne wyzwanie dla dydaktyków tego przedmiotu.

Jednym z głównych problemów przy tworzeniu programów nauczania informatyki jest w tym przypadku dobór treści nauczania, ich szczegółowość na poszczególnych etapach kształcenia oraz dobór odpowiednich metod dydaktycznych. Jednocześnie, wobec dynamicznych zmian, jakim podlega współczesna informatyka istnieje konieczność stałego modernizowania programów tak, aby przekazywane treści najwierniej odzwierciedlały jej aktualny stan. Nie wnikając w rozważania związane z samym formalnym rozumieniem pojęcia – informatyka, które przeprowadzono między innymi w pracach [1, 2, 8, 10, 14, 24, 28, 32, 34], jako najważniejsze elementy przedmiotu informatyka, które uwzględniano w treściach programowych są:

- Budowa i działanie komputera
- Systemy operacyjne
- Programy użytkowe i narzędziowe
- Algorytmika i programowanie
- Sieci komputerowe i Internet
- Bezpieczeństwo systemów komputerowych

Analizując powyższe zagadnienia można zauważyć, że nie wszystkie z nich podlegają jednakowej dynamice zmian. Tym samych ewolucja treści programów dydaktycznych związana była przede wszystkim z rozwojem sprzętu, powstawaniem nowych systemów operacyjnych, rozwojem technik programowania, upowszechnieniem się sieci komputerowych i dostępu do Internetu. W mniejszym stopniu zmiany notowane były w algorytmice czy zasadach działania systemów komputerowych. Jednocześnie należy podkreślić znamienny fakt, że ewolucja programów nauczania informatyki dokonywała się niejako w dwóch aspektach. Pierwszy z nich związany był ze zmianami sprzętu i oprogramowania, sieciami komputerowymi i Internetem i wynikał z konieczności przedstawienia najnowszych osiągnięć w informatyce i teleinformatyce. Konieczność ich wprowadzenia nie budziła zastrzeżeń. Ewentualne spory dotyczyły stopnia szczegółowości prezentowania poszczególnych zagadnień i rozłożeniu akcentów pomiędzy teorią i praktyką. Drugi dotyczący programowania i algorytmiki miał charakter bardziej fundamentalny i sprowadzał się do sposobu postrzegania informatyki jako dziedziny nauki niezwiązanej bezpośrednio ze sprzętem i sposobami jego wykorzystania. Efektem takiego podejścia było między innymi wygenerowanie, w systemie oświaty, nowego przedmiotu, jakim stała się technologia

informacyjna ukierunkowana przede wszystkim na praktyczną stronę wykorzystania informatyki w życiu codziennym.

Przedstawione powyżej problemy uwidaczniają trudności, z jakimi muszą się zmierzyć twórcy nowoczesnych programów dydaktycznych z informatyki i nauczyciele tego przedmiotu.

2. Edukacja informatyczna w polskich szkołach

Interesującego podziału, na pięć etapów w dydaktyce edukacji informatycznej w polskich szkołach dokonał A. Piecuch [27]. Autor za kryterium podziału przyjął obowiązujące na danym etapie programy nauczania. W niniejszym opracowaniu wykorzystano ten podział, przyjmując stosowane przez jego autora nazewnictwo wraz z proponowanymi przez niego przełomowymi programami nauczania, dodając jednocześnie do każdego z nich podstawowe informacje dotyczące używanych wówczas procesorów, systemów operacyjnych i technologii sieciowych:

1. Obejmuje lata siedemdziesiąte.

Był to okres, w którym dominowały duże maszyny cyfrowe (o znikomej jak na obecne warunki mocy obliczeniowej). Znajdowały się one najczęściej w ośrodkach obliczeniowych (ZETO) obsługujących przedsiębiorstwa państwowe oraz na uczelniach wyższych. Komputery klasy IBM 360, ODRA i RIAD obsługiwali zawodowi informatycy kształceni w nielicznych ośrodkach w kraju. Używane systemy operacyjne to najczęściej różne wersje UNIX-a. Studenci, którzy mieli w programie studiów informatykę, najczęściej realizowali ją w postaci zajęć teoretycznych, a pisane przez siebie programy w takich językach jak: Algol 60, Cobol czy Fortran uruchamiali na tych maszynach najczęściej w systemie wsadowym, w którym pośrednikiem był zawodowy operator danego komputera. Dostęp do tych maszyn dla uczniów szkól średnich, poza nielicznymi wyjątkami dotyczącymi dużych ośrodków akademickich, był praktycznie niemożliwy. W drugiej połowie lat siedemdziesiątych pojawia się procesor 8086 firmy Intel. Zostaje zapoczątkowana era komputerów osobistych. Na rynek zostają wprowadzone komputery klasy IBM PC XT, a następnie PC AT z procesorem 80286.

2. Rozpoczynający się od 1985, który dotyczył już bezpośrednio oświaty i związany był z powstaniem pierwszego programu nauczania *Elementów informatyki* dla szkół średnich.

W tym czasie komputery PC AT zostają wyposażone w procesory 80386, o nowych niespotykanych dotychczas możliwościach. Koniec lat dziewięćdziesiątych ubiegłego wieku to rozpoczynająca się era procesorów 80486 umożliwiających przetwarzanie potokowe.

3. Rozpoczął się w 1990 roku wprowadzeniem do użytku szkolnego przedmiotu *Elementy informatyki* dla klas VIII szkoły podstawowej wraz z odpowiednim programem nauczania i trwał do połowy dziesięciolecia.

Jest to burzliwy okres rozwoju sprzętu i oprogramowania. W 1992 roku firma Microsoft wprowadza na rynek system operacyjny Windows 3.11. Pojawia się również nowość system Windows for Workgroups przeznaczony do pracy w sieciach lokalnych. Na rynku coraz popularniejsze stają się modele komputerów IBM PC AT z procesorem 80486. W 1993 roku firma Intel wprowadza na rynek procesor Pentium o nowych możliwościach obliczeniowych.

4. Początek to przełom lat 1994/1995. W tym okresie zatwierdzono do użytku szkolnego trzy programy nauczania *Elementów informatyki*.

Jest to również okres wprowadzenia systemu Windows 95, a zwłaszcza Windows NT usprawniających pracę w sieciach lokalnych. Popularność z zyskuje nowa usługa dostępna w Internecie – WWW. Przy końcu tego etapu pojawia się system Windows 98. W 1997 roku firma Intel wprowadza na rynek procesor Pentium II. W szkołach coraz częściej pojawiają się sieci lokalne w technologii Ethernet oparte o cienki kabel koncentryczny

5. To rok 1999, w którym wchodzi w życie reforma systemu oświaty a wraz z nią przedmioty informatyczne *Technologia informacyjna* i *Informatyka*.

Początek tego etapu to wprowadzenie systemu Windows 98 Second Edition. Coraz popularniejszy Internet, można było, dzięki temu systemowi (usługa – Udostępnianie połączenia internetowego) współdzielić z jednego połączenia zewnętrznego na kilka komputerów. Połączenie z Internetem poprzez modem analogowy. W szkołach sieci lokalne budowane są w technologii Ethernet gdzie jako medium transmisyjne wykorzystuje się skrętkę. W 1999 roku firma Intel wprowadza na rynek procesor Pentium III, dzięki niemu łatwiejsza staje się praca z grafiką 3D i aplikacjami muzycznymi. Microsoft udostępnia system Windows 2000 i w krótkim czasie Windows XP. Na rynku pojawia się procesor Pentium 4, a Microsoft udostępnia system operacyjny Windows 2003. W roku 2005 rozpoczyna się era procesorów wielordzeniowych wprowadzeniem na rynek przez firmę Intel dwurdzeniowego procesora Intel Pentium D. W następnych latach pojawiają się ulepszone wersje procesorów dwu i wielordzeniowych. Pojawia się pilne zapotrzebowanie na oprogramowanie umożliwiające wykorzystanie ich mocy obliczeniowej. Stopniowo poprzez wymianę urządzeń aktywnych w szkolnych sieciach komputerowych prędkość ich pracy zwiększa się 10 Mbps na 100 Mbps Lata 2007 i 2008 przynoszą dwie kolejne wersje systemu Windows, odpowiednio system Windows Vista i Windows Server 2008. Systemy te przynoszą istotne zmiany w stosunku do poprzednich wersji, szczególnie w dziedzinie bezpieczeństwa systemów i pracy w sieci. Większość szkół posiada już szerokopasmowy dostęp do Internetu. Czy ten okres zaowocuje nowym etapem edukacji informatycznej?

6. Przełom lat 2007/2008?

Funkcjonujące w poszczególnych etapach programy nauczania informatyki czy technologii informacyjnej powstawały w oparciu a aktualną w danym okresie podstawę programową kształcenia ogólnego. Podstawy te już w okresie wdrażania reformy oświatowej były wielokrotnie aktualizowane. W celu prześledzenia zmian, jakie wprowadzano w podstawach programowych, a które dotyczyły przedmiotów informatyka i technologia informacyjna dokonano porównania podstawy programowej wprowadzonej w 2003 roku (Rozporządzenie MEN z 6 listopada 2003 r.) i aktualnej podstawy programowej z 2009 roku (Rozporządzenie MEN z 23 grudnia 2008 r.).

Do podstawowych zagadnień ujętych w tych programach na poszczególnych etapach, zgodnie z podstawą programową z 2003 roku [30], należą:

a) II etap edukacyjny klasy IV – VI

Przedmiot - Informatyka

- Zasady bezpiecznego posługiwania się komputerem
- Komputer jako źródło wiedzy i komunikowania się. Zastosowanie komputera w życiu codziennym.
- Opracowanie za pomocą komputera prostych tekstów, rysunków i motywów.
- Korzystanie z elementarnych zastosowań komputerów do wzbogacania własnego uczenia się i poznawania różnych dziedzin wiedzy.
- Poznawanie zastosowań komputerów i opartych na technice komputerowej urządzeń spotykanych przez ucznia w miejscach publicznych

b) III etap edukacyjny – gimnazjum

Przedmiot - Informatyka

- Posługiwanie się sprzętem i korzystanie z usług systemu operacyjnego
- Rozwiązywanie problemów za pomocą programów użytkowych
- Rozwiązywanie problemów w postaci algorytmicznej
- Modelowanie i symulacja za pomocą komputera

c) Licea ogólnokształcące, licea profilowane, technika

Przedmiot - Technologia informacyjna

- Opracowywanie dokumentów o rozbudowanej strukturze zawierających informacje pochodzące z różnych źródeł
- Rozwiązywanie zadań z zakresu różnych dziedzin nauczania z wykorzystaniem programów komputerowych i metod informatyki
- Podstawowe formy organizowania informacji w bazach danych spotykanych w otoczeniu ucznia. Wyszukiwanie informacji w bazach danych, formułowanie rozbudowanych zapytań
- Korzystanie z informacji związanych z kształceniem, pochodzących z różnych źródeł oraz komunikowanie się poprzez sieć

- Wspomaganie prezentacji prac uczniów z zastosowaniem programów komputerowych. Prezentacja w sieci.
- Rozwój zastosowań komputerów. Prawne i społeczne aspekty zastosowań komputerów

Przedmiot - Informatyka - kształcenie w zakresie rozszerzonym (liceum ogólnokształcące)

- Algorytmika i programowanie
- Bazy danych
- Multimedia. Sieci komputerowe

d) Szkoły ponadpodstawowe

Przedmiot - Elementy informatyki

- Posługiwanie się sprzętem komputerowym i korzystanie z usług systemu operacyjnego
- Stosowanie programów użytkowych do wykonywania zadań szkolnych
- Algorytmy rozwiązywania zadań
- Społeczne, etyczne i ekonomiczne aspekty rozwoju informatyki

Przedmiot – Technologia informacyjna

- Podstawowe formy organizowania informacji w bazach danych spotykanych w otoczeniu ucznia. Wyszukiwanie informacji powyżej bazach danych
- Korzystanie z informacji związanych z kształceniem, pochodzących z różnych źródeł oraz komunikowanie się poprzez sieć
- Wspomaganie prezentacji prac uczniów z zastosowaniem programów komputerowych.
- Rozwój zastosowań komputerów. Prawne i społeczne aspekty zastosowań informatyki

Natomiast podstawa programowa z 2009 roku [31] do podstawowych zagadnień, które przedstawiono tak jak je ujmuje podstawa programowa, dla wybranych etapów kształcenia, zalicza:

e) III etap edukacyjny – gimnazjum

Przedmiot - Informatyka

- Bezpieczne posługiwanie się komputerem i jego oprogramowaniem, korzystanie z sieci komputerowej.
- Wyszukiwanie i wykorzystywanie (gromadzenie, selekcjonowanie, przetwarzanie) informacji z różnych źródeł; współtworzenie zasobów w sieci.
- Komunikowanie się za pomocą komputera i technologii informacyjno-komunikacyjnych.

- Opracowywanie za pomocą komputera rysunków, tekstów, danych liczbowych, motywów, animacji, prezentacji multimedialnych.
- Rozwiązywanie problemów i podejmowanie decyzji z wykorzystaniem komputera, stosowanie podejścia algorytmicznego.
- Wykorzystywanie komputera oraz programów i gier edukacyjnych do poszerzania wiedzy
- i umiejętności z różnych dziedzin.
- Wykorzystywanie komputera i technologii informacyjno-komunikacyjnych do rozwijania zainteresowań; opisywanie innych zastosowań informatyki; ocena zagrożeń i ograniczeń, aspekty społeczne rozwoju i zastosowań informatyki.

f) IV etap edukacyjny – liceum

Przedmiot – Informatyka (zakres podstawowy)

- Bezpieczne posługiwanie się komputerem, jego oprogramowaniem i korzystanie z sieci komputerowej.
- Wyszukiwanie, gromadzenie, selekcjonowanie, przetwarzanie i wykorzystywanie informacji, współtworzenie zasobów w sieci, korzystanie z różnych źródeł i sposobów zdobywania informacji.
- Uczeń wykorzystuje technologie komunikacyjno-informacyjne do komunikacji i współpracy z nauczycielami i innymi uczniami, a także z innymi osobami, jak również w swoich działaniach kreatywnych.
- Opracowywanie informacji za pomocą komputera, w tym: rysunków, tekstów, danych liczbowych, animacji, prezentacji multimedialnych i filmów.
- Rozwiązywanie problemów i podejmowanie decyzji z wykorzystaniem komputera, stosowanie podejścia algorytmicznego.
- Wykorzystywanie komputera oraz programów edukacyjnych do poszerzania wiedzy i umiejętności z różnych dziedzin.
- Wykorzystywanie komputera i technologii informacyjno-komunikacyjnych do rozwijania zainteresowań, opisywanie zastosowań informatyki, ocena zagrożeń i ograniczeń, aspekty społeczne rozwoju i zastosowań informatyki.

Przedmiot – Informatyka (zakres rozszerzony)

- Posługiwanie się komputerem i jego oprogramowaniem, korzystanie z sieci komputerowej.
- Wyszukiwanie, gromadzenie, selekcjonowanie, przetwarzanie i wykorzystywanie informacji, współtworzenie zasobów w sieci, korzystanie z różnych źródeł i sposobów zdobywania informacji.

- Komunikowanie się za pomocą komputera i technologii informacyjno-komunikacyjnych.
- Opracowywanie informacji za pomocą komputera, w tym: rysunków, tekstów, danych liczbowych, animacji, prezentacji multimedialnych i filmów.
- Rozwiązywanie problemów i podejmowanie decyzji z wykorzystaniem komputera, stosowanie podejścia algorytmicznego.
- Wykorzystywanie komputera oraz programów edukacyjnych do poszerzania wiedzy i umiejętności z różnych dziedzin
- Wykorzystywanie komputera i technologii informacyjno-komunikacyjnej do rozwijania zainteresowań, opisywanie zastosowań informatyki, ocena zagrożeń i ograniczeń, docenianie aspektów społecznych rozwoju i zastosowań informatyki

Przedstawione powyżej zagadnienia, dla podstawy programowej obowiązującej od roku szkolnego 2009/2010 podane wraz z wymaganiami szczegółowymi, powinny być oczywiście realizowane w programach szkolnych. Obsługa komputera rozumiana jako umiejętność posługiwania się edytorem tekstu, arkuszem kalkulacyjnym, programem graficznym, elementarną bazą danych czy prostym programem komunikacyjnym jest dzisiaj tak niezbędna jak nauka czytania i pisania i są to stwierdzenia, tak obiegowe jak i akceptowalne przez wszystkich. Umieszczanie jednak tych umiejętności w ramach przedmiotu informatyka budzi wątpliwości. Wydaje się, że poprawniejszą nazwą, obejmującą między innymi te zagadnienia była – technologia informacyjna.

Porównując ogólne cele kształcenia ujęte w obu podstawach programowych można zauważyć istotne podobieństwa. Jest to zrozumiałe, gdyż dotyczą one tej samej materii. Stąd wniosek, że wobec dynamicznych zmian zachodzących we współczesnej technologii teleinformatycznej wprowadzenie ich do programów nauczania musi być dokonywane na poziomie wymagań (treści nauczania) szczegółowych. Niemniej te szczegółowe wymagania mogą być różnie interpretowane i rozpatrywane na różnym poziomie szczegółowości. Ponadto umieszczanie, przynajmniej części wymagań szczegółowych może budzić pewne wątpliwości, co do celowości ich wprowadzania. Tym bardziej, że na przykład dla IV etapu kształcenia wprowadzana podstawa programowa ma obowiązywać dopiero od roku szkolnego 2012/2013 (I klasa liceum, liceum profilowanego i technikum). Z dużym prawdopodobieństwem należy bowiem przyjąć, że do tego czasu wiedza młodzieży dotycząca zagadnień związanych z teleinformatyką znacząco wzrośnie w drodze samokształcenia. Przykładowo. Dwa pierwsze cele kształcenia tak dla etapu III jak i IV związane z bezpiecznym posługiwaniem się komputerem i wyszukiwaniem i gromadzeniem informacji wydają się być zbyt trywialne dla dzisiejszej młodzieży, która spędza wiele godzin przy komputerach i Internecie w domu. Sprawność pozyskiwania przez nią informacji można ocenić na podstawie przygotowywanych, często profesjonalnych ściąg i gotowców na różne tematy. Trudno oczekiwać od

ucznia, aby na lekcji informatyki poświęconych na przykład nauce pozyskiwania materiałów wideo z Internetu wykazywał zainteresowanie tym tematem, gdy w domu ma uruchomiony program do ściągania z sieci filmów, czy muzyki i tylko czeka by po powrocie do domu je sobie obejrzeć. Ponadto w domu ma najprawdopodobniej lepszy sprzęt niż ten, którym dysponuje szkoła.

Analogiczna sytuacja występuje przy realizacji celu trzeciego – komunikowanie się za pomocą komputera i technologii informacyjno – komunikacyjnych. Wymaganie od uczniów nauki zakładania konta pocztowego i udziału w forach dyskusyjnych w przypadku, gdy większość z nich już takie konta posiada i to często na różnych serwerach, wydaje się być niecelowe. Natomiast umiejętność komunikowania się z rówieśnikami przez sieć, większość z nich ma opanowane do perfekcji przy okazji uczestniczeniu w grach sieciowych.

Przykładów tych można by przytoczyć więcej. Czy to znaczy, że podstawa programowa jest zła? Nie. Wszystko zależy od uszczegółowienia poszczególnych tematów i rozłożeniu głównych akcentów na tych zagadnieniach, z którymi młodzież radzi sobie gorzej. Klasycznym przykładem tutaj jest algorytmika, która jest kanwą informatyki, a której należałoby poświęcić zdecydowanie więcej miejsca. Tak jak uczyniono to dla etapu IV w wersji rozszerzonej. Podejście takie jest zresztą zgodne z poglądem coraz to większej liczby nauczycieli informatyki, że nie należy ulegać modom, ale przekazywać młodzieży treści będące bazowymi i ponadczasowymi elementami informatyki, a więc właśnie algorytmikę i programowanie. Treści, do których sama młodzież sięga rzadko i przy nauce, których potrzebuje najwięcej pomocy ze strony nauczyciela. Jest to więc niejako koncepcja powrotu do źródeł informatyki

Dobór tych treści programowych jest kluczowym zagadnieniem do rozwiązania, ale w tym przypadku zgodność specjalistów jest godna podziwu Do kanonu obu działów należą następujące zagadnienia [1, 2, 8, 10, 14, 24, 28, 32, 34]:

- Typy danych
- Modele danych oparte na drzewach, listach , zbiorach, grafach, relacjach
- Wyrażenia i instrukcje
- Operacje I/O
- Procedury i funkcje
- Cechy algorytmu: poprawność, skończoność, efektywność
- Wybór algorytmu i czas działania programu (notacja „dużego O")
- Iteracja – algorytmy iteracyjne np. sortowanie przez wybór, algorytm bąbelkowy, sortowanie kubełkowe i pozycyjne, algorytmy Hornera, Euklidesa, Eratostenesa
- Rekurencja – algorytmy rekurencyjne np. sortowanie przez scalanie, algorytm Euklidesa, (problemy: Wieże Hanoi, liczby Fibonacciego)

Zagadnienia te są uwzględnione w celach szczegółowych dla IV etapu kształcenia dla przedmiotu informatyka na poziomie rozszerzonym w podstawie programowej z roku 2009.

W realizacji zadań szczegółowych należy również zwrócić baczną uwagę na zagadnienia związane z ochroną własności intelektualnej w Internecie. Wobec powszechnej praktyki, wśród młodzieży, nielegalnego pobierania z sieci filmów, utworów muzycznych czy wręcz prac przejściowych i dyplomowych problemy ochrony własności intelektualnej i odpowiedzialności karnej za jej naruszanie winny być objęte szczególną uwagą w procesie dydaktycznym.

3. Nauczyciel i jego rola wobec zachodzących zmian w programach nauczania informatyki. zmian dydaktyce informatyki i technologii informacyjnej

Jest faktem, nie podlegającym dyskusji, że tylko dobrze i nowocześnie wykształceni nauczyciele są wstanie przygotować młode pokolenie do pełnego korzystania z dobrodziejstw życia w światowym społeczeństwie informacyjnym, w którym informatyka i technologia informacyjna odgrywa jedną z kluczowych ról.

Rolą nauczyciela, jest więc przygotowanie młodego pokolenia do krytycznej analizy, oceny i twórczego wykorzystania napływającej zewsząd informacji. Aspekty techniczne pozyskiwania i przekształcania jednej formy informacji w drugą są sprawą drugorzędną. Ignorowanie jednak konieczności praktycznego (technicznego) przygotowania nauczyciela do korzystania z nowoczesnych środków informatyki i teleinformatyki wydaje się działaniem, co najmniej pochopnym. Musi się on bowiem sprawnie posługiwać się infrastrukturą informatyczną dostępną w szkole. Infrastruktura ta, na którą składa się sprzęt i oprogramowanie, podlega ciągłej modyfikacji (unowocześnianiu). Występuje tutaj zjawisko sprzężenia zwrotnego pomiędzy dydaktyką przedmiotu i wykorzystywania jego technik dla realizacji tejże dydaktyki. Fakt ten stawia przed nauczycielem dodatkowe wyzwania.

Dotyczą one przede wszystkim biegłej umiejętności wykorzystywania oddanych mu do dyspozycji środków technicznych i łączenia ich możliwości z potrzebami procesu dydaktycznego. Nauczyciel staje się niejako spoiwem łączącym technikę teleinformatyczną z metodyką nauczania.

Kwestia merytorycznego przygotowania nauczycieli i ich umiejętności radzenia sobie w sytuacjach częstokroć trudnych i stresujących jest kluczowym elementem procesu dydaktycznego. Z obserwacji i doświadczenia autorów wynika, że największy dyskomfort psychiczny odczuwa nauczyciel, który musi prowadzić zajęcia na kiepskim sprzęcie (uczniowie posiadają w domu komputery lepsze o kilka generacji), a w klasie znajduje się jeden lub kilku uczniów górujących wiedzą i umiejętnościami nad pozostałymi uczniami (a

częstokroć i nad nauczycielem). W takim przypadku, tylko umiejętne pozyskanie sobie przez nauczyciela tych najzdolniejszych uczniów jako pomocników (asystentów) w prowadzonym procesie dydaktycznym i wyzbycie się własnych częstokroć wygórowanych ambicji umożliwi bezkonfliktowe prowadzenie lekcji. Próby konfrontacji najczęściej prowadzą do klęski i załamania się procesu dydaktycznego.

Zastosowanie nowoczesnych technologii w dydaktyce zmusza wszystkich nauczycieli a przede wszystkim nauczycieli informatyki do ciągłego doskonalenia swoich umiejętności. Nauczyciele w procesie tym nie powinni zostać osamotnieni. Niezbędna jest tutaj instytucjonalna pomoc państwa np. w postaci grantów na studia podyplomowe czy kursy dokształcające.

4. Podsumowanie

Tempo zmian zachodzące we współczesnym świecie wymusza stałe modyfikowanie procesu dydaktycznego. Związane jest to z jednej strony z ciągłym dynamicznym przyrostem wiedzy, a z drugiej z konkurencją w przemyśle i nauce na poziomie globalnym, do której powinni być przygotowani absolwenci polskiego systemu oświaty.

Problemy te ze szczególnym nasileniem pojawiają się w dydaktyce przedmiotów wysokich technologii, do jakich należy zaliczyć informatykę. Konieczność rozwiązywania tych problemów skłania do sformułowania następujących wniosków, które przedstawione w pracy [11] nie straciły nic na swojej aktualności:

- Istnieje pilna potrzeba przedstawiania młodzieży, szczególnie na niższych etapach kształcenia, całego bogactwa informatyki jako nauki w celu umożliwienia jej świadomego wyboru przyszłego zawodu informatyka i wyeliminowania licznych, występujących obecnie, rozczarowań i nieporozumień wynikających z rozbieżności pomiędzy społeczną percepcją pojęcia informatyki a jej rzeczywistą funkcją jako dziedziny wiedzy.
- Zmiany programów nauczania informatyki, co do treści jak i form dydaktycznych powinny być przeprowadzane systematycznie w rytm zmieniających się uwarunkowań zewnętrznych. Powinny one być poprzedzone szeroką dyskusją wśród osób zawodowo związanych z informatyką i jej dydaktyką.
- Duża rozpiętość wiedzy i umiejętności z informatyki wśród uczniów i studentów jest czynnikiem utrudniającym proces dydaktyczny i zmusza do poszukiwania nowych skutecznych metod nauczania tego przedmiotu.
- Należy bezwzględnie przyzwyczajać uczniów i studentów do systematycznego i intensywnego przyswajania wiedzy. Wyrobienie u młodzieży tej umiejętności jest warunkiem niezbędnym do dalszego

samokształcenia się, co w przypadku tak dynamicznie rozwijającej się dziedziny, jakim jest informatyka ma znaczenie kluczowe.

- Wobec zróżnicowania w predyspozycjach intelektualnych młodzieży (niezawinionych przecież przez nauczycieli) należy w trybie pilnym rozważyć możliwość podziału szkół (klas) na szkoły (klasy) realizujące podstawowy standard programowy i na takie gdzie rodzice godząc się na wyższą dyscyplinę programową i wychowawczą chcą zapewnić swoim dzieciom uzyskanie wykształcenia adekwatnego do ich możliwości emocjonalnych i intelektualnych (licznym przeciwnikom takiego poglądu należy wskazać szkoły mistrzostwa sportowego, do których przyjmowano przecież młodzież o określonych predyspozycjach fizycznych i nie budziło to społecznego sprzeciwu)
- Szkoła o charakterze ogólnym nie jest w stanie przekazać wymaganych treści nauczania wynikających z rozwoju cywilizacyjnego współczesnego świata ze względu na:
  - braki w wyposażeniu
  - niską dyscyplinę nauczania
  - małe wymagania w stosunku do uczniów (nauczanie bezstresowe)
- Nauczyciel informatyki musi bezwzględnie i systematycznie uzupełniać swoje wykształcenia. W procesie tym powinien być wspierany przez odpowiednie programy ministerialne np. bezpłatne studiowanie w trybie studiów podyplomowych. Pozostawienie kwestii samokształcenia jako „prywatnego" problemu nauczyciela przedmiotów wysokich technologii jest niedopuszczalne i nieodpowiedzialne.
- Rzeczywista realizacja treści szczegółowych zawartych w podstawie programowej z 2009 roku dotyczącej przedmiotu informatyka zależy przede wszystkim od wiedzy i umiejętności dydaktycznych nauczycieli oraz wyposażenia szkoły.

Autorzy mają świadomość, że część przedstawionych wniosków może wzbudzi kontrowersje i polemiki, niemniej przedyskutowanie ich teraz powinno umożliwić wypracowanie wytycznych pozwalających lepiej przystosować polskie szkolnictwo do wyzwań globalizującego się świata.

LITERATURA

1. ACM Model High School Computer Science Curriculum, na stronie http://www.acm.org.
2. Aho A.V., Ullman J.D., Wykłady z informatyki z przykładami w języku C, Wyd. Helion, Gliwice 2003.
3. Association for Computing Machinery, strona organizacji http://www.acm.org.
4. Cele i kierunki rozwoju społeczeństwa informacyjnego w Polsce. Warszawa 28 listopada 2000, na stronie http://kbn.icm.edu.pl/cele/index.html:

5. Computing Curricula. Final Draft –December 15, 2001 (CC2001 Report), na stronie http://www.acm.org.
6. Denning P.J., Comer D.E., Gries D., Mulder M. C., Tuckner A., Turner J., Young P.R., Computing as a Discipline, materiały z konferencji organizowanej przez ACM, 32:1, str. 9-32, styczeń 1989.
7. Edukacja informatyczna 2002, Ministerstwo Edukacji Narodowej i Sportu, Warszawa 2002 na stronie www.men.waw.pl/.
8. Gurbiel E., Kołczyk E., Krupicka H., Łukojć K., Płoski Z., Sysło M.M., Zuber R., Elementy informatyki. Rozwiązania zadań pod redakcją M.M. Susły, Wyd. Naukowe PWN, Warszawa 1994.
9. Gurbiel E., Hardt-Olejniczak G., Kołczyk E., Krupicka H., Sysło M.M., Technologia informacyjna. Kształcenie w zakresie podstawowym. Podręcznik dla liceum ogólnokształcącego, liceum profilowanego i technikum, WSiP S.A., Warszawa 2002.
10. Harel D., Rzecz o istocie Informatyki. Algorytmika, Wyd. Naukowo-Techniczne, Warszawa 1992.
11. Iskierka S.,Krzemiński Krzemiński., Weżgowiec Z., Aspekt ekonomiczny wprowadzania nowoczesnych programów nauczania technologii informacyjnej, XIX Konferencja Informatyka w Szkole. Polskie szkoły w eEuropie. Szczecin, 10 – 13 września 2003 r. Materiały konferencyjne, ss.449 – 452
12. Iskierka S.,Krzemiński Krzemiński., Weżgowiec Z., Wybrane problemy dydaktyki informatyki i technologii informacyjnej, w Dydaktyka informatyki. Problemy uczenia się i nauczania informatyki i technologii informacyjnych, pod red. Piecucha A., Wydawnictwo Uniwersytetu Rzeszowskiego , Rzeszów 2006, ss.13 - 29
13. Kąkolewicz M., „Standardy wyposażenia i obudowy medialnej przedmiotów ogólnokształcących" na stronie www.ptm.edu.pl/standardy/
14. Koba G., Bock J., Informatyka. Podstawowe tematy. Podręcznik dla gimnazjum. Poradnik metodyczny, Wyd. Szkolne PWN, Warszawa-Wrocław 1999.
15. Koba G., Technologia informacyjna dla szkół ponadgimnazjalnych, Wyd. Migra Sp. z o.o.
16. Komisja Europejska Dyrektoriat Generalny ds. Edukacji i Kultury. Edukacja w Europie: różne systemy kształcenia i szkolenia - wspólne cele do roku 2010. Program prac dotyczący przyszłych celów systemów edukacji; na stronie www.menis.gov.pl/
17. Krawczyński E., Talaga Z., Wilk M., Technologia informacyjna nie tylko dla uczniów, Wydawnictwo Szkolne PWN, Warszawa 2002.
18. Lewicki J., Informatyka w szkole. Podręcznik. Cz. I, Oficyna Edukacyjna * Krzysztof Pazdro. Warszawa 1999.
19. Nowakowski Z., Sikorski W., Informatyka bez tajemnic. Cz. I obsługa mikrokomputerów, Wyd. IV. Wyd. MIKOM, Warszawa 1996.
20. Nowakowski Z., Sikorski W., Informatyka bez tajemnic. Cz. II użytkowanie mikrokomputerów, Wyd. II. Wyd. MIKOM, Warszawa 1995.
21. Nowakowski Z., Sikorski W., Informatyka bez tajemnic. Cz. III programowanie mikrokomputerów. Wyd. IV, Wyd. MIKOM, Warszawa 1995.
22. Nowakowski Z., Dydaktyka informatyki w praktyce. Wybrane zagadnienia. Informatyka bez tajemnic. Cz. IV, Wyd. I, Wyd. MIKOM, Warszawa 1996.

23. Nowakowski Z., Sikorski W., Informatyka dla gimnazjalisty bez tajemnic, Wyd. MIKOM, Warszawa. 2000.
24. Paluszyński W., Kurs Informatyki z ćwiczeniami. Unix, Pascal i struktury danych, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1995.
25. Papert S., Burze mózgów. Dzieci i komputery, Wyd. Naukowe PWN, Warszawa 1996.
26. Pawłowski K.: Polskie grzechy wyższe; Tygodnik Powszechny, nr 4/2003, na stronie http://kiosk.onet.pl/1109725,1,2,242,druk.html
27. Piecuch A., Edukacja informatyczna na początku trzeciego tysiąclecia, Wyd. Oświatowe FOSZE, Rzeszów 2008
28. Płoszajski G., Elementy informatyki. Program szkoły średniej zawodowej oraz ogólnokształcącej, Wyd. Szkolne i Pedagogiczne, Warszawa 1995.
29. Raport – Program nauczania informatyki. ACM/IDEE-CS Curriculum Task Force, na stronie http://www.acm.org
30. Rozporządzenie Ministra Edukacji Narodowej i Sportu z dnia 6 listopada 2003 r. zmieniające rozporządzenie w sprawie podstawy programowej wychowania przedszkolnego oraz kształcenia ogólnego w poszczególnych typach szkół. Dz. U. z 2003 r. Nr 210 poz. 2041.
31. Rozporządzenie Ministra Edukacji Narodowej z dnia 23 grudnia 2008 r. w sprawie podstawy programowej wychowania przedszkolnego oraz kształcenia ogólnego w poszczególnych typach szkół. Dz.U. z 2009 r. Nr 4 poz. 17
32. Sysło M.,M., Algorytmy, Wyd. Szkolne i Pedagogiczne, Warszawa 1997.
33. Walat A.: Elementy informatyki dla szkół średnich. Cz. I. Wydawnictwo edukacyjne. Warszawa. 1993.
34. Wirth N., Algorytmy + Struktury danych = Programy, Wyd. Naukowo-Techniczne, Warszawa 1989.

Polskie Towarzystwo Informatyczne
Oddział Górnośląski

ul. J. Lompy 2/10, 40-040 Katowice
www.pti.katowice.pl
Katowice@pti.org.pl

ISBN: 978-83-60810-28-6