



Gliwice, 01.03.2014 r.

**Recenzja pracy doktorskiej mgr inż. Stanisława Widła  
„Analiza architektury i wydajności Systemów Odpowiedzi Fonicznej”**

**Kontekst pracy.**

Celem rozprawy jest analiza architektury i wydajności sieciowego systemu, który do komunikacji z użytkownikiem wykorzystuje mowę. Rozprawa porusza kilka wątków, łącząc prace Autora dotyczące systemów fonicznych, sieci komputerowych, systemów przemysłowych z ograniczeniami czasowymi i telefonii komputerowej. System Odpowiedzi Fonicznej (ang. *Interactive Voice Response*) w którym wykorzystuje się mowę (ciąg poleceń głosowych) do komunikacji użytkownika z systemem komputerowym. Aktualnie wdrażane systemy odpowiedzi fonicznej są proste, obejmują kilka – kilkanaście modułów dialogowych, działają w oparciu o sieci telekomunikacyjne i są dobrze skalowalne. Uważa się, że wkrótce powstaną znacznie bardziej zaawansowane systemy, zdolne do rzeczywistego dialogu.

Rozwój platform telefonicznych o otwartym kodzie źródłowym oraz otwartych standardów, w szczególności opis serwerowych aplikacji za pomocą VoiceXML spowodowały, że systemy odpowiedzi fonicznej są w stanie wykorzystywać sieci TCP/IP. Aplikacje foniczne muszą zachować rygor pracy czasu rzeczywistego nie tylko w stosunku do warstwy transportowej przesyłanego dźwięku, lecz także na poziomie współpracy elementów oprogramowania, które mogą się znajdować w odległych systemach komputerowych. Dlatego ważna jest ocena efektywności pracy takich systemów i dążenie do wyeliminowania błędów ich pracy, co z kolei wymaga m.in. opracowania kryterium ilościowej oceny wydajności systemów interaktywnej odpowiedzi fonicznych, opracowania metod pomiarowych i przeprowadzenia pomiarów w systemach referencyjnych oraz opracowania modelu systemu.

Autor przeszedł w swoich pracach długą drogę od projektu „Multimedialny rozproszony system zarządzania Urzędem Miasta Pszów” oraz projektu odpowiedzi fonicznej w Urzędzie Miejskim w Gliwicach do badania efektywności i modelowania systemów odpowiedzi fonicznej. Jak zauważa, wykorzystanie mowy do budowy dialogu człowieka z systemem komputerowym prowadzi do rozwoju nowych metod projektowania aplikacji, które różnią się od aplikacji wykorzystujących GUI i wymagają nowych architektur systemowych spełniających wymagania czasowe wynikające z zastosowań fonicznych, zapewniających równoległą obsługę wielu użytkowników oraz możliwość modyfikacji interfejsu z użytkownikiem w trakcie działania aplikacji. Te zagadnienia doprowadziły go do analizy czasów reakcji wielowątkowego systemu odpowiedzi fonicznej opartego na sprzętowej implementacji równoległego

przetwarzania rozproszonego z przesylem komunikatów dla zmiennych parametrów wejściowych, takich jak czasy obsługi w poszczególnych zasobach, czas trwania komunikatu fonicznego i liczba wątków.

## Układ pracy

Rozprawa składa się z sześciu rozdziałów.

We wstępnym pierwszym rozdziale przedstawione są geneza pracy, tezy i układ rozprawy.

Tezy rozprawy są następujące:

- 1. Architektura systemu odpowiedzi fonicznej definiuje sposób implementacji wielokanałowości, wielowątkowości i konfigurowalności systemu. Technika programowania zwana maszyną stanów pozwala zbudować system odpowiedzi fonicznej niezależnie od mechanizmów wielowątkowości systemu operacyjnego, w tym również na jednozadaniowym systemie operacyjnym.*
- 2. Ilościowym parametrem wydajności systemu odpowiedzi fonicznej jest jego czas odpowiedzi, który można wyznaczyć, również dla procesu niestacjonarnego, na podstawie średniej liczby klientów, przepustowości systemu i średniego czasu trwania komunikatu fonicznego.*
- 3. Liczba klientów, czas trwania komunikatu fonicznego i czasy obsługi na stanowiskach wyznaczają przestrzeń zmiennych wejściowych dla modelu systemu odpowiedzi fonicznej. Dla zmiennego czasu obsługi wybranego stanowiska i liczby klientów rozwiązanie modelu pozwala wyróżnić następujące stany wydajności systemu: nieobciążenia, maksymalnego obciążenia, maksymalnego obciążenia, które spowodowało wybrane stanowisko, stanu zrównoważenia.*

W rozdziale drugim Autor przedstawia wiadomości dotyczące dialogu użytkownika z systemami komputerowymi za pomocą mowy. Zdefiniowano pojęcie fonicznego interfejsu z użytkownikiem jako ciągu poleceń i odpowiedzi, w których do wymiany informacji używa się toru fonicznego. Przyjęto, że w pracy będą wykorzystane istniejące generatory syntezy mowy z tekstu. W zależności od charakteru udostępnianej bądź odbieranej od użytkownika informacji,

wyróżniono sposoby działania systemów odpowiedzi fonicznej: AudioTekst, foniczny dostęp do baz danych, poczta głosowa, automatyczny operator. Jeżeli system realizuje funkcje określonej aplikacji, to trzeba określić podzbiór algorytmów translacji na tekst foniczny, za pomocą których można wygenerować dowolny komunikat wymagany w danej klasie aplikacji. Dynamiczna generacja nawet prostych komunikatów wymaga określonej wiedzy lingwistycznej. W pracy przedstawiono algorytmy budowania gramatyczne poprawnej wypowiedzi w języku polskim, na podstawie danych przechowywanych w systemach komputerowych (dane liczbowe, pola w bazach danych).

Obszerny trzeci rozdział dotyczy architektury systemu odpowiedzi fonicznej. Omówiono podstawowe założenia dla projektowania takiego wielodostępnego systemu oraz technicznych cech zasobów sprzętowo – programowych umożliwiającym udostępnianie usług poprzez sieć za pośrednictwem interfejsu fonicznego. Zdefiniowano pojęcie serwera fonicznego jako sys-

temu komputerowego, zwykle rozproszonego, wyposażonego w wielokanałowe interfejsy komunikacyjne, takie jak łącza telefonii stacjonarnej, komórkowej bądź sieci komputerowej, który do dialogu z użytkownikiem posługuje się mową. Omówiono funkcje kanału transmisyjnego, kart komunikacyjnych i oprogramowania systemowego, porównano oprogramowanie systemowe dostarczane wraz z kartami różnych producentów. Przeanalizowano zdarzenia zachodzące w karcie fonicznej i struktury danych służących do przekazywania informacji. Przedstawiono szczegółowo opracowaną przez Autora oryginalną bibliotekę UH (Universal Hardware) i argumentowano jej przewagę nad znanym standardem Telephony Application Programming Interface. Biblioteka pozwala na przenośność oprogramowania serwerów fonicznych pomiędzy kartami różnych producentów i pozwala na uruchomienie zaprojektowanych aplikacji fonicznych na kartach wybranych producentów. Omówiono sposoby implementacji współbieżnego sterowania. Program serwera fonicznego może obsługiwać zdarzenia przychodzące z kanału poprzez funkcje blokujące działanie aplikacji na czas obsługi zdarzenia. Możliwe też jest zastosowanie funkcji, które tylko inicjują obsługę zdarzenia, a następnie oddają sterowanie aplikacji, co pozwala na napisanie wielowątkowej aplikacji niezależnej od mechanizmów obsługujących wielozadaniowość w konkretnym systemie operacyjnym. Zaproponowano wielokanałową, wielowątkową i konfigurowalną architekturę oraz zbudowano system, który miał zastosowania komercyjne. Dla własnej implementacji systemu przyjęto założenia dotyczące autonomicznej pracy karty, opisu aplikacji fonicznej, jednoczesnej pracy torów fonicznych, przestrzegania ograniczeń czasowych.

W rozdziale czwartym przedstawiono analizę wydajności systemu odpowiedzi fonicznej.

Przyjęto czas odpowiedzi systemu jako podstawowy parametr jego wydajności. Zaproponowano model kolejkowy, w którym zasoby systemu odpowiadają stanowiskom obsługi. Dane dla modelu - lista operacji wraz z czasem ich wykonania zostały zaczerpnięte z dziennika, którego format zmodyfikowano tak, by zawierał wszystkie niezbędne dane dla modelu. Jednocześnie, aby zmniejszyć błędy pomiarowe uruchomiono system na jednozadaniowym systemie operacyjnym. Przedstawiono górne i dolne ograniczenia asymptotyczne dla czasu odpowiedzi systemu. Wykorzystano analizę wartości średnich (MVA, *Mean Value Analysis*) jako metodę modelowania i pokazano wpływ doboru czasu trwania komunikatu fonicznego, na czas reakcji systemu jako podstawę poprawy pracy systemu bez konieczności zmiany kodu lub architektury systemu. Pokazano, że odczyt czasu wykonania operacji z dziennika obarczony jest błędami systematycznym – metody i losowym. Pokazano, jak korygować te błędy, a przedstawiona analiza dyskretnych funkcji gęstości prawdopodobieństwa czasów odpowiedzi może mieć znacznie szersze zastosowania niż tylko korekta błędów zapisu danych do dziennika. Jeżeli z sygnału czasowego pozyska się dyskretne funkcje gęstości prawdopodobieństwa, to wynikiem złożenia niezależnych systemów jest spłot dyskretnych funkcji gęstości prawdopodobieństwa. Rozpatrywany spłot i rozplot dyskretnych funkcji gęstości prawdopodobieństwa parametrów systemu komputerowego są interesujące i mogą być stosowane np. w analizie najgorszego przypadku w systemach czasu rzeczywistego. Ponieważ okazało się, że metoda rozplotu wykazywała jest bardzo wrażliwa na błędy pomiarowe, udo-

skonalono ją poprzez wykorzystanie metody najmniejszych kwadratów. Zbadano też wpływ szerokości przedziału histogramu na dokładność uzyskiwanego rozkładu oraz problem niestacjonarnych pomiarów.

Rozdział piąty dotyczy przestrzeni parametrów określających wydajność badanego modelu systemu. Wykorzystując analityczny model w postaci sieci zamkniętej stanowisk, w której krąży stała liczba klientów, obliczono wartości średnie parametrów opisujących stan systemu. Obliczenia przeprowadzono dla wszystkich punktów znajdujących się na płaszczyźnie wyznaczonej

przez czas obsługi wybranego stanowiska i liczbę klientów. Przeprowadzono analizę wykorzystania zasobów, wprowadzono definicję maksymalnego wykorzystania zasobu oraz porównano wykorzystanie zasobu z wynikami otrzymanymi dla systemu otwartego. W przestrzeni rozwiązań modelu zostały wyróżnione obszary w których system jest nieobciążony, maksymalnie wykorzystany, maksymalnie wykorzystane jest wybrane stanowisko, system jest zrównoważony. Podano zależność określającą, kiedy wybrane stanowisko nie będzie wąskim gardłem systemu. Pokazano, że aby otrzymać zakładany poziom czasu odpowiedzi systemu nie tylko przyspieszając wszystkie jego elementy systemu, ale również zmieniając czas trwania komunikatu głosowego. Przedstawiono w postaci licznych wykresów zależność przyrostu czasu odpowiedzi systemu od czasu obsługi klienta i liczby klientów.

Rozdział szósty to podsumowanie i wnioski. Bibliografia zawiera 200 pozycji.

#### **Uwagi.**

Praca napisana jest starannie, widoczny jest duży wkład pracy Doktoranta w uzyskanie wyników, jego ciekawość badawcza, dociekliwość i konsekwencja w drążeniu tematu. Rezultaty są zarówno implementacyjne i pomiarowe (doświadczalne), jak i analityczne, składając się na interesującą całość. Wyniki prac zostały opublikowane – bibliografia zawiera dwadzieścia kilka pozycji autorstwa lub współautorstwa Doktoranta. W części analitycznej umiejętnie wykorzystana metodę modelowania opartą na analizie wartości średnich, nasuwa się jednak pytanie czy zastosowanie innej metody, np. łańcuchów Markowa czy aproksymacji dyfuzyjnej nie zmieniłoby nieco rezultatów w postaci uzyskanych obszarów pracy systemu.

#### **Podsumowanie:**

Uważam, że rozprawa doktorska mgr inż. Stanisława Widła w pełni spełnia warunki pracy doktorskiej. Autor wykazał się dużą wiedzą z zakresu projektowania i analizy systemów fonicznych; potrafił zaproponować oryginalne i konstruktywne rozwiązania. Wnioskuje o przyjęcie tej pracy jako rozprawy doktorskiej i dopuszczenie jej do publicznej obrony.



T. Widła