STUDIA INFORMATICA 2003
Volume 24 Number 2B (54)

Arkadiusz LIBER
Politechnika Wrocławska, Wydział Informatyki i Zarządzania

PSEUDOLOSOWE SIECIOWE SYSTEMY DEZINFORMACYJNE

Streszczenie. W pracy przedstawiono wyniki badań w zakresie konstrukcji pseudolosowych systemów dezinformacyjnych będące kontynuacją badań opisanych w [1], [2], [3]. Na gruncie teorii informacji przeprowadzono analizę porównawczą systemu dezinformacyjnego z systemem informacyjnym. W oparciu o system bazy wiedzy technicznej przedstawiono sposoby generacji zaburzonej losowo wiedzy z zakresu nauk matematycznych i fizycznych. Szczególny nacisk położono na niedostrzegalność zmian dokonywanych w zawartości systemów informacyjnych. W pracy przedstawiono sposoby rozpowszechniania nieprawdziwych lub niepełnych informacji przygotowywanych dynamicznie. W końcowej części pracy pokazano metody ochrony serwisów informacyjnych przed atakami dezinformacyjnymi za pomocą wprowadzenia do serwisów ukrytych warstw dodatkowych.

Słowa kluczowe: systemy dezinformacyjne, dezinformacja pseudolosowa.

PSEUDORANDOM NETWORK DISINFORMATION SYSTEMS

Summary. Because of the growing importance of electronic information it is necessary to intensify research aimed at ensuring its authenticity. It is very easy to create electronic systems containing falsified and incomplete information. Such systems, called disinformation systems, are most harmful if they become components of knowledge systems. In this paper, the results of research on the construction of disinformation systems are presented.

Keywords: disinformation systems, pseudorandom disinformation.

1. Wprowadzenie

Dezinformacja towarzyszy człowiekowi od początków jego istnienia. Najbardziej rozpowszechnione przypadki dezinformacji związane są z błędami i niedokładnościami powstają-

cymi w procesie komunikacji. Coraz częściej błędne informacje ulegają propagacji w celu zapewnienia przewagi militarnej, ekonomicznej lub zachowania własnej wiedzy. W kryptografii znane są powszechnie algorytmy ukrywania informacji bez jej szyfrowania. Algorytmy wiedzy zerowej pozwalają na wykazanie swojej wiedzy lub kwalifikacji bez ich ujawniania. Nowoczesne systemy informacyjne stanowią całokształt zasobów sprzętowych, programowych oraz wiedzy udostępnianej za pomocą mechanizmów sieciowych [4], [5]. Każdy system informacyjny składa się ze zbioru danych, fizycznych nośników danych, systemów ich przetwarzania oraz systemów prezentowania informacji. Należy podkreślić, że systemy informacyjne opisywane są w większości pozycji literaturowych w aspekcie pozytywnym. Zakłada się mianowicie, że zawarta w systemie informacyjnym wiedza wykorzystywana jest dla dobra przedsiębiorstwa, instytucji państwowej, organizacji społecznej czy ogólnie dla dobra i rozwoju człowieka.

2. Systemy informacyjne i systemy dezinformacyjne

Klasyczna definicja informacji oparta jest na procesach losowych, których elementami są zdarzenia losowe X_i. Zdarzenia te, w przypadku skończonym, tworzą zbiór zdarzeń $S=\{X_1, X_2,..., X_n\}$ i mogą odpowiadać pojawieniu się symboli (liter, cyfr itp.), zjawisk fizycznych, zjawisk społecznych i innych zjawisk w otoczeniu człowieka. Elementy takiego zbioru mogą być używane do konstruowania komunikatów przekazywanych pomiędzy uczestnikami procesu wymiany informacji. Uczestnikami takiego procesu mogą być przykładowo: człowiek-człowiek, człowiek-komputer, komputer-komputer, komputer-urządzenie peryferyjne. Przyjmując, że zbiór zdarzeń $S=\{X_1, X_2,..., X_n\}$ zawiera wszystkie zdarzenia elementarne oraz prawdopodobieństwo każdego ze zdarzeń jest znane i wynosi $p_i=p(X_i)$, otrzymujemy zbiór prawdopodobieństw $P=\{p_1, p_2,..., p_n\}$. Suma wszystkich prawdopodobieństw ze zbioru P daje wartość 1. Miarą pojedynczej informacji niesionej przez zdarzenie X_i jest tzw. ilość autoinformacji stowarzyszonej ze zdarzeniem lub entropia indywidualna:

$$I(X_i) = -\log_a p(X_i), \tag{1}$$

gdzie $a > 0$.

Podstawa logarytmu a związana jest z jednostką informacji i najczęściej przyjmuje wartość $a=2$. Formalną miarą ilości informacji przechowywanej w systemie informacyjnym lub w przesyłanej wiadomości jest entropia:

$$H(x) = -\sum_{i=1}^{n} p(X_i) \log_2 p(X_i), \tag{2}$$

gdzie: X_i – warianty treści wiadomości, $p(X_i)$ – prawdopodobieństwo wystąpienia wiadomości X_i.

Jak już wspomniano, z definicji informacji nie wynika ani jej pozytywny, ani negatywny charakter. W celu rozróżnienia systemów informacyjnych od systemów dezinformacyjnych przyjmijmy, że systemem informacyjnym jest system pierwotny skonstruowany do udostępniania pewnego zbioru informacji:

$$Z = \{z: z \in Z\}, \tag{3}$$

będącego pewnym podzbiorem zbioru wiedzy całkowitej X:

$$Z \subseteq X. \tag{4}$$

Jako system dezinformacyjny przyjmijmy system skonstruowany w celu udostępniania pewnego zbioru informacji Y, w którym przynajmniej jeden z elementów nie należy do zbioru wiedzy całkowitej X:

$$Y = \{ y : y_j \in Y \subseteq X \wedge \exists\, y_i: y_i \notin X\}. \tag{5}$$

Zbiór Y można rozłożyć na sumę podzbiorów:

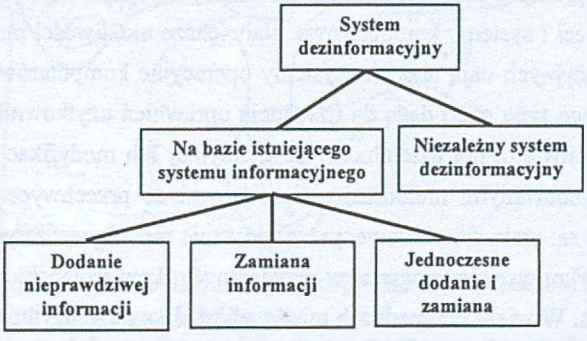
$$Y = Y_{inf} \cup Y_{dinf}, \tag{6}$$

takich że:

$$Y_{inf} = \{ y : y \in Y \cap X\}, \tag{7}$$

$$Y_{dinf} = \{ y : y \notin Y \cap X\}, \tag{8}$$

gdzie: Y_{inf} – zbiór elementów będących podzbiorem zbioru wiedzy X, Y_{dinf} –zbiór elementów dezinformacyjnych. Najprostszy system dezinformacyjny może stanowić niezależny system udostępniania wiedzy. Może jednak być utworzony na bazie istniejącego systemu informacyjnego. Utworzenie systemu dezinformacyjnego na bazie systemu informacyjnego możliwe jest przez: (a) dodanie do systemu informacyjnego elementu $y_i \notin X$, (b) zamianę w systemie informacyjnym elementu z_j: $z_j \rightarrow y_i \notin X$, (c) dodanie zbioru elementów nie należących do X z jednoczesną zamianą elementów należących do Z.



Rys. 1. Schemat konstrukcji systemów dezinformacyjnych
Fig. 1. The schema of the construction of disinformation systems

W celu porównania systemów informacyjnych z systemami dezinformacyjnymi należy wprowadzić miarę μ. Najprostszą miarą dezinformacji,jaką można zaproponować, jest miara oparta na liczności zbiorów informacji $\overline{\overline{Y}}_{\inf}$ oraz $\overline{\overline{Y}}_{d\inf}$:

$$\mu_d = \overline{\overline{Y}}_{d\inf} = \overline{\overline{Y}} - \overline{\overline{Y}}_{\inf}. \tag{9}$$

Na bazie miary dezinformacji można wprowadzić współczynnik zniekształceń (przekłamania) α jako:

$$\alpha = \frac{\overline{\overline{Y}}_{d\inf}}{\overline{\overline{Y}}}. \tag{10}$$

Współczynnik ten przyjmuje wartość zero dla systemu informatycznego nie zawierającego elementów dezinformacyjnych oraz wartość jeden dla systemu informacyjnego, który zawiera wyłącznie przekłamane wiadomości.

Miarą uwzględniającą naturę losową informacji jest miara entropijna. Dla systemu dezinformacyjnego można ją przyjąć jako wartość funkcji entropii wyrażonej zależnością (2) określoną na zbiorze $\overline{\overline{Y}}_{d\inf}$:

$$\mu_{Hd} = H\left(Y_{d\inf}\right). \tag{11}$$

3. Ataki na systemy informacyjne i lokalizacja systemów dezinformacyjnych

Teoretycznie najskuteczniejszym atakiem na system informacyjny jest modyfikacja zbioru danych przy zachowaniu nośników danych, systemu przetwarzania danych i systemu prezentacji informacji. W praktyce najprostszym atakiem może okazać się atak na system prezentacji informacji. W praktyce ataki na systemy informacyjne oparte są najczęściej na klasycznych atakach na sieci i systemy komputerowe. Największe możliwości modyfikacji działania systemów informacyjnych dają ataki na systemy operacyjne komputerów umieszczonych w sieci. Wszystkie tego typu ataki dążą do uzyskania uprawnień użytkownika systemu wystarczających do przeprowadzenia działalności destrukcyjnej lub modyfikacji zawartości systemu. Najczęściej stosowanymi metodami zmierzającymi do przechwycenia identyfikatora i hasła użytkownika są: ataki słownikowe, przeszukiwanie metodą wyczerpującą, podglądanie, podsłuch łącza, wykorzystanie programów rezydentnych, koni trojańskich oraz luk w systemach operacyjnych. W wielu przypadkach możliwe jest skłonienie użytkownika do pobrania i wykonania programu modyfikującego jego zasoby. Program taki przesyłany jest pocztą elektroniczną, z zawartością pobieranych stron WWW za pomocą cookies. W przypadku działań

destrukcyjnych i przeprowadzonych na dużą skalę łatwo jest zidentyfikować rodzaj ataku i jego źródło. W przypadku gdy intruz dokonuje jedynie niewielkich zmian, może to pozostać niezauważone.

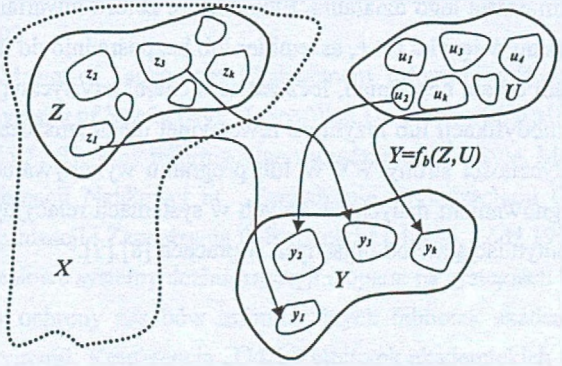
4. Systemy dezinformacyjne z pseudolosową modyfikacją informacji

Najbardziej efektywną techniką tworzenia systemów dezinformacyjnych jest modyfikacja istniejącego systemu informacyjnego. Podejście takie zapewnia możliwość sterowania poziomem zniekształceń α udostępnianej wiedzy. Jest to istotne ze względu na możliwość utworzenia informacyjnego systemu wielofunkcyjnego sterowanego liczbą dostępów do serwisów.

Przyjmijmy, że mamy pierwotny system informacyjny zawierający podzbiór elementów ze zbioru wiedzy $Z \subseteq X$. Załóżmy, że mamy zbiór elementów dezinformacji U. Utworzenie systemu dezinformacyjnego na bazie Z i U polega na określeniu odwzorowania f:

$$Y = f(Z,U). \tag{12}$$

Odwzorowania f, ze względu na skończone moce zbiorów Z, U w rzeczywistych systemach, tworzą rodzinę indeksowaną $\{f_n\}$. Ograniczając odwzorowania do kategorii zamiany k - elementowego zbioru elementów $Z'=\{z_1,z_2,...,z_k\}$ na k - elementowy zbiór elementów $U'=\{u_1,u_2,...,u_k\}$, można wprowadzić wektor binarny $b=b_1b_2...b_k$ którego bity określają numery elementów podlegających zamianie. W taki sposób wektor zerowy określa brak modyfikacji pierwotnego systemu informacyjnego. Wektor b składający się z samych jedynek określa wymianę wszystkich elementów zbioru Z' na elementy ze zbioru U'.



Rys. 2. Ilustracja syntezy systemu dezinformacyjnego na bazie systemu informacyjnego za pomocą funkcji f sterowanej wektorem b

Fig. 2. The illustration of the synthesis of the disinformation system on the basis of the data system by means of the function f of controlled with the vector b

Wybór wektora b w konkretnym systemie może być zdeterminowany przez użytkownika lub generowany automatycznie. Automatyczna generacja wektora b może być wykonana za pomocą generatora liczb pseudolosowych. Z przeprowadzonych wcześniej badań [1], [2], [3], [5], [7] wynika, że system dezinformacyjny staje się trudniejszy do zlokalizowania, jeżeli po wygenerowaniu pseudolosowego wektora b poddawany jest on nieliniowemu przekształceniu, które sterowane jest ilością dostępów w do serwisów:

$$b = g(w, p), \tag{13}$$

gdzie w – funkcja identyfikacji dostępów do systemu przez użytkowników, p – funkcja pseudolosowej generacji ciągu binarnego o zadanym rozkładzie.

5. Zabezpieczanie systemów informacyjnych przed modyfikacją informacji

Zabezpieczenie serwisów informacyjnych przed atakami dezinformacyjnymi jest niezwykle trudne. Zmiana informacji dokonywana jest nie tylko przez zewnętrznych włamywaczy. Często zmiany dokonywane są przez użytkowników autoryzowanych, którzy dokonują zmian bez weryfikacji źródła danych. Najbardziej efektywną formą uwierzytelniania serwisów jest dołączanie do nich podpisów i certyfikatów oraz sygnatur. Najciekawszą formą sygnatur cyfrowych są sygnatury niewidoczne umiejscowione w obrazach lub dźwiękach umieszczonych na stronach WWW [8],[9]. Oprócz takiej formy możliwe jest zamieszczenie niewidocznych sygnatur w oprogramowaniu komputerowym oraz w kodzie HTML stron internetowych. Istotą takiego zabezpieczenia jest wybór inwariantów, które dodane do kodu programu lub dokumentu HTML nie zmieniają jego działania. Elementy ze zbioru inwariantów wstawiane są w kod źródłowy programu w języku C++, assembler lub bezpośrednio do dokumentu HTML. Po modyfikacji program działa bez zmian, lecz zawiera charakterystyczną sygnaturę. Sygnatura taka przy próbie modyfikacji lub inżynierii rewersyjnej ulega zniszczeniu i możliwe jest stwierdzenie nieautentyczności strony WWW lub programu wykonywalnego. Trwają prace nad niewidocznym sygnowaniem danych zawartych w systemach relacyjnych i obiektowych baz danych, będące kontynuacją metod opisanych w pracach [8],[9].

6. Wnioski

Z punku widzenia teorii informacji wykonanie podstawowego systemu dezinformacyjnego jest identyczne jak systemu informacyjnego. Jedyną różnicą jest arbitralny wybór elementów należących do zbioru Z oraz do zbioru Y. W ramach pracy wykonano modyfikację wcze-

śniejszego modelowego systemu informacyjnego opartego na bazie wiedzy technicznej z zakresu matematyki [3]. Mimo stosunkowo małego rozmiaru bazy wiedzy ograniczonej do zależności symbolicznych, obejmujących kilkaset całek oznaczonych i nieoznaczonych, stwierdzono dużą odporność wykonanego systemu informacyjnego na ataki statystyczne. Pokazano, że na bazie prawidłowo działającego systemu informacyjnego można łatwo utworzyć system dezinformacyjny. Jako formy ochrony systemu dezinformacyjnego można przyjąć identyczne rozwiązania jak w systemach informacyjnych. Możliwe jest skonstruowanie autonomicznego systemu dezinformacyjnego zmieniającego automatycznie zawartość stron WWW w Internecie. Bardzo skuteczną metodą propagacji serwisów dezinformacyjnych jest umożliwienie pobrania przez użytkownika kopii systemu w celu umieszczenia jej w jego własnym systemie informacyjnym. Szczególnie podatne na rozpowszechnianie są systemy dezinformacyjne o dużych bazach wiedzy udostępniane bezpłatnie. Prowadzone są dalsze prace nad teorią oraz budową systemów dezinformacyjnych [1],[3].

LITERATURA

1. Liber A.: Watermark and Invisible Information Layers in Electronic Publications and Information Systems. From Printed Book to Digital, 13-14.10.2002, Tartu University Library.
2. Liber A.: Systemy informacyjne z wiedzą ukrytą w obrazach i dźwiękach. Multimedialne i Sieciowe Systemy Informacyjne, MISSI Kliczków, 19-20 września 2002.
3. Liber A.: Inteligentne sieciowe systemy dezinformacyjne oparte na systemie bazy wiedzy technicznej. Multimedialne i Sieciowe Systemy Informacyjne, MISSI Kliczków, 19-20 września 2002.
4. Durrani T.: Master of Technology Management. Information systems for advanced engineering. University of Strathclyde, Heriot-Watt University, STAMP.
5. Lewandowski J.: Systemy informacyjne zarządzania produkcją. Materiały V Międzynarodowej Konferencji Naukowej nt. „Zarządzanie Organizacjami Gospodarczymi", cz. I, Wydział Organizacji i Zarządzania Politechniki Łódzkiej, Łódź 1998.
6. Liber A.: Sieciowe systemy dezinformacyjne oparte na systemach baz wiedzy technicznej oraz metody ochrony zasobów informacyjnych bibliotek akademickich przed atakami dezinformacyjnymi. Konferencja „Udział bibliotek akademickich w kształtowaniu społeczeństwa informacyjnego w Polsce – potencjał, możliwości, potrzeby". Bydgoszcz-Klonowo 2002.
7. Liber A.: Statystyczne bazy danych w sieciowych systemach informacyjnych jako kompromis między zaspokajaniem potrzeb informacyjnych społeczeństwa a ochroną informa-

cji. Konferencja „Udział bibliotek akademickich w kształtowaniu społeczeństwa informacyjnego w Polsce – potencjał, możliwości, potrzeby". Bydgoszcz-Klonowo 2002.
8. Liber A.: Niewidoczne sygnatury obrazowe w rozproszonych systemach identyfikacji biometrycznej. III Krajowa Konferencja Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim. Kraków 19-21.11. 2002.
9. Liber A.: Niewidoczne Pseudolosowe Sygnatury Obrazowe w Zabezpieczaniu Publikacji Elektronicznych. Problemy Ochrony Zbiorów i Systemów Komputerowych w Bibliotekach. Białystok-Wigry, 6-8 czerwca 2001.

Recenzent: Dr inż. Andrzej Białas

Wpłynęło do Redakcji 5 kwietnia 2003 r.

Abstract

This research work is based on the results obtained in [1], [2] and [3], constituting their extension and supplementation. Because of the growing importance of electronic information it is necessary to intensify research aimed at ensuring its authenticity. It is very easy to create electronic systems containing falsified and incomplete information. Such systems, called disinformation systems, are most harmful if they become components of knowledge systems. In this paper, the results of research on the construction of disinformation systems are presented. Theory of disinformation system with pseudorandom mapping are presented. Possibilities of using hidden information layers for protecting of electronic information systems against disinformation attacks are described.

Adres

Arkadiusz LIBER: Politechnika Wrocławska, Wydział Informatyki i Zarządzania, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Polska, liber@ci.pwr.wroc.pl