MECHANIZMY PROFILOWANIA UŻYTKOWNIKÓW W SIECI – WYKORZYSTANIE, ANALIZA, ZAPOBIEGANIE

Michał Zarychta

Plan prezentacji

- Prywatność i anonimowość w Internecie
  - Aspekty prawne
  - Tworzenie profilu
  - Wykorzystanie profili
- Google
- Atak na prywatność
- Ochrona prywatności
- Projekt (element pracy magisterskiej)

Prywatność i anonimowość w Internecie

- **Prywatność** – „osobista własność, niepodlegająca państwu ani żadnym instytucjom publicznym, dotycząca spraw osobistych i rodzinnych”
- **Anonimowość** – „niemożność identyfikacji tożsamości jednostki pośród innych członków danej społeczności, wprost w odniesieniu do osoby albo do pochodzącego od niej przedmiotu”

Prywatność i anonimowość w Internecie – aspekty prawne (1/4)

- Ustawa z dnia 29 sierpnia 1997 roku o ochronie danych osobowych
  - Każdy ma prawo do ochrony dotyczących go danych osobowych
  - Przetwarzanie danych osobowych to ich zbieranie, udostępnianie, opracowywanie, przechowywanie, utrwalanie, zmienianie i usuwanie – za zgodą osoby lub w szczególnych warunkach

Prywatność i anonimowość w Internecie – aspekty prawne (2/4)

- Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 29 kwietnia 2004 roku
  - Sposób prowadzenia i zakres dokumentacji, która opisuje jak należy przetwarzać dane osobowe oraz środki techniczne i organizacyjne, których celem jest zapewnie ochrony przetwarzanych danych osobowych
  - Warunki techniczne i organizacyjne w odniesieniu do urządzeń i systemów informatycznych służących do przetwarzania danych osobowych
  - Wymagania dotyczące odnotowywania sytuacji udostępniania danych i bezpieczeństwa w przetwarzaniu danych osobowych

Prywatność i anonimowość w Internecie – aspekty prawne (3/4)

- Rozporządzenie Parlamentu Europejskiego i Rady z dnia 18 grudnia 2000 roku
  - „Sankcje dla tych, którzy naruszają przepisy i monitorowanie przez niezależny organ nadzoru”
  - „Spójne i jednolite stosowanie zasad ochrony podstawowych praw i wolności osób fizycznych w odniesieniu do przetwarzania danych osobowych powinno być zapewnione w całej Wspólnocie”

Prywatność i anonimowość w Internecie – aspekty prawne (4/4)

- Dyrektywa Parlamentu Europejskiego i Rady z dnia 12 lipca 2002 roku
  - Określenie praw dla sektora komunikacji elektronicznej
  - Zapobieganie dostępowi do komunikatów w publicznych sieciach komunikacyjnych w celu ochrony poufności komunikacji

Prywatność i anonimowość w Internecie – Tworzenie profilu

- Analiza danych
  - Częste ścieżki nawigacji
  - Maksymalne odwołania w przód
- Odkrywanie wzorców
  - Odkrywanie reguł asocjacji
  - Odkrywanie reguł sekwencji
- Klasyfikacja
  - Analiza wzorców
  - Grupowanie

Prywatność i anonimowość w Internecie – Wykorzystanie profilu

- Marketing
  - Spersonalizowana **reklama**
  - Sprzedaż profili (FBI, RIAA, firmy ubezpieczeniowe)
- Kradzież tożsamości
  - Dane pozwalające na fałszywe uwierzytelnienie (socjotechnika)
  - Terroryzm
- Kontrola pracowników
  - Zarządzanie dostępem do Internetu
  - Śledzenie aktywności
- Kontrola obywateli
  - Projekt Złota Tarcza
  - Carnivore, SORM-2, System RIP, Echelon

Google – teoria spiskowa?

- E-mail
- Google Earth/Google Maps
- YouTube
- Google Docs
- Google Calendar
- Google Analytics
- iGoogle
- Google Desktop

Atak na prywatność – Śledzenie z wykorzystaniem cookies

- Fragment tekstu przechowywany na komputerze użytkownika
- Wysyłany jako fragment nagłówka HTTP przez przeglądarkę do serwera
- Wykorzystanie
  - Ciasteczka stron trzecich
  - Przechwycenie cookies
  - Cross-site
  - Oszukane cookies

Atak na prywatność – URL i plik logu serwera

- Analiza pliku logu serwera/historii przeglądania
- Znajomość hierarchii strony/serwisu
- Brak wrażliwości na szyfrowanie
- Ograniczony dostęp
  - Administratorzy sieci
  - Dostawcy usług internetowych
  - Administratorzy sieci korporacyjnych
  - Współużytkownicy stacji roboczej

Atak na prywatność – wykorzystanie Web Beacons

- web bug, tracking bug, tracking piksel, piksel tag, 1x1 gif, clear gif
- HTML – img src
- Przekierowania na dowolny serwer
- Zalety
  - Łatwość implementacji
  - Dokładność mechanizmu
- Wady
  - Popularność

Atak na prywatność - AdSerwer

- Wykorzystanie zebranych profili
  - Upload reklam do witryn internetowych
  - Śledzenie reklam w celu profilowania
  - Kierowanie reklam zgodnie z założeniami
  - Optymalizacja i zarządzanie baza profili
  - Raportowanie skuteczności poprzez zliczanie akcji (kliknięć, odwiedzin)
- DoubleClick, Gemius

Atak na prywatność – Malware (1/3)

- **Exploit** – przejęcie kontroli poprzez wykorzystanie luk w oprogramowaniu
- **Dialer** – zmiana numeru dostępowego w modemie na 0-700 lub zagraniczny
- **Robak** – „wirusy sieciowe”
- **SQL/URL Injections** – atak na bazy danych i serwery poprzez wykorzystanie luk w oprogramowaniu

Atak na prywatność – Malware (2/3)

- **Trojan** – instaluje się na maszynie ofiary, otwiera port komunikacyjny i pozwala na przejęcie kontroli na systemem operacyjnym
  - **Spyware** – oprogramowanie szpiegujące
    - Scumware – „zamieszanie” w systemie operacyjnym
    - Adware – reklamy, pop-up
    - Hijacker BHO – dodatki do przeglądarki internetowej
    - Keylogger – zapisywanie wciskanych klawiszy
    - Stealware – przejęcie konta bankowości elektronicznej
  - **Rootkit** – ukrywanie procesów i plików
  - **BackDoor** – wykonywanie poleceń z poziomu administracyjnego
  - **Wabbit** – zapełnienie przestrzeni dyskowej

Atak na prywatność – Malware (3/3)

- **Wirus** - program lub fragment kodu, który dołącza się do innego oprogramowania w celu infekcji komputera ofiary oraz reprodukcji
  - **BOOT Sektor** – umiejscowiony w sektorze rozruchowym dysku
  - **Pasożytniczy** – wykonywalne wraz z innymi programami
  - **Wieloczęściowy** – atak wieloma sposobami
  - **Towarzyszący** – nazwy podobne do znanych programów, ale inne rozszerzenie
  - **Makro** – kod VBA

Atak na prywatność – Sniffing

- Przechwytywanie informacji
  - Poszukiwanie danych wrażliwych
  - Możliwość określenia stron komunikacji
  - Możliwość analizy sesji WWW
- Wardriving
  - Atak na bezprzewodowe sieci zabezpieczone i niezabezpieczone
- Zalety
  - Filtrowanie całego ruchu
  - Niezależność
  - Świeżość danych
- Wady
  - Skomplikowana instalacja
  - Obróbka danych (surowe pakiety)
  - Koszt

Atak na prywatność – JavaScript

- Rozbudowany mechanizm
- Zalety
  - Tani i łatwy w implementacji
  - Często jedyna możliwa technika
  - Brak wrażliwości na buforujące proxy
- Wady
  - Konieczność obsługi JavaScript
  - Przeładowanie stron skryptami
  - Popularność mechanizmu

Ochrona prywatności – blokowanie cookies

- Odpowiednie ustawienie przeglądarki internetowej lub „tryb prywatny”
- Dodatki do przeglądarek internetowych
  - Better Privacy, Cookie Button
- Możliwość ograniczenia usług
  - Koszyk w sklepie internetowym
  - Spersonalizowany widok w serwisach

Ochrona prywatności – serwery anonimizujące

- Wykorzystanie zaufanego serwera proxy
- Szyfrowanie (TLS/SSL)
- Wady
  - Duże opóźnienia
  - Przeniesienie ruchu w inne miejsce w sieci
  - Możliwość kompromitacji

Ochrona prywatności - VAST

- Jeden węzeł pośredniczący
  - Współpraca z agentem (aplet Java)
  - Połączenie szyfrowane wykorzystujące TSL/SSL
  - Przekazywanie adresów właściwych i nadmiarowych do serwera pośredniczącego
  - Selekcja otrzymanych odpowiedzi
- Generowanie ruchu nadmiarowego
  - Słownik haseł/terminów
  - Odpowiednie sesje tematyczne

Ochrona prywatności – serwery pośredniczące

- Sieci miksujące
  - Wiele serwerów
  - Szyfrowanie
  - Ruch nadmiarowy
- Anonimowe P2P
  - Opennet
    - Bez konfiguracji
    - Przypadkowe połączenia
    - Decyzja na temat bezpośrednich połączeń
  - Darknet – F2F
    - Pełna konfiguracja
    - Zaufane węzły

Ochrona prywatności - Firewall

- Kompleksowe rozwiązanie (kombajn)
  - Filtrowanie pakietów – reguły filtrowania zgodne z polityką bezpieczeństwa
  - Brama aplikacyjna – bezpieczna komunikacja dla aplikacji
  - Utrzymanie bezpiecznego połączenia
  - Serwer proxy – ukrywanie adresu sieciowego

Projekt - koncepcja

- Projekt i implementacja bezpiecznego oraz anonimowego komunikatora
  - Wykorzystanie idei MixNet
    - Ukrycie lokalizacji
    - Połączenie bez znajomości tożsamości sieciowej
    - Wzajemne świadczenie usługi serwera
  - Szyfrowanie RSA
    - Bezpieczeństwo przekazu

Projekt - rozwój

- Dodanie ruchu nadmiarowego
  - Odpowiedni harmonogram
- Algorytm określania tras
  - W danym okresie czasowym używanie jednego zdefiniowanego zbioru
- Dystrybucja kluczy szyfrowania
  - Serwer kluczy
  - Diffie-Hellman

Bibliografia

- Akhil Sahai, Sven Graupner; Web Services in the Enterprise: Concepts, Standards, Solutions, and Management
- Avinash Kaushik; Web Analytics: An Hour a Day
- Carla Schroder; Linux Networking Cookbook
- Chris Nicoll, Corien Prins, Miriam van Dellen; Digital Anonymity and the Law: Tensions and Dimensions
- David Chaum; Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms
- David Flanagan; JavaScript: The Definitive Guide, Fifth Edition
- Gene K. Landy, Amy J. Mastrobattista; The IT / Digital Legal Companion: A Comprehensive Business Guide to Software, IT, Internet, Media and IP Law
- George Meghabghab, Abraham Kandel; Search Engines, Link Analysis, and User's Web Behavior
- Igor Margasiński, Krzysztof Szczypiorski; Wszechstronna anonimowość klienta HTTP
- J.R. Okin; The Internet Revolution: The Not-for-Dummies Guide to the History, Technology, and Use of the Internet
- James F. Kurose, Keith W. Ross; Sieci komputerowe. Od ogółu do szczegółu z internetem w tle. Wydanie III
- Samuel J. Best, Brian S. Krueger; Internet Data Collection
- Jinyang Li, Frank Dabek; F2F: Reliable Storage in Open Networks
- Kancelaria Sejmu; Dz.U. 1997 Nr 133 poz. 883, USTAWA z dnia 29 sierpnia 1997 r. o ochronie danych osobowych
- Krzysztof Brzeziński, Igor Margasiński, Krzysztof Szczypiorski; Prywatne wojny w sieci: poddaj się, okop, negocjuj lub stań do walki
- Marek Wojciechowski; Odkrywanie wzorców zachowań użytkowników WWW
- Michael A. Caloyannides; Privacy Protection and Computer Forensics, Second Edition
- Peter Brusilovsky, Alfred Kobsa, Wolfgang Nejdl; The Adaptive Web: Methods and Strategies of Web Personalization
- Rannenberg Kai, Royer Denis, Deuker André; The Future of Identity in the Information Society
- RSA Laboratories; PKCS #1 v2.1: RSA Cryptography Standard
- Simson Garfinkel, Gene Spafford; Web Security, Privacy and Commerce, Second Edition
- Stuart McClure, Joel Scambray, George Kurtz; Hacking exposed 6: Network Security Secrets & Solutions
- Yuan Gao; Web systems design and online consumer behavior