Zeszyty Naukowe Wydziału
Elektroniki i Informatyki
Politechniki Koszalińskiej
ISSN:1897-7421
Nr 13, Listopad 2018
Str: 37-45

Grzegorz Górski
Paweł Koziołko
Zakład Systemów Multimedialnych i Sztucznej Inteligencji
Wydział Elektroniki i Informatyki
Politechnika Koszalińska

Analiza skuteczności wybranych metod ochrony anonimowości stosowanych w przeglądarkach internetowych

Słowa kluczowe: Podatności, aplikacje internetowe, zabezpieczenia, cookies, inwigilacja

1. Wprowadzenie

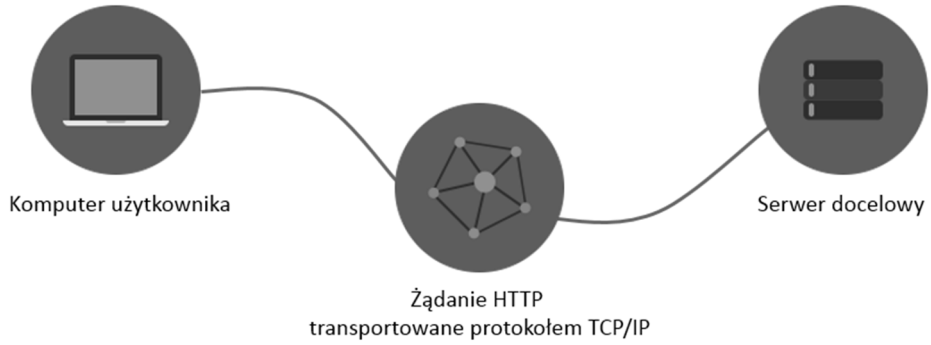
Na podstawie informacji Głównego Urzędu Statystycznego, w 2017 roku w Polsce ponad 78% gospodarstw domowych posiadało Internet szerokopasmowy [1]. Powszechny dostęp do usług oferowanych przez sieć globalną powoduje, że informacje o aktywności poszczególnych użytkowników są narażone na te same zagrożenia, co inne informacje przetwarzane w sieci Internet. Bardzo często stosowane są mechanizmy, których głównym celem jest ułatwienie dostępu użytkownika do wybranych zasobów. Jest to możliwe tylko i wyłącznie dzięki gromadzeniu informacji o preferencjach konsumenta, które są przez niego podawane. Taki zbiór danych - gromadzony jest zarówno na serwerach usługodawców, jak i w systemach operacyjnych, aplikacjach użytkowanych komputerów oraz urządzeń mobilnych stanowi bardzo dokładny opis preferencji użytkownika i może być skutecznie wykorzystany do jego identyfikacji, poszukiwania lub w najlepszym przypadku przesyłania reklam spersonalizowanych tymi informacjami.

Przeglądarki internetowe zapewniają użytkownikom dostęp do wyświetlania wysokiej jakości usług www, które wymagają informacji nie tylko na temat przeglądarki ale także środowiska systemu operacyjnego użytkownika. Dzięki różnym współczesnym narzędziom informatycznym do serwera usługowego bez wiedzy użytkownika przesyłane są szczegółowe informacje związane z konfiguracją

sprzętu i oprogramowania w celu optymalizacji korzystania użytkownika z zasobów Internetu. Pewien zasób informacji o swoich klientach jest niezbędny do świadczenia usług, jednak wiele portali gromadzi nadmiarowe informacje, które same w sobie stają się towarem oferowanym innym usługodawcom bez zgody i wiedzy użytkowników. Należy zauważyć, że nawet niewielkie różnice w informacjach o użytkownikach gromadzonych przez punkty dostępu do usług mogą być skutecznie wykorzystane do przeprowadzenia ataku na urządzenie klienta.

2. Główne obszary inwigilacji użytkownika Internetu

W literaturze naukowej zajmującej się ochroną bezpieczeństwa zdefiniowano trzy główne obszary inwigilacji tj.: urządzenie końcowe wykorzystywane przez użytkownika (komputer lub urządzenie mobilne), treść żądania http przesyłana protokołem TCP/IP a także serwer docelowy. W artykule opisano wybrane grupy zagrożeń z pierwszego z wyżej wymienionych obszarów.



Rys. 1. Główne obszary inwigilacji użytkownika Internetu

3. Podatności po stronie klienta

Urządzenie końcowe klienta jest jednostką najbardziej narażoną na ataki ze strony osób trzecich. Poprzez instalowanie zewnętrznego oprogramowania, przeglądanie stron www oraz podłączanie do niezaufanych, ogólnodostępnych sieci Wi-Fi użytkownicy nieświadomie zwiększają prawdopodobieństwo wycieku danych wrażliwych.

Do przeglądania zasobów sieci Internet wymagana jest przeglądarka internetowa, która tłumaczy język HTML, JavaScript oraz kaskadowe arkusze styli (CSS) i wyświetla wcześniej zaprogramowane strony www. Przeglądarki, oprócz tłumaczenia kodu dynamicznie prezentowanych stron www, zapamiętują również pewne informacje w celu zwiększenia komfortu korzystania z Internetu.

3.1. Pliki cookies

Pliki cookies, zwane także ciasteczkami, są to tworzone na komputerze klienta podczas komunikacji przeglądarki z serwerem, na którym znajduje się strona internetowa. Zawierają dane umożliwiające identyfikację użytkowników. Są jednym ze sposobów zapamiętywania pól formularzy, takich jak login czy dane adresowe, dzięki czemu nie ma konieczności ponownego ich wypełniania przy każdym pobieraniu lub odświeżaniu strony www. Poza polami formularzy, coraz częściej zapamiętywane są zachowania użytkownika na określonej witrynie w celu tworzenia tzw. odcisków palca użytkownika. Mechanizm ten jest ściśle powiązany z zagadnieniami UX (ang. User Experience). Zbieranie i zapisywanie informacji o kolejnych krokach może być wykorzystane do skierowania użytkownika do zachowań oczekiwanych przez właściciela strony internetowej. Najczęstszym przykładem są sklepy internetowe. Na urządzeniu klienta zapisywane są grupy artykułów, które wyświetlił odwiedzający sklep. W przypadku kolejnego podłączenia do strony zostają wyświetlane produkty z wcześniej wyświetlonej grupy w określonych miejscach na stronie internetowej, które zawierają reklamy lub podpowiedzi.

22 marca 2013 roku weszły w życie znowelizowane przepisy ustawy Prawa Telekomunikacyjnego. Najważniejsza zmiana nakłada na właścicieli serwisów internetowych obowiązek informowania użytkowników strony o plikach, które serwis umieszcza w komputerze użytkownika oraz o tym, w jakim celu to robi.

3.2. Pliki supercookies

Pliki super cookies spełniają te same funkcje, co zwykłe ciasteczka, ponieważ mogą zawierać prawie wszystkie informacje, w tym historię przeglądania, dane uwierzytelniające lub dane kierowane do reklamodawców.

Zasadniczą różnicą między plikami cookies, a super cookies jest fakt, że te ostatnie mogą być zapisywane są w ukrytej lokalizacji na komputerze klienta, przez co stają się trudne do usunięcia. Dodatkowo nie są stosowane mechanizmy automatycznego, okresowego usuwania takich plików. Biorąc pod uwagę rozmiar pojedynczego pliku super cookie często znacznie większy niż 100kB (w przypadku zwykłych plików cookies jest to z reguły ok. 4kB) przy intensywnym korzystaniu z usług sieci może to mieć wpływ na wykorzystanie zasobów w urządzeniu użytkownika. Dodatkowe zagrożenie powoduje fakt, że informacje zawarte w plikach super cookies dostępne są nie tylko w obrębie witryny uruchamianej na konkretnej przeglądarce, ale w obrębie każdej domeny, niezależnie od przeglądarki, na której zostały zapisane.

3.3. Pliki zombie cookies i evercookie

Pliki określane mianem „zombie cookies" to ciasteczka samoodnawialne po zniszczeniu (np. skasowaniu przez użytkownika). Kopie zapasowe ciasteczek tworzone są poza dedykowaną pamięcią przeglądarki przeznaczoną do zapamiętywania plików cookies lub w specyficznych przypadkach - online. Pliki tego typu mogą zostać zapisane na urządzeniu klienta bez jego wiedzy i akceptacji, nawet przy włączonej opcji blokady zapisu plików cookies w przeglądarce.

Kolejną odmianą plików gromadzących informacje o użytkowniku będące implementacją w języku JavaScript narzędzi informatycznych (API) do tworzenia tzw. permanentnych ciasteczek. Ten rodzaj pliku wykrywa akcję usunięcia ciasteczka i tworzy je na nowo z kopii zapasowej. Kopie zapasowe są tworzone w zależności od dostępności modułu na urządzeniu klienta [4] w następujących miejscach:

- Standardowych plikach cookies HTTP
- HTTP Strict Transport Security (HSTS) Pinning
- Local Shared Objects (Flash Cookies)
- Silverlight Isolated Storage
- Automatycznie generowanych wartościach RGB przy pomocy force-cached PNG i HTML5 Canvas
- Web History
- HTTP ETags
- Web cache
- window.name caching
- Internet Explorer userData storage
- HTML5 Session Storage
- HTML5 Local Storage
- HTML5 Global Storage
- HTML5 Database Storage via SQLite
- HTML5 IndexedDB
- Java JNLP PersistenceService
- Java CVE-2013-0422 exploit (applet sandbox escaping)

3.4. Skrypty JavaScript

Najczęściej wykorzystywanym mechanizmem do pozyskiwani informacji o użytkowniku są skrypty napisane w języku JavaScript. Obecnie około 95% stron www posiada zaimplementowane takie skrypty [2]. Podstawową funkcjonalnością jest wprowadzenie interaktywności między użytkownikiem a witryną. Aplikacje JS pomagają w identyfikacji użytkownika, ale poprzez dostęp do informacji o specy-

ficznej konfiguracji komputera, systemu operacyjnego, czy wersji przeglądarki, można dokonać implementacji, która stanie się narzędziem do inwigilacji i przekazywania informacji o użytkowniku.

Przykład takiej implementacji prezentuje artykuł [15], w zaprezentowane dwa podstawowe aspekty związane z poprawną identyfikacją użytkownika – pozytywną cechą unikalnej identyfikacji jest możliwość ochrony przed kradzieżą sesji zestawionej, np. z bankiem. Z drugiej jednak strony, można wykorzystać takie narzędzie do permanentnego śledzenia aktywności użytkownika.

Mechanizm „Fingerprinting", czyli oznaczanie unikalnych elementów wykorzystywanych przez klienta, w przytoczonym artykule dotyczy nie tylko rozpoznania wersji przeglądarki użytkownika, ale także określa wersji systemu operacyjnego i architektury komputera.

4. Wybrane metody ograniczające możliwość zbierania informacji o użytkowniku

Twórcy rozwiązań i technologii informatycznych, analizując podstawowe obszary architektury komputera oraz aplikacji internetowych monitorowane pod kątem zbierania informacji o użytkowniku, opracowali metody utrudniające lub ograniczające dostęp do tego typu informacji przez podmioty do tego nieuprawnione.

4.1. Tryb incognito

Obecnie każda przeglądarka posiada tzw. tryb „incognito". Według założeń producentów taki tryb prywatny zapewnia anonimowość przeglądającemu strony www. Przeglądarka z aktywną tą funkcjonalnością nie gromadzi historii przeglądania, powinna blokować zapis plików cookies oraz tworzyć czyste środowisko za każdym razem kiedy jest uruchamiana. Żaden tekst podany w polach tekstowych lub w polach wyszukiwania nie powinien zostać dodany do listy podpowiedzi automatycznego wypełniania formularzy. Tryb prywatny powinien odseparować dane przeglądarki w trybie standardowym i nie dopuścić do wycieku wrażliwych informacji [5].

Producenci przeglądarek internetowych dodają informację o trybie prywatnym, która mówi, że tryb ten nie zabezpiecza podatności związanych z architekturą protokołu TCP/IP w związku z czym dostawcy Internetu jak i osoby trzecie nasłuchujące komunikacji wychodzącej z komputera użytkownika trybu prywatnego mogą śledzić jego aktywność w sieci. Badania opublikowane w pracy [7] pokazują, że w roku 2107 jedynie przeglądarka Opera była w stanie zablokować większość podatności związanych z wyciekiem tożsamości użytkownika.

DIGITAL CITIZEN	Google Chrome	Internet Explorer	Microsoft Edge	Mozilla Firefox	Opera
Deleted data and files					
Cookies	Yes	Yes	Yes	Yes	Yes
Data from forms	Yes	Yes	Yes	Yes	Yes
Temporary files (cache)	Yes	Yes	Yes	Yes	Yes
Browsing history	Yes	Yes	Yes	Yes	Yes
Search history	Yes	Yes	Yes	Yes	Yes
Recovers closed tabs	No	Yes	No	Yes	No*
Disables add-ons and toolbars	Yes	Yes	Yes	No	Yes
Blocks trackers and advertising	No	Optional	No	Yes	Optional
VPN encryption	No	No	No	No	Optional

Rys. 2. Badania skuteczności trybu prywatnego najbardziej popularnych przeglądarek internetowych w 2017 r.

4.2. Dodatki do przeglądarek

Dodatki (ang. Plug-in) do przeglądarek www to miniaplikacje rozszerzające domyślne funkcjonalności. Takie rozszerzenia powinny być instalowane wyłącznie ze sklepów producentów. Z jednej strony mogą utworzyć interfejs ułatwiający odczytywanie wiadomości email, dodać przycisk pobierający pliki audio i wideo bezpośrednio z odtwarzacza HTML5 umieszczonego na stronie www, ale także blokować mechanizmy tworzące wirtualne odciski palców, czyli informacje pozwalające na identyfikację użytkownika na podstawie jego aktywności.

Jednym z takich dodatków jest Ghostery, którego głównym zadaniem jest blokowanie skryptów śledzących aktywność użytkownika oraz wtyczek umożliwiających korzystanie z portali społecznościowych czy komentarzy (tzw. trackerów) [6]. Wpisanie adresu strony do przeglądarki z tym uruchomionym dodatkiem powoduje przeanalizowanie i wyświetlenie zablokowanej listy skryptów śledzących. Szczegółową listę trackerów wraz z ich opisem i dokładnym adresem URL można zobaczyć klikając w ikonę Ghostery (zawierającą informację o liczbie blokad), następnie przechodząc do zakładki Widok szczegółowy.

Podobnie jak w przypadku aplikacji antywirusowych, prawidłowe działania dodatku Ghostery wymaga nieustannej aktualizacji biblioteki skryptów śledzących.

Rys. 3. Widok szczegółowy z listą zablokowanych skryptów śledzących

5. Podsumowanie

Zaprezentowane w artykule zagrożenia oraz metody ochrony informacji o aktywności użytkownika implementowane w przeglądarkach internetowych pokazały, że nie istnieją rozwiązania pozwalające na wyeliminowanie tego typu zagrożeń wyłącznie za pomocą oprogramowania instalowanego na urządzeniach końcowych użytkownika. Jest to jedyny element systemu informatycznego opartego na usługach w sieci globalnej na którym klient posiada częściową kontrolę. Ciągłe mutacje skryptów śledzących użytkownika [8] zmuszają producentów przeglądarek oraz dodatków w postaci rozszerzeń zapewniających ochronę tożsamości do aktualizowania swojego oprogramowania. Przygotowanie rozwiązania uniwersalnego wymagałoby opracowania analizatora kontekstowo-zależnych skryptów uruchamianych przez przeglądarki www. Jest to oczywiście niemożliwe, gdyż wymagałoby implementacji automatu o nieskończonej liczbie stanów. Metody

ochrony anonimowości w sieci Internet powinny stanowić większy system pośredniczący w dostępie do usług sieciowych. W takim rozwiązaniu odpowiednio skonfigurowane urządzenie klienta jest tylko częścią mającą z jednej strony funkcję prezentacyjną a z drugiej separującą dane przechowywane na komputerze użytkownika przed usługodawcą.

Bibliografia

1. Główny Urząd Statystyczny, Społeczeństwo informacyjne w Polsce 2017: https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5497/2/7/1/spoleczenstwo_informacyjne_w_polsce_w_2017.pdf
2. Beauty and the Beast: Diverting Modern Web Browsers to Build Unique Browser Fingerprints, Pierre Laperdrix; Walter Rudametkin; Benoit Baudry, 2016 IEEE Symposium on Security and Privacy
3. https://w3techs.com/technologies/details/cp-javascript/all/all
4. Sandvine, Global Internet Phenomena Raport, 2014: https://www.sandvine.com/downloads/general/globalinternet-phenomena/2014/1h-2014-global-internetphenomena-report.pdf
5. https://support.mozilla.org/pl/kb/Przegl%C4%85danie%20w%20trybie%20prywatnym
6. Sam Macbeth, Tracking the Trackers: Analysing the globaltracking landscape with GhostRank
7. https://www.digitalcitizen.life/what-is-private-browsing-which-browser-is-best
8. A. Narayanan, D. Reisman, S. Englehardt, C. Eubank, P. Zimmerman, Cookies that give you away: Evaluating the surveillance implications of web tracking
9. New Web Code Draws Concern Over Privacy Risks, https://www.nytimes.com/2010/10/11/business/media/11privacy.html?hp
10. Soltani, S. Canty, Q. Mayo, L. Thomas, C.J. Hoofnagle, Flash Cookies and Privacy
11. J. Mayer and J. Mitchell "Third-Party Web Tracking: Policy and Technology", IEEE Symposium on Security and Privacy, 2012
12. Karwowski Maciej, "Ocena skuteczności mechanizmów zapewnienia prywatności w sieci Internet", praca dyplomowa wrzesień 2014
13. M. Ayenson, D. Wambach, A. Soltani, N. Good, C. Hoofnagle " Flash Cookies and Privacy II: Now with HTML5 and ETag Respawning„
14. G. Aggarwal, E. Burzstein, C. Jackson, D. Boneh, "An Analysis of Private Browsing Modes in Modern Browsers
15. K. Mowery „Fingerprinting Information in JavaScript Implementations" – Proceedings of W2SP 2011. IEEE Computer Society, May 2011.

Streszczenie

Praca prezentuje metody śledzenia użytkownika przeglądającego Internet poprzez właściwości przeglądarek internetowych jak i słabości wynikające z architektury sieci Internet. Ponadto, w pracy przedstawiono zestaw narzędzi maskujących tożsamość użytkownika. W pracy przedstawiono innowacje w HTML5 zapewniające dostęp do wysoce wyróżniających się atrybutów użytkownika, w szczególności za pomocą interfejsu Cookie, który opiera się na wielu warstwach systemu użytkownika.

Abstract

The work presents methods of tracking the user browsing the Internet through the properties of web browsers as well as weaknesses resulting from the architecture of the Internet. In addition, the work presents a set of tools masking the user's identity. The paper presents innovations in HTML5 that provide access to highly-distinguished attributes of the user, in particular using the Cookie interface, which is based on many layers of the user's system.

Keywords: Vulnerabilities, web applications, security, cookies, surveillance