STUDIA INFORMATICA 2003
Volume 24 Number 2B (54)

Mirosław SKRZEWSKI
Politechnika Śląska, Instytut Informatyki

OCENA BIERNYCH ZAGROŻEŃ INTERNETOWYCH

Streszczenie. Systemy komputerowe przez połączenie z Internetem wystawione są na kontakty, mogące stanowić dla nich różnego rodzaju zagrożenia. Można je podzielić na zagrożenia aktywne, nadchodzące z sieci, oraz na zagrożenia pasywne (bierne), wymagające do ich uaktywnienia ingerencji ze strony użytkownika. W celu wykrycia ich obecności niezbędne jest zastosowanie specjalizowanych narzędzi, np. systemów IDS lub analizatorów protokołów. Artykuł przedstawia informacje o możliwych rozmiarach strat związanych z zagrożeniami i metodach ich oceny.

Słowa kluczowe: Internet, zagrożenia, straty zasobów.

INTERNET PASSIVE THREAT EVALUATION

Summary. Computer systems due to their Internet connections are subject to many kinds of threats. One can distinguish active threats, coming from the network without user involvement, and passive threats, activated by user intervention. To find their presence one need special tools like IDS systems or protocol analyzers (sniffers). The paper presents some information about the amount of losses, related to the passive treats and the ways of their estimation.

Keywords: Internet, threats, resouces losses.

1. Wstęp

Systemy komputerowe przez połączenie z Internetem wystawione są na kontakty z innymi systemami, mogące stanowić dla nich różnego rodzaju zagrożenia. Systemy operacyjne realizują różne usługi komunikacyjne i obsługują różne protokoły bez interwencji użytkownika. Odbierane w ich wyniku dane mogą zawierać niepożądaną lub szkodliwą zawartość, mogącą narazić użytkownika na różnego rodzaju straty, od zmniejszenia efektywności połączeń komunikacyjnych po utratę informacji przechowywanych w systemie.

Możliwe rodzaje zagrożeń sięgają od niepożądanej poczty elektronicznej (spamu), wormów (robali internetowych), wirusów, programów szpiegujących i gromadzących dane o użytkowniku po ataki skryptowe i próby przejęcia kontroli nad systemem w celu wykorzystania jego mocy obliczeniowej lub jako stacji do prowadzenia ataków typu DDoS.

Większość tego typu zagrożeń jest niewidoczna dla użytkownika. W celu wykrycia ich obecności niezbędne jest zastosowanie specjalizowanych narzędzi, np. systemów Intrusion Detection Systems - IDS, monitorujących przepływ pakietów w sieci, klasyfikujących zgromadzoną informację i rejestrujących wystąpienia znanego typu zagrożeń lub podejrzanych sytuacji, czy też narzędzi analizujących zawartość pobieranych stron html w poszukiwaniu programów rejestrujących różne informacje o użytkownikach odwiedzających daną stronę.

O ile fakt istnienia większości tego typu zagrożeń jest nagłaśniany w wielu publikacjach i można przyjąć, że staje się znany przeciętnemu użytkownikowi, o tyle informacje o rozmiarach związanych z zagrożeniami strat (zmniejszenie przepustowości łącza, zajętość czasu i zasobów systemu, utrata prywatności) nie są już tak powszechnie dostępne. Celem niniejszego artykułu jest prezentacja informacji o tej, często ignorowanej, stronie zagrożeń internetowych.

2. Podział zagrożeń internetowych

Występujące zagrożenia można podzielić na zagrożenia aktywne, pojawiające się z sieci bez naszej ingerencji, np. w postaci robali internetowych (wormów), skanujących otoczenie sieciowe w poszukiwaniu systemów podatnych na atak, ataków skryptowych próbujących zainstalować w systemie „konie trojańskie", czy też programów skanujących sieć w poszukiwaniu systemów z określonymi „dziurami bezpieczeństwa" w oprogramowaniu systemowym, oraz na zagrożenia pasywne (bierne), wymagające do ich uaktywnienia ingerencji ze strony użytkownika.

Przykładami takich pasywnych zagrożeń są programy wykorzystujące do swojego rozprzestrzeniania „legalne" protokoły, takie jak *smtp* czy *http* – wirusy i robale internetowe rozprzestrzeniające się jako różnego rodzaju załączniki do przesyłek e-mail-owych, jako dodatkowa zawartość elementów graficznych stron czy też w postaci linków śledzących zachowanie (zwyczaje) użytkowników – web bugs (robale webowe) [3], wskazujących na serwery internetowe inne niż odwiedzany.

Otwarcie strony z web-bugs lub kliknięcie na niej link-u uruchamia nie zamierzoną wymianę informacji z tym dodatkowym serwerem, np. w postaci wyskakujących okien reklamowych, przesyłu informacji o systemie uzytkownika czy też wymiany zawartości „ciasteczek"- cookies, zapisywanych lokalnie w systemie użytkownika przez odwiedzany serwer,

zawierających różnego rodzaju informację o użytkowniku, co w efekcie może mieć następstwo w postaci serii „nie zamawianych" przesyłek e-mail-owych z różnego rodzaju ofertami reklamowymi (spamu).

Ochrona przed niechcianą pocztą (filtracja spamu) może również powodować straty zasobów użytkownika. Filtry antyspamowe konfigurowane są zwykle w programie klienta poczty elektronicznej, który dopiero po ściągnięciu poczty z serwera dokonuje jej selekcji. Lepszym rozwiązaniem jest filtracja poczty przez program na serwerze, przed przekazaniem poczty użytkownikowi. Dostępne mechanizmy filtracji nie są zbyt skuteczne, część spamu dociera do użytkownika. Wiekszość reklam przesyłana jest w formacie html, w którym wiele elementów graficznych zastąpionych jest linkami do serwerów. W momencie otwierania (podglądu) takiego dokumentu uruchamiane jest ściąganie grafiki z sieci, zwiększając ilość zasobów marnowanych przez spam.

Zetknięcie się z każdym z zagrożeń internetowych pociąga za sobą pewne koszty w postaci strat przepustowości łącza, zajętych przez transmisje pakietów nie inicjowane celowo przez użytkownika, zajęcia części czasu i zasobów procesora. Ponieważ zagrożenia typu biernego wykorzystują akceptowane przez system mechanizmy komunikacyjne i legalne programy, nie są one blokowane w żaden sposób przez dostępne mechanizmy zabezpieczające typu personal firewall czy program antywirusowy.

3. Wykrywanie biernych zagrożeń

Projektanci stron www stosują technologie zbierające jak najwięcej informacji o odwiedzających je użytkownikach oraz tworzące dynamicznie zawartość stron związaną z ich zainteresowaniami. Do zbierania informacji o użytkownikach wykorzystywane są *cookies* (ciasteczka), zapisujące na komputerze klienta informacje o nim i jego preferencjach, wykorzystywaną przy kolejnych wizytach użytkownika na danej stronie. Do śledzenia sposobu poruszania się po stronach serwera stosowana jest technologia *clear gif*, niewidocznych małych obrazków (1x1 pixel), których odczyt powoduje zarejestrowanie tego faktu na serwerze, np. przez uruchomienie skryptu. Normalnie, obie technologie powodują wymianę informacji między stacją użytkownika a odwiedzanym serwerem internetowym lub innym, w sieci tej samej firmy.

Analizując dostepne informacje o web-bugs (robalach webowych) można określić je jako ukryte (nie rzucające się w oczy) elementy stron internetowych, uruchamiające przy ich odczycie dodatkowe połączenia z serwerami z innych domen niż odwiedzana strona. Typowymi elementami stron, umożliwiajacymi podejmowanie tego typu operacji są małe obiekty graficzne, np. o rozmiarze 1 na 1 pixel, w kolorze otaczającego tła, z którymi

związany jest url wskazujący inną niż odwiedzana domenę i przekazujący w wywołaniu dużą ilość parametrów, mogących zawierać w sobie również inne url-e.

Ponieważ normalne elementy graficzne umieszczane są na stronach również w postaci wskazania (url-a) na dokument w odpowiednim formacie na serwerze internetowym, podstawą wykrycia podejrzanego elementu może być jego mały rozmiar, długość związanego z nim linku oraz jego przynależność do innej domeny internetowej. Ocena stopnia ewentualnej szkodliwości, tzn. tego, jakie informacje i w którą stronę przesyłane są w ramach otwieranego w ten sposób połączenia, wymaga dokładnej analizy zawartości przesyłanych ramek. Obecnie nieliczne programy, np. [4] sygnalizują tylko występowanie tego typu elementów na stronach internetowych, natomiast nie prowadzą analizy ich skutków ani nie oferują możliwości blokady otwierania sygnalizowanych połączeń.

4. Badania stron internetowych

W celu zebrania więcej informacji o występowaniu robali webowych zdecydowano się na przeszukanie wybranych stron internetowych oraz na zarejestrowanie i przeanalizowanie ruchu pakietów związanego z przeglądaniem tych stron. Do wstępnej selekcji stron www zawierających poszukiwane elementy wykorzystano program (wtyczkę do przeglądarki Internet Explorer) bugnosis. Wybrane strony otwierano ponownie, rejestrując ruch pakietów analizatorem protokołów.

Zarejestrowano całość ruchu pakietów przy odczycie wybranych stron kilku firm (Intel, Motorola, Microsoft, Oracle), czasopism (EDN, EPN, PC World, Infoworld, NetWorld) oraz portali internetowych (Buy.com, BestBuy, Compusa, Allegro.pl, msn.com) i analizowano, z iloma różnymi adresami (podsieci) otwierane były połączenia oraz jaki był udział tych dodatkowych podsieci w całym ruchu. Wyniki tej analizy przedstawia tabela 1.

Tabela 1

Ilość otwieranych połączeń oraz udział dodatkowego ruchu (web bugs) przy otwieraniu wybranych stron internetowych

Strona www	Ilość połączeń	Udział głównej strony	Udział web-bug
Intel.com	6	89,42%	9,31%
Motorola.com	3	99,19% (30,56%)	0,00% (68,64%)
Oracle.com	6	81,60%	15,59%
Microsoft.com	2	99,00%	0,00%
msn.com	8	91,89%	5,43%
BestBuy.com	7	94,12%	4,36%
Buy.com	7	96,59% (29,46%)	2,03% (69,16%)
Compusa.com	6	95,81%	3,44%
Allegro.pl	5	89,61%	9,62%

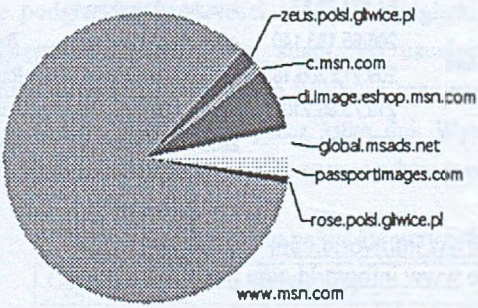
cd. tabeli 1

EDN_mag.com	7	80,64%	17,93%
EPN_mag.com	6	95,94% (89,19%)	3,54% (10,29%)
InfoWorld.com	6	66,03%	33,04%
NetWorld.com	8	89,75%	9,07%
TechRepublic.com	6	98,37% (68,51%)	0,00% (29,86%)
PCWorld.pl	6	81,31%	18,04%

Przy analizie ruchu pakietów dla niektórych stron (Motorola, TechRepublic, Buy.com, EPN-mag.com) wystąpił problem z zakwalifikowaniem części ruchu sieciowego związanego z podsiecią *akamai.net*. Odczyt stron tych firm odbywa się w dużej części z tej podsieci (od 10 do prawie 70%), co odpowiada jednej z cech szukanych web bugs (połączenia do serwerów z innej niż odczytywana strona domeny), z drugiej strony właściciel tej podsieci, firma Akamai Technologies oferuje usługi i technologie związane z e-biznesem, realizowane na potrzeby klientów przez udostępniane serwery, więc być może należy ten ruch traktować jako pochodzący z podstawowego serwera, stąd w tablicy zamieszczono obie interpretacje (jako web bugs w nawiasie).

Jak widać, wiekszość stron (z wyjątkiem Microsoft, może Motoroli i TechRepublic) wykorzystuje różne formy rejestrowania ruchu odwiedzających i gromadzenia o nich informacji z udziałem serwerów innych firm, spoza ich firmowej podsieci. Ten dodatkowy ruch sięga od kilku do ponad 30% całkowitego ruchu (w bajtach) związanego z odczytem danej strony. Udział głównej podsieci nie osiąga 100% ze względu na rejestrowany ruch związany z tłumaczeniem nazw dns na adresy IP.

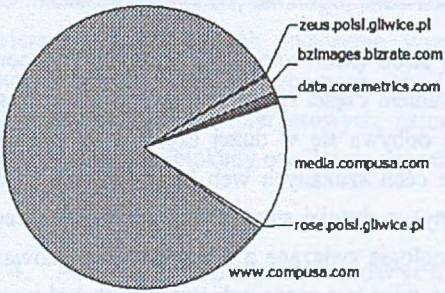
Na rys. 1 do 4 przedstawiono dla wybranych adresów (url) udział różnych serwerów w dostarczaniu zawartości odczytywanej strony.



207.68.172.245	www.msn.com	
207.68.173.244	www.msn.com	
207.68.177.126	c.msn.com	
66.147.182.37	di.image.eshop.msn.com	
65.54.225.156	passportimages.com	(?)
217.32.247.3	global.msads.net	(?)
157.158.57.35	rose.polsl.gliwice.pl	(dns)
157.158.1.4	zeus.polsl.gliwice.pl	(dns)

Rys. 1. Składowe ruchu związanego z odczytem strony www.msn.com
Fig. 1. Traffic components related to the reading of www.msn.com site

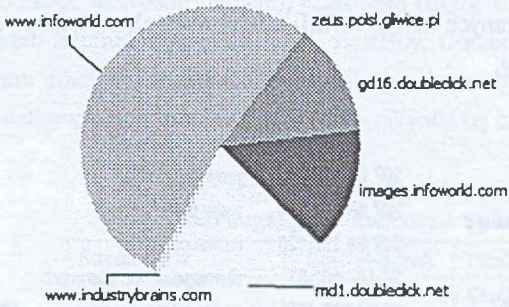
W przypadku strony www.msn.com przy tworzeniu zawartości strony uczestniczy w połączeniach sześć adresów, z których cztery leżą w domenie msn.com, chociaż w różnych podsieciach IP. Dwa pozostałe (5% ruchu) związane są prawdopodobnie z usługą passport i ogłoszeniami reklamowymi microsoftu, mogą też być związane z rejestracją ruchu użytkowników na stronie, ale cała zawartość strony pochodzi z różnych sieci tej samej firmy.



198.22.121.140	www.compusa.com	
198.22.121.142	media.compusa.com	
216.52.244.252	bzimages.bizrate.com	(web bug?)
64.245.50.115	data.coremetrics.com	(web bug?)
157.158.57.35	rose.polsl.gliwice.pl	(dns)
157.158.1.4	zeus.polsl.gliwice.pl	(dns)

Rys. 2. Analiza ruchu związanego z odczytem strony www.compusa.com
Fig. 2. Traffic analysis of the download of the www.compusa.com site

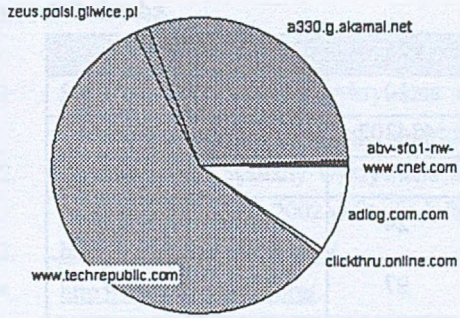
Strona www.compusa.com w zdecydowanej większości (96%) pochodzi z jednej podsieci IP i z jednej domeny. Około 3% ruchu stanowi komunikacja z serwerami z zupełnie innych domen, które można potraktować jako przykłady web bug-ów, resztę stanowi komunikacja z serwerami dns.



64.124.72.32	www.infoworld.com	
64.124.72.36	images.infoworld.com	
206.65.183.140	gd16.doubleclick.net	(bug ?)
209.213.205.163	www.industrybrains.com	(bug ?)
216.73.86.210	md1.doubleclick.net	(bug ?)
157.158.1.4	zeus.polsl.gliwice.pl	(dns)

Rys. 3. Składowe strony www.infoworld.com
Fig. 3. Components of the www.infoworld.com page

W przypadku site-u www.infoworld.com tylko 67% ruchu pochodzi z odwiedzanej domeny, natomiast 31% informacji przy ściąganiu tej strony jest wymieniane z serwerami z domeny doubleclick.net, związanej z usługami reklamowymi i gromadzeniem informacji o odwiedzających strony internetowe - w tablicy 1 zaliczono ten ruch jako web bugs.



206.16.0.191	www.techrepublic.com	
206.16.0.45	adlog.com.com	(?)
206.16.0.147	clickthru.online.com	(?)
206.16.0.150	abv-sfo1-nw-www.cnet.com	(?)
193.219.28.104	A30.g.akamai.net	(web bug ?)
157.158.1.4	zeus.polsl.gliwice.pl	(dns)

Rys. 4. Analiza ruchu związanego z odczytem strony www.techrepublic.com
Fig. 4. Traffic analysis of the www.techrepublic.com page

Strona www.techrepublic.com jest bardziej kłopotliwa do oceny. Większość ruchu pochodzi z jednej podsieci klasy C, ale reprezentuje cztery różne domeny internetowe. Może to odpowiadać hostowaniu serwerów różnych domen na sprzęcie dostawcy usług internetowych. Według adresów IP 30%, według nazw domen dns do 41% ruchu ramek nie jest związane z domeną techrepublic.com .

5. Ocena strat zasobów związanych ze spamem

Dla oceny wielkości zasobów zajmowanych przez spam docierajacy do programu klienta poczty elektronicznej rejestrowano analizatorem protokołów ruch ramek w czasie odbioru poczty z serwera pocztowego, a nastepnie rejestrowano ruch pakietów w trakcie przeglądania zestawienia odebranej poczty, bez jej otwierania w osobnym oknie i usuwania spamu. W programie outlook express, w momencie ustawienia kursora na wybranej przesyłce, pojawiał się podgląd jej zawartosci w oknie podglądu. Po wskazaniu przesyłki kursorem była ona następnie kasowana jako spam. Po usunięciu grupy wybranych listów rejestracja była zatrzymywana i analizowano wielkość oraz rodzaj ruchu w sieci w trakcie jej wykonywania. Operacje te powtarzano przez kilka dni. Wyniki pomiarów natężenia ruchu przy odbiorze kilku porcji poczty i usuwaniu spamu zebrano w tabeli 2.

Tabela 2

Analiza ruchu pakietów przy odbiorze poczty i usuwaniu spamu

Objętość ściąganej poczty [B]	847823	999612	491775
Ilość ściąganych przesyłek	43	95	28
Ilość przesłanych pakietów	1123	1312	634
Udział protokołu smtp [%]	90	100	99
Bajtów na przesyłkę e-mail	19716	10522	17563

cd. tabeli 2

Pakietów na przesyłkę e-mail	26	14	23
Ilość skasowanych przesyłek (spamu)	10	40	10
Wielkość ruchu związanego ze spamem	434302	1430892	362513
Ilość przesyłanych pakietów	1044	2927	810
Ilość otwartych połączeń w trakcie przeglądania spamu	24	49	22
Udział protokołu http w przeglądaniu spamu [%]	97	98	96
Bajtów na usuwaną przesyłkę (spam)	43430	29201	36251
Pakietów na usuwaną przesyłkę	104	60	81
Koszt spamu [bajtów]	63146	39723	53814
Koszt spamu [pakietów]	130	74	104

W końcowych wierszach tabeli 2 przedstawiono orientacyjny „koszt" jednej przesyłki e-mailowej, usuwanej w programie klienta poczty elektronicznej. Składają się na niego „koszt" ściągnięcia przesyłki z serwera poczty (średnio 10-20 kB) oraz „koszt" dodatkowego ruchu w sieci, związany z samoczynnym otwieraniem przesyłki w oknie podglądu i ściąganiem elementów listów w formacie html z serwerów internetowych. Ta druga składowa w badanych przykładach jest 2-3 razy większa (większość spamu reklamowego zawiera wiele elementów graficznych). Sumaryczny ruch w sieci związany z odbiorem i usuwaniem w kliencie poczty elektronicznej pojedynczej przesyłki - spamu wyniósł od 40 do 63 kB, co odpowiadało wymianie od 74 do 130 pakietów w sieci.

6. Podsumowanie

Z przedstawionych przypadków analizy ruchu w sieci związanego z przeglądaniem stron internetowych i odbiorem poczty elektronicznej zawierającej spam wynika, że w wielu wypadkach ponosimy nieplanowane koszty swojej działalności. Przy zwiedzaniu stron internetowych od kilku do ponad 30 % przesyłanej przy tym informacji może nie pochodzić z odwiedzanej strony i może związane być z gromadzeniem informacji o działalności użytkownika i jego zainteresowaniach. Usuwanie spamu w programie klienta poczty elektronicznej, np. outlook lub outlook express może pociągać za sobą przesył 40 - 60 kB (wymianę ponad 100 pakietów). Na szczęście większość dostawców usług internetowych nie pobiera opłat od objętości przesyłanej siecią informacji, ale w innych przypadkach bierne zagrożenia internetowe mogą mieć zupełnie wymierny, zauważalny dla użytkownika koszt.

LITERATURA

1. Skrzewski M.: Systemy wykrywania zagrożeń komputerów BW 486/Rau2/2002. Sprawozdanie z tematu badawczego, Instytut Informatyki Politechniki Śląskiej, grudzień 2002.
2. Skrzewski M.: Systemy wykrywania zagrożeń komputerów. IX Konferencja Sieci Komputerowe, Zakopane 2002 w Studia Informatica, czerwiec 2002.
3. http://www.pestpatrol.com/.
4. http://www.bugnosis.org/.
5. Koziński M.: Ochrona danych - Groźne oblicza spyware, http://www.pckurier.pl/ archiwum/autor.asp?IDAutor=330&Rok=2002.
6. Fratto M.: New Security Threats - Stronger Defenses, http://www.informationweek.com /story/1211/1211f2.html.
7. http://www.msn.com/.
8. http://www.techrepublic.com/.
9. http://www.infoworld.com/.
10. http://www.compusa.com/.

Recenzent: Dr inż. Andrzej Białas

Wpłynęło do Redakcji 16 kwietnia 2003 r.

Abstract

Computer systems due to their Internet connections are subject to many kinds of threats, invisible for user during normal operation. One can distinguish active threats, well known, coming from the network without user involvement, and passive threats, which must be activated by user intervention. To find their presence one need special tools like Intrusion Detection Systems - IDS or protocol analyzers (sniffers). Two kinds of passive threats, web bugs and spam related losses, were investigated on selected examples of web sites and daily mail processing routines. Results of experiments, the amount and kind of generated network traffic, obtained with the use of protocol analyzer, are summarized in table 1 and 2. Details of the web traffic are presented on figs. 1 through 4 and discussed through the paper.

Adres

Mirosław SKRZEWSKI: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16, 44-101 Gliwice, Polska, mireks@polsl.gliwice.pl .