Anonymity on the Web

Threats for the anonymity

- **IP address**
  - Public IP address – assigned by Internet Services Provider (ISP) – they can identify the user exactly; others may identify approximate location by the IP (*e.g. try GeoIP in the browser*); if the user is also logged in to a service like Gmail or Facebook, their providers can identify both the user and his/her location
- **Cookies**
  - A cookie is a small portion of data created by a website and stored in a web browser - it keeps various pieces of information about the user's activities in the past, which can be any time uploaded back to website – purpose: identification of the user, his/her browsing history on the website, past searches on the website etc. Some cookies may store sensitive information – passwords, information from forms, credit card number etc.
  - Cookies can be also stolen by hackers and used for unauthorized access to user's data and services
- **Smartphones and tablets**
  - Android and iOS require (as a part of the initialization) the logon to their stores – the sign up involves sensitive information
  - Majority of smartphones is also able to get precise position of the user (GPS) and the telephone number
- Services like face or voice recognition (Facebook, Google Street View)
- DNS – the provider of DNS knows what websites were visited from given IP address
- Malware

Tools for improving anonymity

- Login names different from the real ones (☺)
- Proxy servers
- Anonymizing networks – TOR, I2P
- Virtual Private Networks (VPN)
- Public computers
- Public networks
- Encryption of the whole communication

Why support anonymity?

- Access to uncensored information in countries, which deny freedom of speech
- Whistleblowing – publishing of sensitive information about government, corporations, which may endanger the publisher (Wikileaks)
- Attempt to minimize digital trace (information collected by providers of online services)
- In anonymity people can discuss sensitive topics (health, physical abuse, sexuality)

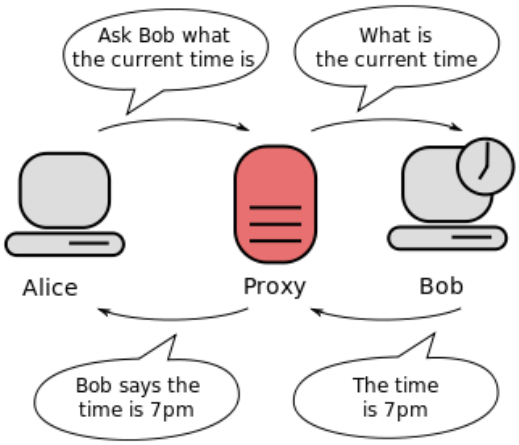
Why not?

- Detection and identification of those who commit crimes over the network
  - Spam, Hacking, Terrorism
- Cyber bullying, trolling

Proxy Servers

A proxy server is a server (computer/application) that acts as an *intermediary* for requests from clients seeking resources from other servers.

1. A client connects to the proxy server, requesting some service, such as a file, connection, web page, or other resource available from a different server.
2. Proxy server identifies the server (if necessary –e.g. using DNS), and sends request for the service.
3. The connected server sends response.
4. The response is re-transferred to the client.



Effects:

- The server, which provided service, does not know the client's IP address –the IP address it knows is the proxy IP only (used when getting services allowed in certain countries only – e.g. to get to Hulu the proxy must be in the USA).
- DNS is contacted by the proxy, so the DNS server does not know IP address of the client (that is regularly used to monitor user's habits or restrict access to some websites).
- Through proxy passes all traffic – if the traffic is not encrypted, the proxy may control or modify anything (often used in companies to monitor users and restrict access).
- Browser addons (Flash, Java) can be used to reveal client's IP address.

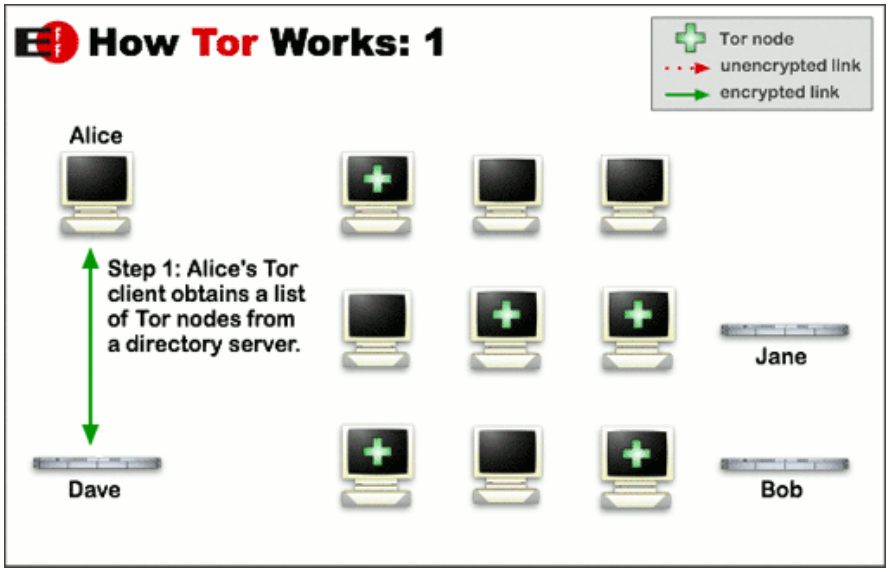
Proxy servers available to public offer their services

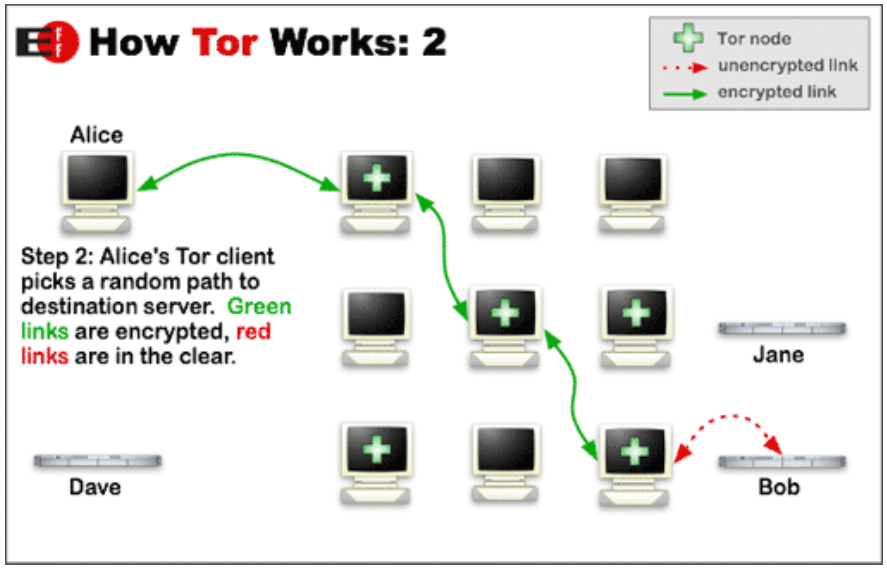
- for free – usually slow, a lot of advertisement;
- paid – faster a stable connection, yet not as fast as direct connections.

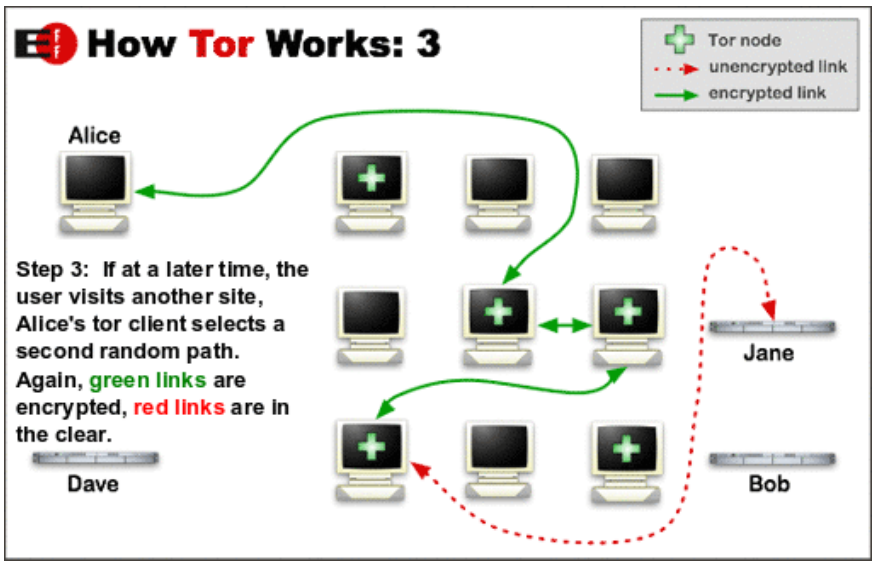
Tor and JonDonym

Anonymity service – Tor = independent network and software utilizing the network – Tor client software directs internet traffic through a worldwide volunteer network of servers – the result is the user's location or usage cannot be found from anyone conducting network surveillance or traffic analysis.

It was originally developed with the U.S. Navy in mind, for the primary purpose of protecting - government communications.







Tor clients:

- *Tor browser* – modified Firefox browser
- *Vuze* (torrent client) – Tor support
- *Orbot* – application for Android

Alternative - **JonDonym.**

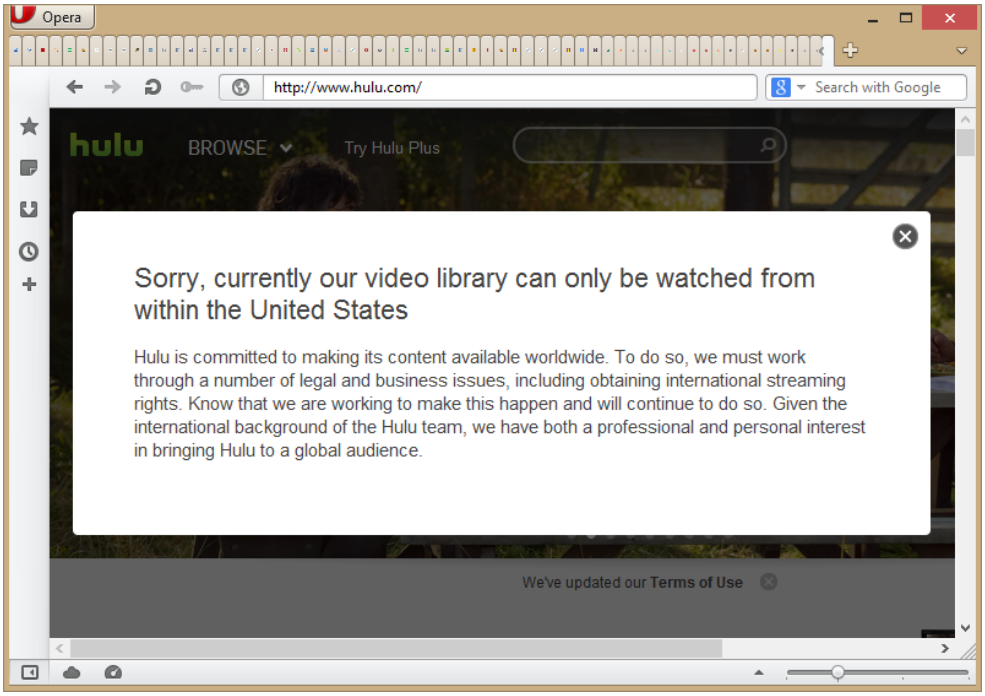


Figure 1: HULU from Slovakia in standard browser

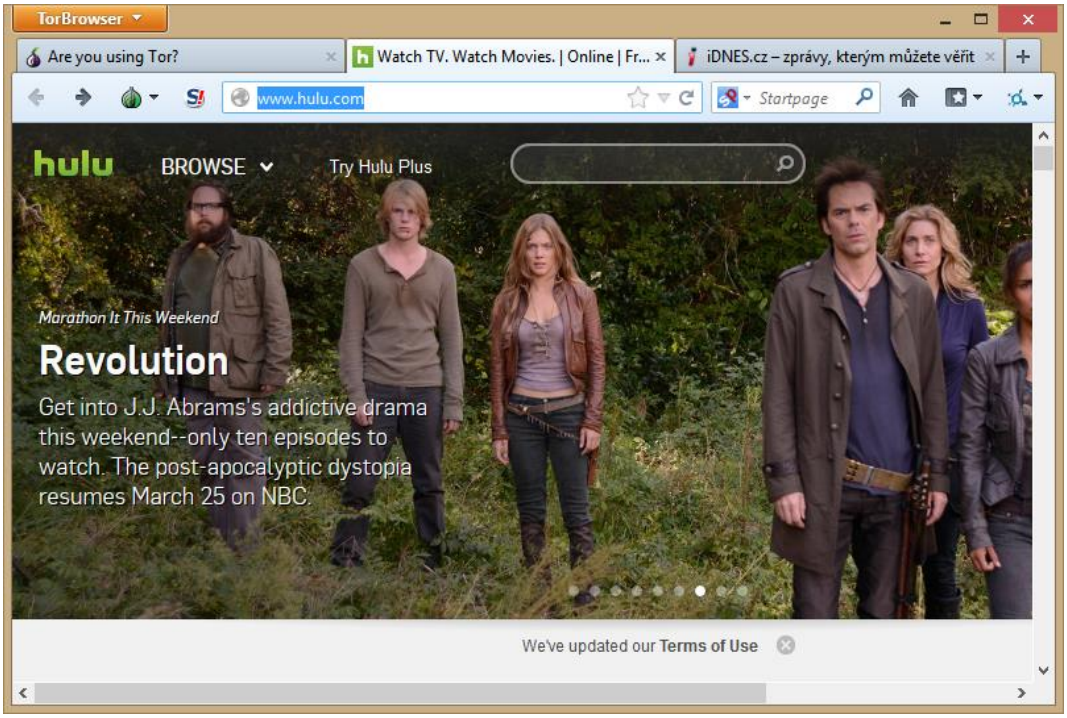


Figure 2: HULU from Slovakia using Tor

Darknet

Web is currently split into 3 groups:

- **Surface web** – normally accessible and searchable over Google or similar search engines
- **Deep web** – content, which is not publicly accessible, but it can be accessed over standard software (e.g. ordinary browsers) – access is restricted by authentication; here belong your private messages, email, databases …

- **Darknet** - a small portion of the deep web that is intentionally hidden and made inaccessible without specific software (the Tor network, only accessible via Tor browser).

Within the dark net, both web surfers and website publishers are entirely anonymous. Whilst large government agencies are theoretically able to track some people within this anonymous space, it is very difficult, requires a huge amount of resources, and isn't always successful.

Military, government, and law enforcement organisations are still amongst the main users of the hidden Internet. This is because ordinary internet browsing can reveal your location, and even if the content of your communications is well-encrypted, people can still easily see who is talking to whom and potentially where they are located. For soldiers and agents in the field, politicians conducting secret negotiations, and in many other circumstances, this presents an unacceptable security risk.

The darknet is also popular amongst journalists and political bloggers, especially those living in countries where censorship and political imprisonment are commonplace. Online anonymity allows these people, as well as whistleblowers and information-leakers, to communicate with sources and publish information freely without fear of retribution. The same anonymity can also be used by news readers to access information on the surface web which is normally blocked by national firewalls, such as the 'great firewall of China' which restricts which websites Chinese Internet users are able to visit.

Activists and revolutionaries also use the darknet so that they can organise themselves without fear of giving away their position to the governments they oppose.

Of course, this means that terrorists also use it for the same reasons, and so do the darknet's most publicized users—criminals. The darknet's large criminal marketplaces are well known. Here, you can buy everything from drugs to assassinations.

Virtual Private Networks

A Virtual Private Network (VPN) is a network technology that creates a secure network connection over a public network such as the Internet or a private network owned by a service provider.

VPN technology employs sophisticated encryption to ensure security and prevent any unintentional interception of data between private sites. All traffic over a VPN is encrypted using algorithms to secure data integrity and privacy. That is its key advantages in comparison to the proxies – **encrypted is ALL traffic** (unlike in case of proxies, where ISP may capture and see the traffic between you and the proxy).

There are services (usually paid), which offer VPN for almost anyone. Examples are:

- *HideMyAss*
- *IPVanish*
- *ExpressVPN*

In this case is installed an application, which – when started – connects ALL applications over the VPN.

There are also free solutions:

- **Opera offers built-in VPN** – this one works within Opera only(!)
- **Hamachi**
- Skilled users may employ their own VPN (e.g. using OpenVPN).

Cookies

Cookies are strings of text that are saved on a computer by browsing different web pages. They allow small bits of information to be stored.

Cookies are used to save personal preferences among other things. They are also used to track demographics and browsing habits. This information is sent to the user's computer and then uploaded to web databases without the user's approval.

Some cookies hold the session information, so the users do not have to login again. It contains session ID (a huge number), which is unique for a session, i.e. a specific connection of a user and a server. Stealing the session cookie has similar consequences to a stolen password – it is possible to open the account without the password, while the session ID is valid, and extract information or take actions, which do no need additional authentication (password, SMS).

Social Engineering

Social engineering is essentially the art of **gaining access** to buildings, systems or data **by exploiting human psychology**, rather than by breaking in or using technical hacking techniques (*e.g. instead of trying to find a software vulnerability, a social engineer might call an employee and pose as an IT support person, trying to trick the employee into divulging his password*).

Techniques

Phishing - It means to make a target go on an artificial emulation of the genuine site, where the target puts his username and password. A common method is to send a **phishing email** that appears to come from a legitimate business—a bank, or credit card company—requesting "verification" of information and warning of some dire consequence if it is not provided. The e-mail usually contains a link to a fraudulent web page that **seems** legitimate—with company logos and content—and has a form requesting everything from a home address to an ATM card's PIN.

Figure 3: Example of phishing

Dumpster Diving – search of valuable information about the target by searching inside their trash, either physical or virtual (old harddisks, flashdisks, CDs, printed materials …).

Spoofing IP - similar to proxies or VPN, it allows to use a random IP address.

Pharming – a technique, when user is redirected to fake version of a website while typing **legitimate url.** The scam is possible by malevolent modification of the DNS, which can be done:

- **manually** – user's DNS settings are modified on the attacked computer (rare, inefficient);
- **spoofed DNS in DHCP** – public places, malevolently opened WiFi networks;
- **modified by malware** – usually done by Trojan horses – efficient until the infection is detected by antimalware software;
- **DNS poisoning** – manipulation of IP addresses returned by a regular DNS server – misuses a security hole or a flaw in software of the DNS server (very efficient – it can affect all users of the attacked DNS server).

Process of Pharming[1]

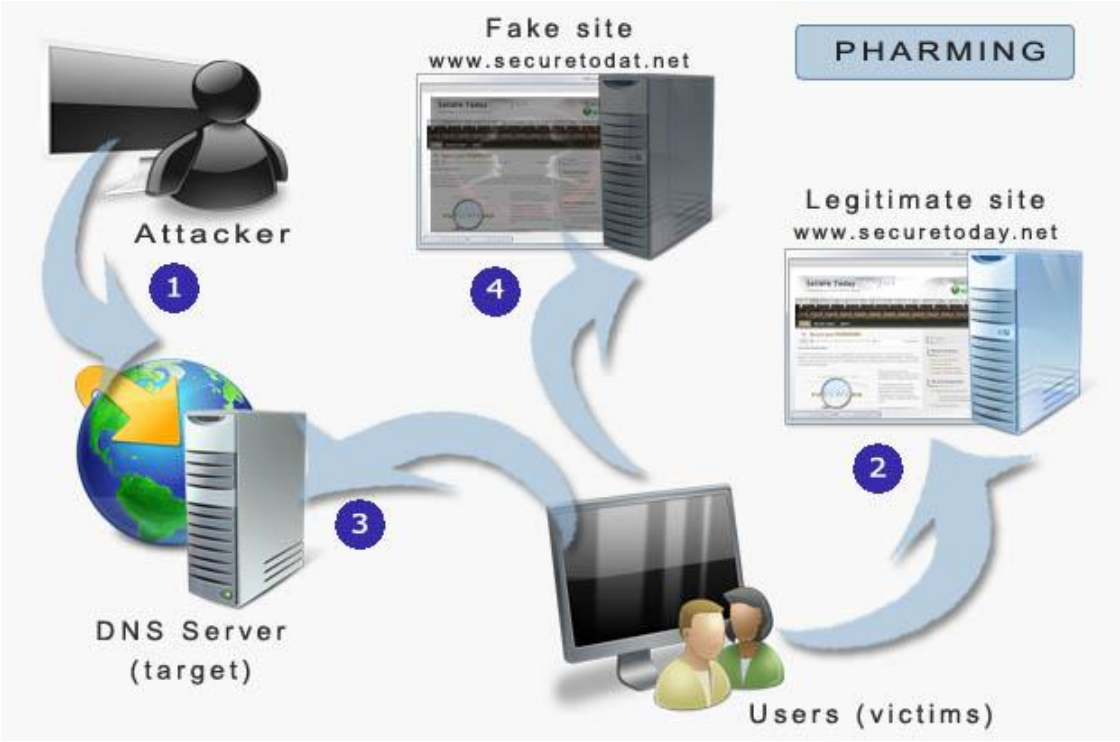


Figure 4: Pharming (source: http://www.securetoday.net/tag/pharming/)

1. *An attacker exploits vulnerabilities of a DNS. Using crafted responses or take advantage of a vulnerability, an attacker can poisoned the DNS cache and can change valid entries. Internally, a disgruntled engineer can even manipulate the host lookup on these servers. Externally, attackers can take advantage of the operating systems vulnerabilities.*
2. *A user wants to go to a website securetoday.net and enter in the browser.*
3. *The user's computer queries the DNS to resolve the site. Now, DNS being poisoned resolved the site to the nefarious fake website and redirected to securetodat.net.*
4. *User unaware of what happened thinks he is on the correct website.*

[1] http://www.securetoday.net/tag/pharming/

Prevention and Detection

Phishing

- Do not respond to emails, which ask for personal information
- Check, whether the page URL you were sent to from an email, is correct – even one different symbol is suspicious.
- Typical indicators of phishing: general title – no specific user information; urgency – "*something bad happened*"; "*Click here*" links

Pharming

- Set well-known DNS servers (Google: 8.8.8.8, 8.8.4.4; OpenDNS: 208.67.222.222; 208.67.220.220);
- Check, if the web browser shows indicators of secure and trusted connection (green bar, padlock in the address bar);
- Check the certificate displayed by the page;
- Do not use HTTP web pages to send sensitive information.
- Keep the firewall on, update the system and antimalware software (a reliable one).

Reliability and Trustworthiness of Information

- Information can be published virtually by anybody – often anonymously – the reliability of information is in doubt
- Signs of reliable information
    - Known author's identity – real name, photo along with brief information about the author in online resources (catalogues, educational institutions, LinkedIn)
    - Assigned verifiable identifier (ISBN number, ISSN number)
- Thinks to keep in mind
    - *Who is the sponsor of the website/webpage*? The information may miss some details or both pros and cons of the topic (information about nicotine at the website of Philip Morris)
    - *Is there contact information*? If the information is provided by trustworthy organization, there should be published full contact information (address, email, phone); if there is just name of the organization, but no contact, it is no reliable
    - *Is it primary or secondary information*? Primary – topic of an article, secondary – information in the discussion related to the article

Hoax, Spam, Junk Mail

Spam

- unsoliticed electronic messages (emails, instant messages, SMS) – usually advertisement, but it can be also simple scam (Nigerian scam)
- problems
    - goods advertised in spam bring usually little to no benefit in exchange for money
    - spam overloads communication channels
    - users/spam filters spend a lot of time sorting the messages

Hoax

- information that is not true, however it is masqueraded as truth – it often causes undesirable panic (a virus that blows monitors, various companies offering money for emails send to multiple friends, recycled milk etc.)
- Easy verification on pages, which collect known hoaxes – e.g. **hoax.cz**

Junk Mail

- Another name for email spam, i.e. spam send via emails

Some Extra Reading

Source: **http://technet.idnes.cz/spam-pravopisne-gramaticke-chyby-dqt-/sw_internet.aspx?c=A121126_164036_sw_internet_pka**

Proč je nevyžádaná pošta plná chyb? Aby na ni reagovali jen troubové

27. listopadu 2012

Příšerné pravopisné a gramatické chyby, nesmyslné požadavky nebo nereálné sliby, to jsou téměř nepřehlédnutelné známky toho, že se nás nevyžádaným e-mailem někdo snaží podvést. Proč je spam tak hloupý? Aby na něj reagovali jenom hlupáci, u kterých je reálná šance, že pošlou peníze.

"Pozdravy den nádherný, nigerijský jsem ministr a rád jsem bych požádal Vy o transakci pomoc." Je možné, aby i po přinejmenším dvaceti letech existence e-mailu někdo naletěl na tak očividný podvod?

Spam v číslech

Spam představuje až 89 procent elektronické pošty (zdroj: MAAWG)

Nejvíce spamu pochází z Indie (13,8 %), USA (10,5 %), Jižní Koreje (6,4 %) a Ruska (5,4 %)

Čísla ukazují, že rozesílání nevyžádané pošty, takzvaného spamu, je stále výnosné podnikání. Podle odhadů tvoří spam dvě třetiny všech poslaných e-mailů, což znamená pokles oproti situaci před dvěma lety. Automatizované filtry se stále vylepšují, a do našich schránek tak naštěstí pronikne jen zlomek této laviny. Jen samotný fakt, že spam stále chodí, svědčí o jeho ziskovosti.

Spamboti (počítačové programy posílající spam) svým tvůrcům vydělávají miliony. Spam je často pečlivě testovaný nejen po technické, ale i po psychologické stránce. Spameři vytvoří řadu verzí, které rozesílají na různé adresy a podle procenta odpovědí (míra odrazu, míra konverze apod.) texty dále zdokonalují.

Dvě veze jedné spamové zprávy. Stejné věty jsou formulovány různě "špatnou" angličtinou. Žlutě jsou označeny některé rozdílné formulace v obou e-mailech. Červeně jsou zakroužkované vážnější prohřešky proti anglické gramatice (drobné chyby nejsou vyznačeny).

Když jsou tedy zprávy tak pečlivě testované, jak je možné, že dokážete na první pohled rozpoznat, že jde o spam? Odpověď vás možná překvapí: spameři jsou rádi, že se nenachytáte. Ušetří si čas, který mohou věnovat naivnějším příjemcům, tedy těm, u kterých je větší šance, že z nich nějaké peníze dostanou.

Inteligentní čtenáři jsou pro spamera ztracený čas

Abychom pochopili, jakou cenu mají různí příjemci pro rozesílatele spamu, podívejme se, jak může fungovat na obrázku výše uvedená žádost Eleny z Ruska. Jde mimochodem o minimálně pět let starý trik, zřejmě úspěšný, protože se stále objevují aktivní varianty.

Spamer: rozešle žádost nebohé Eleny milionům příjemců. Elena je neprosí o peníze, ale o to, aby jí do Ruska poslali přenosná kamna. Adresu prý pošle.

Příjemce A: pozná na první pohled, že jde o podvod, zprávu smaže nebo označí za spam a jde dál.

Příjemce B: zprávu si přečte a je mu Eleny líto. Napíše jí, aby se jí zeptal, jak jí může pomoci. Spamer se zaraduje a začne z něj nenápadně tahat peníze. Příjemce B zpozorní, konverzaci ukončí, protože (správně) cítí zradu. Spamer na něj svůj čas plýtval zcela zbytečně. Při velkých počtech jde o ohromné ztráty.

Příjemce C: i on je dojat Eleniným příběhem. Vymění si s ní několik e-mailů. Zjistí, že platit poštovné za těžká kamna do Ruska je nesmysl. Nabídne, zda by místo toho mohl poslat peníze. Elena neví, musí se zeptat. Zato spamer ví: narazil na perfektního příjemce spamu. Dříve či později z něj peníze vytáhne.

Spamerům jde pochopitelně o to, aby svou zprávu dostali k co nejvíce příjemcům typu C, protože právě z nich (a pouze z nich) mají peníze.

Důležité však je podívat se na náklady spamerů. Rozeslat miliony zpráv je nestojí skoro nic. Ale odpovídání na reakce, to už je časově náročná, špatně automatizovatelná činnost. A to je důvod, proč se snaží spameři přesunout co nejvíce příjemců ze skupiny B do skupiny A. Béčka pro ně totiž představují zbytečné náklady a zbytečně strávený čas.Spam, spam, spam, spam!

Slovo "spam" označuje konzervovaný masný výrobek podobný šunce, jeden z mála masitých pokrmů dostupných v Británii během druhé světové války.

Konzerva - SPAM

Monty Python - Létající cirkus

Pro nevyžádanou poštu se tento pojem používá v narážce na scénku slavné britské skupiny Monty Python. "Nemám ráda spam!" deklaruje hrdinka (již zesnulý herec Graham Chapman) v reakci na nabídku restaurace, kde se spamu nemůže vyhnout, ať si objedná cokoli. Scénka měla premiéru v roce 1970 a pro nevyžádané zprávy se tento termín používá přinejmenším od 90. let.

Nigerie odradí skoro všechny. A o to spamerům jde

"Aby útočník maximalizoval zisk, nezaměřuje se na všechny možné příjemce zpráv. Musí vyvážit zisk ze skutečných zásahů a ztrátu z falešných zásahů," píše Cormac Herley, výzkumník společnosti Microsoft. Ve své studii ukazuje, že spameři mají dobrý důvod pro to, aby působili nedůvěryhodně.

Když si do Googlu napíšete slovo "nigeria", hned na čtvrtém místě vám vyhledávač nabídne "nigerian scam", tedy nigerijský podvod. Jde zřejmě o nejznámější typ podvodných e-mailů vůbec. "Proč spameři stále říkají, že jsou z Nigérie, když by si přece mohli vymyslet libovolnou zemi na světě? Není to tím, že jsou hloupí," myslí si Herley. Spameři se snaží napsat e-mail dostatečně nedůvěryhodný, aby na něj odpověděli jen ti nejnaivnější z příjemců. Právě těm má cenu věnovat se v dalších e-mailech a tahat z nich postupně velké peníze.

Úspěšnost spamových výzev se pohybuje hluboko pod jednou tisícinou, a pokud by spameři zněli příliš důvěryhodně, zvýšili by si náklady, aniž by jim stouply příjmy. Spameři pečlivým testováním ladí text tak, aby obsahoval dostatek indicií, že se jedná o spam. Odradí tak průměrně inteligentní příjemce, a zůstanou jen ti nejdůvěřivější, kteří se nepozastavili ani nad špatnou gramatikou, ani nad nelogickými konstrukcemi.

Kdo spamuje spamery?

Dalším důvodem pro špatnou gramatiku v nevyžádaných e-mailech je samozřejmě i strojový překlad či snaha obejít filtrování zpráv i pomocí automatizované či manuální úpravy textu.

Doplňující pohled a alternativní výklad špatné gramatiky nabízí odpověď na Quora: "Existuje skupina lidí, kteří se zaměřují na spamování spamerů. Reagují na jejich e-maily a snaží se je nenápadně přesvědčit a manipulovat." Stránky věnované podvádění podvodníků najdete třeba na adrese 419eater.com ("419" je zažité označení Nigerijského dopisu).

Vyznavači tohoto zvláštního, zřejmě zábavného a možná dokonce užitečného sportu, údajně dávají spamerům řadu neužitečných tipů. Snaží se je například přesvědčit k tomu, aby do svých mailů zařadili více chyb, o kterých jim tvrdí, že jde o "srozumitelný žargon, nikdo už to dneska jinak neřekne, věř mi, mluvím anglicky celý svůj život".

Vtipálci tak podvodníkům ubírají čas a dávají falešnou naději, že by z nich mohly kápnout nějaké peníze. Jak ovšem ukazuje výše citovaná studie, to, že je příští spam plný gramatických chyb, může být pro rozesílatele spíše výhoda.

Autor: Pavel Kasík, www.idnes.cz