Metody pro potlačení spamu

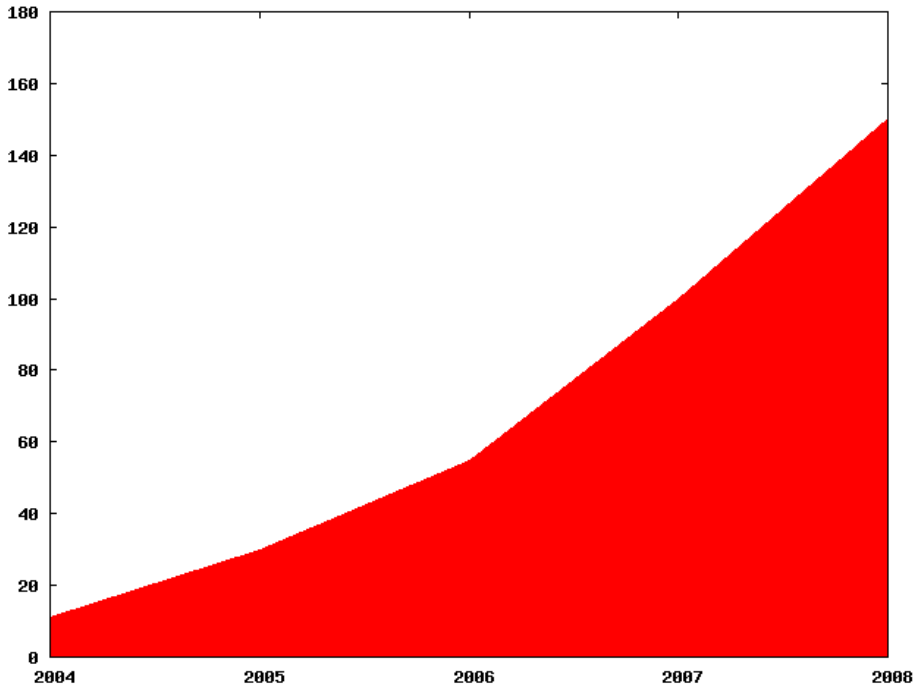
Petr Hruška (petr.hruska@nic.cz)

Obsah prezentace

- statistické údaje, zdroje spamu
- existující metody
- vyhlídky

Statistiky

- drastický nástup spamu v roce 2003
- v roce 2005 překročila míra spamu 80%, nyní dosahuje přibližně 85 až 90%



objem spamu v miliardách denně

Zdroje spamu

- 95% spamu je odesíláno ze zombie počítačů v botnetech
- podle spamhaus.org pochází 80% spamu od cca 400 lidí
- botnety se rozrůstají díky novým nezkušeným uživatelům
- odhaduje se, že v botnetech je nyní asi 12 milionů počítačů
- největší botnety mají miliony zombie

Rozesílání spamu

- zombie počítač je schopen rozeslat tisíce až desetitisíce spamů za hodinu
- často se využívá nastavení Outlooku
- velikost průměrného spamu je 2 až 3 KB
- obsahuje-li spam grafiku, je průměrná velikost okolo 11KB
- ve spamech bývá link (přes 60%)
- spamy mají často podvrženou zpětnou adresu, která skutečně existuje

Odpovídání na spam

- spam se zahazuje bez upozornění odesílatele
- odesílatel je totiž často zfalšován a byl by mu tak vlastně předán spam
- tím bohužel email přichází o spolehlivost
- adresa, ze které přijde odpověď, má pro spammery mnohem větší cenu

Obrana proti spamu

- blacklisty, graylisty, whitelisty
- platební systémy
- reputační systémy
- DCC (distributed checksum clearinghouses)
- DKIM
- elektronický podpis
- SPF
- proof of work
- filtry obsahu

Blacklisty

- založeny na spamtraps, nebo na seznamech dodaných ISP
- v první fázi se filtruje podle ip adresy, lze rovnou odmítnout spojení (75%)
- po přijetí mailu se hledají url spammerských stránek, zbyde asi 8% původního spamu
- blacklisty snižují výpočetní nároky na další filtrování
- mají velkou výhodu v tom, že si každý může zjistit, zda je filtrován

Graylisty

- založeny na předpokladu, že legitimní odesíletel zkusí poslat email znovu
- při prvním pokusu je email odmítnut pro dočasnou chybu
- při dalším pokusu (který nesmí nastat příliš brzy) je mail přijat a doručen
- některé systémy nezvládají další pokusy
- odesílateli může přijít zpět zpráva s upozorněním (still trying)

Whitelisty

- maily jsou přijaty jen od uživatelů vedených ve whitelistu
- žádost o přidání do whitelistu může být také spam
- první kontakt je potřeba provést nějakou jinou metodou
- pro většinu firem je první kontakt (se zákazníkem) velmi důležitý

Platební systémy

- technicky vzato není problém zavést obdobu známek
- problémem je nutnost jakési centrální autority, která by známky vydávala a evidovala
- decentralizace na více vydavatelů sníží spolehlivost
- evidence vlastníků či použitých známek?
- měla by být známka krytá penězi?
- kolik by známka měla stát? (rozdíly mezi bohatými a chudými státy)
- měly by být národní známky?
- vzniknou směnárny?

DCC

- distributed checksum clearinghouses
- sdílený registr hashů zpráv
- vyskytne-li se hash mnohokrát, je označen jako spam
- určitá část spamu má variabilní náhodně vygenerovanou část, právě kvůli DCC

DKIM

- domainkeys identified mail (RFC 4871)
- umožňuje podepsat obsah zprávy
- vkládá speciální hlavičku DKIM-Signature
- veřejný klíč je uložený jako TXT v doméně (např. gamma._domainkey.gmail.com)
- podpis typicky vkládá server, nikoliv klient
- podpis je citlivý na změny zprávy, například přidání informací o antivirové kontrole
- email s chybným DKIM nelze jednoznačně považovat za spam
- moc se nepoužívá, výjimky jsou např. gmail.com a linkedin.com

Elektronický podpis

- podpis uživatelem
- sám o sobě toho moc neřeší
- spammer si může vymyslet vlastní klíč
- opět problém prvního kontaktu

SPF - Sender Policy Framework

- umožňuje majiteli zveřejnit, kdo může odesílat emaily z jeho domény
- má status experimental, pracovní skupina se rozpadla
- v zahraničí 10 až 20% domén, v ČR jen asi 2.5%
- podpora v běžných mailserverech (exim, qmail, sendmail)

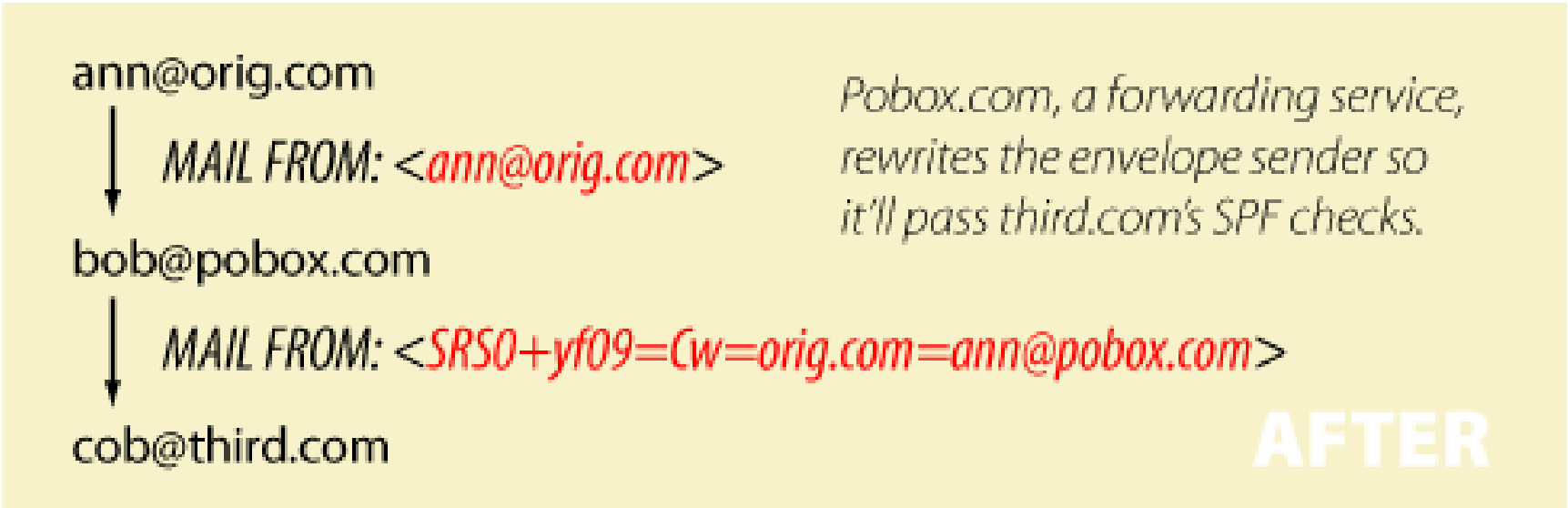
SPF - Sender Policy Framework

- technicky řešeno přes DNS, TXT záznamy začínající na "v=spf1", nebo přímo nový typ SPF
- příjemce kontroluje SPF záznam pro doménu uvedenou odesílatelem v *MAIL FROM*

SPF - Sender Policy Framework

- problémy s verbatim forwarding (forwardování SMTP serverem)

SRS - Sender rewrite scheme



Proof of work

- odesílatel musí vyřešit úkol, který by byl pro spammera příliš náročný
- CAPTCHA, například přepis textu z obrázku na webu příjemce
- výpočetně náročné metody – hashcash

hashcash

- klient musí doplnit řetězec vzniklý z emailové adresy příjemce a datumu odeslání tak, aby po aplikování funkce md5 vzniknul hash začínající určitým počtem nulových bitů na počátku
- obtížnost lze nastavit počtem bitů
- výpočetně náročné, je nutné vyzkoušet mnoho kombinací
- lze výrazně urychlit pomocí grafických karet, ale koncept lze zachovat s použitím jiných algoritmů než je md5

hashcash

- hashcash je podporován SpamAssasinem
- v současné době je možné pomocí hashcash výrazně snížit pravděpodobnost chybného odfiltrování mailu
- hashcash je opatření, které může provést uživatel i správce MTA
- existuje plugin do Thunderbirdu, do Outlooku nikoliv
- důsledné používání hashcash by pravděpodobně snížilo celkový objem spamu

Filtry obsahu

- nejznámější freewareový program SpamAssasin
- používá mnoho kritérií založených na hlavičkách i obsahu mailu
- je trénován tak, aby chybovost nebyla vyšší než 0.1%
- s využitím SpamAssasinu lze spam prakticky okamžitě zlikvidovat
- počet prošlých spamů se sníží na jednotky za den

Vlastnosti úspěšného opatření

- musí mít okamžitý účinek (SA účinkuje ihned)
- nesmí propustit prakticky žádný spam (SA jednotky denně)
- nesmí zahodit žádné legitimní emaily (SA méně než 0.1%)
- SpamAssasin je praktické řešení, díky tomu ostatní metody nemají velkou šanci na úspěch
- navíc velká většina dalších metod může být použita jen jako další kritérium pro SpamAssasin

Závěr

- existuje řada opatření, která by mohla vylepšit situaci
- například SPF v kombinaci s MSA umožňuje majiteli domény zařídit, aby všechny odeslané maily prošly doménovým mailserverem
- okamžitá výhoda je v možnosti rozpoznání falešných zpráv o nedoručení
- bohužel, většina správců tento systém nevyužívá
- cokoliv, co dá jen trochu práce, nemá velkou šanci na úspěch

Možné kroky ze strany CZ.NIC

- propagace SPF, v ČR se používá asi 10x méně než v zahraničí
- pluginy pro hashcash do emailových klientů
- návrh modifikace hashcash pro použití funkcí náročných na paměťové operace