

Spam

**jak nepřivolávat,
nepřijímat a nerozesílat
nevyžádanou poštu**

Martin Adámek

- Prevence, spočívající v obecné ochraně e-mailové adresy
- Ochrana e-mailové adresy při jejím uvádění na webových stránkách
- Ověřování získaných adres a antispamová řešení
- Fyzické principy rozesílání spamu
- „Druhá strana barikády“ – propagace

Upozornění pro čtenáře a uživatele této knihy

Všechna práva vyhrazena. Žádná část této tištěné či elektronické knihy nesmí být reprodukována a šířena v papírové, elektronické či jiné podobě bez předchozího písemného souhlasu nakladatele. Neoprávněné užití této knihy bude **trestně stíháno**.

Používání elektronické verze knihy je umožněno jen osobě, která ji legálně nabyla a jen pro její osobní a vnitřní potřeby v rozsahu stanoveném autorským zákonem. Elektronická kniha je datový soubor, který lze užívat pouze v takové formě, v jaké jej lze stáhnout s portálu. Jakékoliv neoprávněné užití elektronické knihy nebo její části, spočívající např. v kopírování, úpravách, prodeji, pronajímání, půjčování, sdělování veřejnosti nebo jakémkoliv druhu obchodování nebo neobchodního šíření je zakázáno! Zejména je zakázána jakákoliv konverze datového souboru nebo extrakce části nebo celého textu, umisťování textu na servery, ze kterých je možno tento soubor dále stahovat, přitom není rozhodující, kdo takovéto sdílení umožnil. Je zakázáno sdělování údajů o uživatelském účtu jiným osobám, zasahování do technických prostředků, které chrání elektronickou knihu, případně omezují rozsah jejího užití. Uživatel také není oprávněn jakkoliv testovat, zkoušet či obcházet technické zabezpečení elektronické knihy.



Věnováno

Městské knihovně Náchod

Jaroslavi Česenkovi a Evě Bradnové

SPŠ elektrotechnické a VOŠ Pardubice

Fakultě informatiky a managementu Univerzity Hradec Králové

Spam

jak nepřivolávat, nepřijímat a nerozesílat nevyžádanou poštu

Martin Adámek

Vydala Grada Publishing, a.s.
U Průhonu 22, Praha 7
jako svou 3487. publikaci

Odpovědný redaktor Tomáš Vild
Sazba Tomáš Vild
Návrh a grafická úprava obálky Vojtěch Kočí
Počet stran 168
První vydání, Praha 2009

© Grada Publishing, a.s., 2009
Cover Photo © fotobanka allphoto

Autoři i nakladatelství se omlouvají za sníženou kvalitu některých obrázků, způsobenou nižší kvalitou podkladů stažených z internetu.

V knize použité názvy programových produktů, firem apod. mohou být ochrannými známkami nebo registrovanými ochrannými známkami příslušných vlastníků.

Vytiskly Tiskárny Havlíčkův Brod, a.s.
Husova ulice 1881, Havlíčkův Brod

ISBN 978-80-247-2638-0 (tištěná verze) © Grada Publishing, a.s. 2009
ISBN 978-80-247-6310-1 (elektronická verze ve formátu PDF) © Grada Publishing, a.s. 2011

Obsah

Seznam použitých zkratek a slovníček pojmů	9
Předmluva	11
Úvod	13
Obsah a právní úprava spamu.....	15
1.1 Definice spamu.....	15
1.1.1 Co je spam?.....	15
1.1.2 Encyklopedické definice	16
1.1.3 Úvaha: Nejasné hranice spamu a normální pošty	17
1.1.4 Rozšíření spamu	19
1.2 Věcný obsah spamů.....	19
1.2.1 Reklama	19
1.2.2 Ukázka jedné mnohokrát zaslané zprávy	20
1.2.3 Viry a trojské koně	22
1.2.4 Podvodné maily – phishing.....	22
1.2.5 Další druhy podvodných mailů	24
1.3 Antispamový zákon.....	25
1.3.1 Zákon 480/2004 Sb.....	25
1.3.2 Reálná účinnost právních opatření.....	28
1.3.3 Postup při hlášení nevyžádaných zpráv dozorujícímu orgánu	28
Databáze e-mailových adres	31
2.1 Poskytování a získávání e-mailových adres	31
2.1.1 Registrací	32
2.1.2 Sdělením – marketing	33
2.1.3 Zveřejněním – spamboti.....	34

2.1.4	Používáním	36
2.1.5	Existencí – freemail.....	37
2.1.6	Tipováním	37
2.1.7	Obchod s databázemi a úvaha o cílení kampaně.....	39
2.1.8	Důkazy – více než čtyřletý subdoménový výzkum	45
2.2	Ochrana e-mailových adres před spamboty	48
2.2.1	Bez odkazu „mailto:“	48
2.2.2	Zavináč	49
2.2.3	Textový přepis zavináče.....	51
2.2.4	Odstraňte žirafu	54
2.2.5	Polské rébusy a hádanky	55
2.2.6	Bez popisku.....	55
2.2.7	Obecný tvar.....	56
2.2.8	Seznam osob, a samostatně doména	57
2.2.9	Textový popis adresy	57
2.2.10	Zápis adresy ASCII kódem.....	58
2.2.11	Výpis adresy javascriptem	59
2.2.12	Zápis zprava doleva.....	60
2.2.13	Zápis obrázkem.....	61
2.2.14	Webové formuláře.....	63
2.2.15	Vedlejší účinky ochrany adres při zápisu na webu	67
2.2.16	Speciální adresa	67
2.2.17	Dočasnost adresy	68
2.3	Ověřování e-mailových adres.....	68
2.3.1	Obrázky a skripty.....	69
2.3.2	Odhlášení odběru	71
2.3.3	Automatická odpověď v nepřítomnosti.....	75
Rozesílání a filtrování spamu	79	
3.1	Antispamová řešení.....	79
3.1.1	Spamfiltry a blacklisty	79

3.1.2	Whitelisty, a ruční potvrzení zprávy u nových odesílatelů	92
3.1.3	Jednoduché filtry v nastavení mail serveru	96
3.1.4	Jednoduché filtry v nastavení klienta	97
3.1.5	Optická kontrola.....	97
3.2	Snahy o obejití spam-filtrů.....	98
3.2.1	Překlepy a přesmyčky ve filtrovaných slovech	98
3.2.2	Bílé na bílém	101
3.2.3	Nelinearizovatelná tabulka	103
3.2.4	Sdělení v obrázcích	104
3.2.5	Doplnění nesouvisejících slov	105
3.2.6	Zneužívání cizích IP adres	106
3.2.7	Falšování e-mailových adres	106
3.2.8	Forma osobního dopisu	107
3.2.9	Lokalizace do jazyka příjemce.....	110
3.3	Fyzická realizace rozesílání spamu	112
3.3.1	Nabídka rozesílání reklamních e-mailů na klíč – v ČR.....	112
3.3.2	Nabídka rozesílání spamu v Rusku a SNS	113
3.3.3	Distribuované rozesílání ze soukromých počítačů	115
3.3.4	Zneužívání cizích e-mailových adres, falšování odesílatele.....	117
3.4	Náklady na příjem spamu	120
3.4.1	Datová zátěž.....	120
3.4.2	Čas a mzdové náklady	121
3.4.3	Obsah spamu	122
Zásady pro psaní zpráv	123	
4.1	Jak psát osobní zprávy, aby nevypadaly jako spam	123
4.1.1	Adresa odesílatele.....	124
4.1.2	Jméno odesílatele.....	125
4.1.3	Předmět	125
4.1.4	Požadovat potvrzení o doručení	128
4.1.5	Požadovat potvrzení o přečtení	129

4.1.6	Když příjemce nereaguje	130
4.1.7	Telefonická domluva před zasláním zprávy	131
4.2	Druhá strana barikády – informační maily a hromadné nabídky.....	132
4.2.1	Informační zprávy	132
4.2.2	Získávání nových zákazníků	135
4.2.3	Úvaha: Morální aspekt, a právní definovatelnost spamu.....	136
4.2.4	Právní hledisko – aktuální stav, možnosti zneužití.....	137
4.2.5	Přínosy informací pro příjemce, vítaný mail servis.....	141
4.3	Na okraj – příbuzné formy obtěžující reklamy	141
4.3.1	Papírová pošta	142
4.3.2	Ostatní elektronické kanály	144
4.3.3	Zbylé obtěžující reklamní metody.....	147
4.4	Závěrem.....	147
4.4.1	Rekapitulace	147
4.4.2	Poznatky z praxe a výsledky výzkumu	148
4.4.3	Výhledy do budoucnosti	151
	Doslov.....	155
	Seznam použité literatury.....	157
	Zdroje textů a informací.....	157
	Seznam obrázků a jejich zdroje	163
	Rejstřík.....	165

Seznam použitých zkratek a slovníček pojmů

- **BCc, Blind Carbon Copy** (česky „Skrytá kopie“, doslova „Slepá kopie“) – příjemci uvedení v tomto poli zprávu obdrží, aniž by o tom ostatní příjemci zprávy věděli
- **Cc, Carbon Copy** (česky „Kopie“, doslova „průklep“) – od těchto příjemců neočekáváte odpověď, zprávu jim posíláte pouze „na vědomí“ – příjemce, které ve zprávě oslovujete druhou mluvnickou osobou nebo od kterých očekáváte odpověď, uvádějte v poli „To“
- **CSD (Circuit Switched Data)** – vytáčené datové připojení (mobilní nebo pevné; platí se za čas připojení)
- **GPRS (General Packet Radio Service)** – stálé datové připojení přes mobilní telefon GSM (platí se za přenesená data, resp. dle nabídky operátora příp. paušálem za měsíc)
- **EDGE** – rychlejší varianta GPRS (v ceníku či vyúčtování se tyto technologie nerozlišují)
- **From („Od“)** – pole, v němž je uvedena e-mailová adresa odesílatele e-mailové zprávy
- **imap4** – protokol pro oboustrannou synchronizaci elektronické pošty klienta se serverem (odesílání i příjem)
- **IP adresa** – fyzická adresa počítače připojeného k internetu, resp. „adresa zásuvky“
- **Komunikátor** – kombinace kapesního počítače a mobilního telefonu
- **MDA** – původně komerční název pro konkrétní kombinaci PDA a mobilního telefonu, stal se synonymem pro „komunikátor“

- **PC, Personal Computer** – osobní počítač (nikoliv každý počítač)
- **PDA, Personal Digital Assistant** – kapesní počítač
- **pop3** – protokol (a server) pro stahování příchozí pošty
- **Reply-to** – adresa, na kterou se má odesílateli odpovědět (klient na ni adresuje odpověď)
- **SMTP** – protokol (a server) pro odesílání odchozí pošty
- **To** (česky „Komu“) – příjemci zprávy
- **topdoména** – doména nejvyššího řádu – část adresy za poslední tečkou (např. „cz“, „com“)
- **URL** – webová adresa (mj.), kterou vidíte v adresním řádku prohlížeče
- **W3C (World Wide Web Consortium)** – souhrn standardů pro tvorbu validního webu, měly by být závazné alespoň pro profesionální webmastery
- **W4D (Worth World Wide Web Dogma)** – souhrn doporučení pro tvorbu kvalitního webu
- **XHTML 1,0 Strict** – Jeden ze standardů W3C, v současné době doporučovaný k používání

Předmluva

E-mailový systém umožnil lidem snadnou, rychlou a levnou komunikaci. Bobužel tak snadnou a levnou, že mnoho firem nabízí své služby či výrobky komukoliv (všem), ignorujíc poučku Freda Habna „psí krmivo nenabízíme chovatelům koček“ [38]. E-mail je prostě tak přístupný a levný, že se to vyplatí. „Řekněme o sobě každému, však nás to nebolí“.

Jenže ono to bolí příjemce takových zpráv, protože reklamních zpráv je příliš mnoho. Reklamy ničí či zneprůjemňují mezilidskou komunikaci, využívajíce přitom paradoxně prostředek, který sám o sobě komunikaci usnadňuje. Možná namítnete, že kdyby nebyl e-mail, nebyl by ani spam (ne v podobě, jak jej známe dnes), a že tedy spam jen snižuje výboby e-mailu, a neomezuje ty druhy komunikace, které tu byly před e-mailem. Jenže e-mail zkrátka je. A lidstvo je na něm závislé, takže spam lidstvu (lidem, konkrétním člověkům) ničí komunikaci. Osobní i pracovní.

Martin Adámek,

Náchod, červen 2008



Úvod

Cílem této knihy je seznámit laického čtenáře se zásadami ochrany e-mailové adresy, navrhnout webmasterům způsoby bezpečného a současně přístupného zápisu adres na webu, poodbalit zvědavým jedincům vnitřní principy fungování spamu, stručně popsat nejrozšířenější spamfiltry, občas nastínit možnosti příštího vývoje a upozornit na morální a etické aspekty spamu. V neposlední řadě také kniha uvádí zásady, které by měly být dodržovány při psaní osobních zpráv, aby nebyly omylem považovány za spam, a při psaní obchodních zpráv, aby nebyly spamenem.

Ač napsána informatikem, snaží se kniha o komplexní pohled na spam. Autor si uvědomuje, že nedílnou součástí komunikace (vč. spamu) jsou lidé. Proto na spam nablíží nejen z hlediska ICT (informačních a komunikačních technologií), ale i z pohledu marketingového, psychologického, či eticko-morálního.

Kniha vychází z pozorování vyžádaných i nevyžádaných zpráv, přijímaných na různé adresy, jakož i z pozorování chování uživatelů.

Knihou pojednává o stavu problematiky platném v době zpracovávání, byla dokončena v červenci 2008.

Některé jevy se mohou během poměrně krátké doby zásadně změnit.

1.

Obsah a právní úprava spamu

1.1 Definice spamu

1.1.1 Co je spam?

Spam je nevyžádaná reklamní pošta, nebo jiné nevyžádané sdělení, zpravidla komerčního charakteru, rozesílané hromadně mnoha příjemcům. Ale názory na to, co přesně je spam, se liší.

- Za spam můžeme jednoznačně považovat zprávy zasílané automaticky mnoha adresátům, kteří si tyto zprávy neobjednali, a zejména nemají možnost odběr těchto zpráv zrušit. Tyto zprávy zpravidla přicházejí ze zahraničí, jsou psány nejrůznějšími jazyky (převládá angličtina) a nabízejí nejrůznější zboží – nejčastěji nelegální software nebo Viagru a podobné prostředky. Tyto zprávy často chodí opakovaně stejným příjemcům, přičemž mnohdy obsahují velmi podobnou nebo stejnou nabídku – upraveny jsou pouze vnitřně, aby je nezachytil spam-filtr, ale čtenáři se jeví jako stále stejná zpráva.
- Slabší forma spamu je víceméně jednorázové a často ruční zasílání nabídek českými firmami, se spolehlivou možností zrušit odběr dalších zpráv – resp. většinou stačí nedat souhlas se zasíláním dalších zpráv a seriózní česká firma to respektuje. Takové e-maily mne většinou nechávají poměrně klidným (pokud mne jimi odesílatel nezahrnuje opakovaně a pokud nefalšuje adresu odesílatele), a občas odesílateli na oplátku zašlu nabídku svoji. Např. provozovatel nějaké horské chaty našel můj web o cykloturistice, a poslal mi nabídku cyklistického víkendů, který pořádal. Tak jsem mu jako doplňkový program nabídl cestopisnou prezentaci o mé cestě na kole do Monaka. Sice se nakonec ta víkendová akce pro nedostatek zájemců vůbec nekonala, ale jednou třeba může taková relevantní protinabídka přinést ovoce.
- Někdy se také jako spamování hovorově označuje i hromadné přeposílání zajímavých e-mailových zpráv na distribuční seznam. Ale jelikož tito příjemci o přeposílání mají zájem, sami se ho aktivně zúčastňují, a mají možnost se kdykoliv z příjmu zpráv odhlásit, nelze v tomto případě o spamu mluvit.
- Za spam samozřejmě nelze označit ani vyžádaný odběr informací z nějakého serveru nebo i příjem komerčních sdělení (akčních nabídek) nějaké firmy – pokud se příjemce sám zaregistroval, a odesílatel si autentičnost této registrace ověřil, například vhodně napsaným skriptem, který po registraci odesílá kontrolní mail a vyžaduje potvrzení registrace. Samozřejmostí je pak možnost kdykoliv odběr zpráv ukončit.

1.1.2 Encyklopedické definice

Existuje mnoho definic spamu, podíváme se jen na některé z nich:

Server *answers.com* [2] k heslu „spam“ uvádí:

- „Unsolicited e-mail, often of a commercial nature, sent indiscriminately to multiple mailing lists, individuals, or newsgroups; junk e-mail.“ [2].

A zmiňuje i předpokládaný původ tohoto slova:

- „From SPAM (probably inspired by a comedy routine on the British television series Monty Python's Flying Circus, in which the word is repeated incessantly).“ [2].

V české mutaci otevřené encyklopedie Wikipedie [4] nalézám nejen další definici, ale i podrobnější informaci o původu tohoto slova:

- „Spam je nevyžádaný masově rozesílaný e-mail, případně jiné druhy internetové komunikace, nejčastěji reklamní. Používá se též zkratka UBE/UCE (Unsolicited Bulk/Commercial Email). Pro opak spamu, tj. poštu, která je zaslána konkrétní osobou se specifickým jednorázovým účelem a adresát ji považuje za žádoucí, se řidčeji používá termín ham (anglicky šunka).“ [4].

1. Obsah a právní úprava spamu

- „Název pochází ze značky amerických konzerv lančmítu (...), která se vyrábí od 30. let dodnes (v současnosti ale výrobce trvá na psaní velkým písmem SPAM) a za 2. světové války a po ní byla hojně rozšířená a stále méně oblíbená ve Velké Británii. Proto se objevuje v závěrečném skeči 25. dílu seriálu Monty Pythonův létající cirkus, kde všechny položky jídelního lístku v restauraci obsahují spam, i mnohokrát opakovaně, a spory zákazníků s číšnicí o objednávky přerušuje skupina Vikingů zpívajících “Spam, spam, spam...” [4].
- „Označení tak bylo přijato nejprve pro praktiku mnohonásobného rozeslání téže zprávy na Usenetu, ale pak se význam posunul pro zneužívání skupin k šíření různých nepřípadných textů a přímo reklamy a zachoval se i poté, co se těžiště takových aktivit přesunulo do e-mailu.“ [4].



Server *Coto.je* [3] nenachází k výrazu „spam“ ani jednu položku. Tím získáváme důkaz, že „spam“ jako takový je poměrně nový jev, který přišel až s novými elektronickými prostředky komunikace – tento server vyhledává z Ottovy encyklopedie, Malé československé encyklopedie a Encyklopedie Universum.

Pokud budete chtít vidět klip, díky kterému nevyžádané reklamy získaly název „spam“, zadejte např. na serveru *www.youtube.com* do vyhledávacího pole např. „Monty Python Spam“. Pokud tvorbu Monty Pythonů neznáte, tak předem upozorňuji, že skeč je poměrně svérázný. Nicméně právě tento skeč o hostech, kteří se v restauraci dožadovali jídla bez spamu, je předlohou označení „spam“ pro obtěžující reklamy.

Nicméně problém nevyžádané pošty obecně, je starší než „spam“ a dokonce starší než e-mailový systém.

Původ anglického výrazu „junk mail“, který lze volně přeložit jako „braková pošta“, nacházím v knize „Do-it-yourself advertising“ [38]. Podle jejího autora termín „junk mail“ poprvé použil publicista Westbrook Pegler, jako název několika novinových sloupků, ve kterých se zlobil na firmu, která jeho velmi mladé dceři zaslala nevhodné nabídky prádla. Název „junk mail“ se zalíbil i jeho kolegům z dalších novin, a význam tohoto pojmu se přesunul na celé odvětví zásilkového prodeje. K těmto událostem mělo dojít ve třicátých nebo čtyřicátých letech dvacátého století, což znamená, že etické problémy související s poštovními nabídkami, jakož i nutnost přesného cílení nabídky podle bližších informací o adresátech, nejsou žádnou novinkou.

1.1.3 Úvaha: Nejasné hranice spamu a normální pošty

Kde přesně je vlastně hranice spamu?

Spam vznikl proto, že e-mail umožňuje velmi snadnou, levnou, a rychlou komunikaci. Primárně mezi jednotlivci nebo v menších skupinách. Ale při hromadné korespondenci se to vše samozřejmě projevuje ještě mnohem více. Nemusíte tisknout tisíc dopisů, potiskovat tisíc obálek a dopisy do nich vkládat, ani platit tisíc poštovních známek.

Jedna z definic spamu je, že jde o nevyžádanou nabídku. Ale může bez nabídky fungovat obchod? A kolik procent nabídek je vyžádaných? Takže kde je etická hranice, kdy už nevyžádaná nabídka vadí, a tudíž ji můžeme označit jako spam?

Když zahraniční firma rozesílá na náhodné adresy nabídky hodinek a jiných produktů, je to bezpochyby spam.

Ale na druhou stranu, když napíšete knížku, a nabízíte ji postupně různým vydavatelstvím, tak je to myslím v pořádku. A když tu nabídku rukopisu pošlete emailem, vadí to? Myslím, že ne.

Takže teď máme dva příklady, jeden pro každý z extrémů, kdy se o spam jedná, a kdy ne.

Ale co ty případy mezi nimi? Jak určit hranici?

Když výrobce cyklistických potřeb rozešle nabídku cyklistům, které „ručně“ našel na internetu, je to spam?

Když nějaká firma pošle nabídku potenciálním zákazníkům, vybraným např. podle oboru nebo lokality, je to spam?

Když někdo pořádá cestovatelskou expedici, najde na internetu sto potenciálních zájemců, a rozešle jim stejný nabídkový e-mail (byť je jeho zájem nekomerční), je to spam?

Žádná z uvedených nabídek nebyla vyžádaná. Ale jsou všechny nemorální nebo neetické, abychom je mohli označit za spam, nežádoucí projev vyspělé úrovně komunikačních technologií?

Mnohé z těch nabídek jsou nejen „normální“ nebo „přijatelné“, ale často jsou dokonce některými příjemci i vítané, protože jsou pro ně zajímavé. Mezi pojmy „nevyžádaná“ a „nežádoucí“ je totiž dost velký rozdíl.

Hranice mezi spamem a přijatelnou nabídkou je opravdu velmi těsná, navíc ji asi každý může vidět úplně někde jinde.

Pokud bych měl z této krátké úvahy udělat nějaký závěr, případně mi, ze čistokrevný spam je pouze nabídka rozeslaná „hlava-nehlava“ na velké množství adres, získaných nejrůznějšími způsoby, bez ohledu na to, komu dané adresy patří a co dotyčného člověka zajímá.

Zatímco nabídka rozeslaná menší skupině adresátů, vybraných podle klíče „zajímá ho to?“ asi spam není. Ale to vyžaduje zpracování osobních údajů, mnohdy citlivých, což už je další oblast „informačního a komunikačního práva“, a ona cílená reklama pak sice nemusí být tak nepřijemná z hlediska „spam × nespam“, ale může narazit v oblasti osobních údajů.

Navíc takové přesně cílené nabídky, využívající osobní údaje, jsou mnohem účinnější, a slibují větší šanci odezvy adresáta. Z toho plyne další morální či etický problém, protože taková reklama je pak zákeřnější, osobně cílená, a každému nabídne to, co by mohl opravdu potřebovat. Což na jednu stranu znamená, že se příjemce nabídkou necítí být obtěžován, ale na druhou stranu pravděpodobně nabídku podrobněji prostuduje, ztratí s jejím čtením více času, a nakonec ze sebe ještě nechá vylákat peníze, protože si nabízenou věc koupí s větší pravděpodobností, než u necílené reklamy.

Z makroekonomického pohledu to pak sice může znamenat růst HDP, ale ten s sebou také ponese větší plýtvání, způsobené zbytečným vyřazením doposud funkční