Průzkum využití prostředků obrany proti spamu

Součástí mé diplomové práce byl návrh metodiky pro obranu proti spamu, zaměřené zejména na výběr vhodných protiopatření. Jako jedno z východisek pro vytvoření metodiky byl také zvolen průzkum zkušeností uživatelů elektronické pošty s prostředky obrany proti spamu. Hlavním cílem bylo získat dostatečné množství relevantních dat pro nalezení odpovědí na tyto otázky:

- Jaký problém pro uživatele spam představuje – objektivně (počty zpráv) i subjektivně (jak uživatelé spam vnímají)?
- Jaké metody obrany proti spamu jsou používány?
- Jaká je úspěšnost jednotlivých metod v obraně proti spamu?

Nemalou mírou k úspěšnosti průzkumu přispěly odborné internetové magazíny underground.cz, POOH.CZ, Lupa a Technet.cz, kterým bych tímto rád poděkoval. Také bych rád poděkoval všem těm, kteří věnovali pár minut vyplnění a odeslání dotazníku.

Popis použité metody

Odůvodnění podoby dotazníku

Pro sběr dat byla zvolena metoda dotazníkového šetření. Dotazník byl navrhován tak, aby korespondoval se strukturou diplomové práce. Otázky proto byly rozděleny do šesti graficky oddělených bloků:

- Základní informace,
- Prevence,
- Blokování,
- Software,
- Přímý kontakt,
- Právo.

Každý blok obsahuje několik otázek týkajících se zejména konkrétního chování respondenta (např. zda filtruje elektronickou poštu) a je ukončen otázkou, která na pětibodové škále (rozhodně ano, spíše ano, nevím, spíše ne, rozhodně ne) zjišťuje postoj k danému tématu. Každá otázka byla také doplněna krátkým vysvětlujícím textem. Typ otázky byl volen podle charakteru odpovědí – byly použity tyto typy otázek:

- výběr jedné možnosti (v dalším textu označeno jako „1“),
- výběr více možností („n“),
- zadání konkrétní číselné hodnoty („x“),
- u některých otázek bylo součástí také doplňující pole, které umožnilo respondentovi odpovědět i jinou, než nabízenou odpovědí („d“).

Cílem bloku *Základní informace* bylo zjistit základní informace o respondentovi a jeho postoji ke spamu. První otázka (Jaká je vaše pozice?, 1) sloužila k rozdělení respondentů do segmentů pro přesnější analýzu. Následující série třech otázek (Kolik poštovních schránek spravujete?, x; Kolik e-mailových zpráv denně dostáváte?, x; Jaký podíl z těchto mailů tvoří spam?, 1) zjišťovala absolutní čísla o objemu zpráv a podílu spamu. Další otázka (Jste si vědom, odkud získal spammer vaši e-mailovou adresu?, n) testovala, zda si jsou respondenti vědomi odkud byla získána jejich adresa a také k identifikaci nejčastějších způsobů získávání adres. Poslední dvě otázky zjišťovaly postoj uživatelů ke spamu z pohledu aktivita/pasivita v obraně (Bráníte se aktivně proti spamu?, 1) a z pohledu vážnosti situace (Považujete spam za vážný problém?, 1).

Blok *Prevence* obsahoval otázky týkající se vyhýbání se spamu. První otázka (Máte k dispozici rady jak se vyhnout spamu a jak s ním zacházet?, n) zjišťovala jak moc jsou uživatelé elektronické pošty informováni o prevenci spamu. Následující otázka (Používáte pro riskantní situace zvláštní e-mailovou adresu?, 1) pomáhala odhadnout, jak moc uživatelé využívají rozdělení adres podle stupně bezpečnosti. Poslední otázka zjišťovala odhadovanou účinnost prevence (Považujete prevenci spamu za dostatečně účinný prostředek?, 1).

V bloku nazvaném *Blokování* byl zjišťován přístup respondentů k eliminaci podezřelých e-mailů podle různých kritérií – podle klíčových slov (Filtrujete e-maily podle klíčových slov?, 1), nebo podle původu zprávy (Blokujete poštu od jiných než povolených uživatelů?, 1; Blokujete poštu od konkrétních uživatelů/z konkrétních serverů?, nd). Na závěr byl opět zjišťován postoj k účinnosti těchto metod (Považujete blokování spamu za dostatečně účinný prostředek?, 1).

Další blok zkoumal možnosti využití speciálního *Software*. První otázka zjišťovala zda respondent využívá software a v jaké podobě (Používáte nějaký software na obranu proti spamu?, n). Dále jsem se pokusil identifikovat konkrétní softwarové produkty (Který software na obranu proti spamu používáte?, nd) – respondent měl na výběr z pěti známých produktů, mohl doplnit další. Blok ukončovala otázka zjišťující postoj k účinnosti software (Považujete software za dostatečně účinný prostředek?, 1).

Blok *Přímý kontakt* byl zaměřen na kontakt poškozeného se spammerem nebo poskytovatelem jeho připojení a úspěšnost tohoto počínání (Snažil jste se někdy kontaktovat spammera?, 1; Snažil jste se někdy kontaktovat poskytovatele připojení/správce pošty serveru, odkud vám přišel spam?, 1) a na účinnost přímého kontaktu (Myslíte si, že má smysl kontaktovat spammera/jeho ISP?; 1).

Poslední blok nazvaný *Právo* zkoumal možnosti právní cesty obrany proti spammingu. První otázka se zaměřovala na zkušenosti respondentů (Snažil jste se někdy bojovat proti spammerovi právní cestou?, 1), druhá otázka zjišťovala účinnost právní obrany (Myslíte si, že je současné právo dostatečně účinné v boji proti spamu?, 1).

Dotazník byl doplněn textovým polem, do kterého mohli respondenti zadat své připomínky k dotazníku nebo k tématu (které se v dalším průběhu tvorby diplomové práce ukázaly být velice podnětné). Ti respondenti, kteří zadali také e-mailovou adresu, byli o výsledcích průzkumu informováni elektronickou poštou.

Technické provedení

Jako nejvhodnější forma pro dotazník byla vyhodnocena forma webového formuláře, který ukládá data do databáze, a to zejména z těchto důvodů:

- jednoduché zadání údajů uživatelem,
- stálý přehled o aktuálním stavu šetření,
- možnost analýzy dat pomocí přímých dotazů na databázi,
- snadný přenos dat do dalších analytických nástrojů.

Vzhledem ke zkušenostem a snadné dostupnosti byla zvolena kombinace jazyka PHP, databáze MySQL a školního linuxového serveru Sorry. Jako uživatelské rozhraní byla zvolena stránka v jazyce HTML, zejména formulářové prvky tohoto jazyka (FORM, INPUT, SELECT a podobně). Uživatel požádá server o provedení skriptu index.php, který nejprve zobrazí zmíněný formulář. Po vyplnění formuláře uživatel stiskne tlačítko pro odeslání. Pomocí JavaScriptové funkce je provedena kontrola správnosti zadání číselných údajů a data jsou následně metodou POST (odesílaná data nejsou součástí URL, ale součástí HTTP žádosti) předána zpět skriptu index.php, který je uloží do databázové tabulky a informuje uživatele o výsledku, resp. ošetří chybové stavy.

Tabulku, do které se ukládají data, tvořil jeden sloupec pro každou otázku s možností výběru jedné odpovědi nebo zadání konkrétní odpovědi a jeden sloupec pro každou odpověď k otázce s možností výběru více odpovědí. Tabulka dále obsahovala sloupec, do kterého byla ukládána IP adresa odesílatele, z důvodu ochrany osobních údajů zakódována algoritmem MD5, pro kontrolu případných duplicitních záznamů. Tabulka byla několikrát týdně zálohována.

Při analýze výsledků bylo použito dvou metod. Některé údaje, zejména četnosti, byly získávány přímo z databáze pomocí SQL dotazů. Celá tabulka pak byla přenesena přes formát CSV do aplikace Microsoft Excel, kde byly prováděny analýzy, které nebylo technicky možné provést na databázi, a kde byly připraveny tabulky a grafy.

Některé z řádků, které obsahovaly podezřelá data (zejména extrémně vysoké hodnoty počtu schránek kombinované s velice nízkými podíly zpráv na schránku a neuvedenou e-mailovou adresou), byly z databáze vyřazeny, aby nebyly těmito extrémy ovlivněny statistiky absolutních počtů.

Výběr, segmentace a oslovení respondentů

Před zahájením tvorby dotazníku bylo nutné rozhodnout o výběru a segmentaci respondentů. V úvahu připadaly dva postupy – omezit respondenty na konkrétní segmenty nebo výběr respondentů neomezovat a provádět segmentaci dodatečně otázkou v dotazníku. Zvolena byla nakonec druhá varianta, a to zejména z těchto důvodů:

- dal se očekávat větší počet respondentů, a tak i větší vypovídací schopnost statistik,
- snadnější oslovení respondentů – respondenty lze oslovit plošně (pomocí odkazů na webových stránkách, v diskusních skupinách a podobně); druhá varianta by vyžadovala oslovení přímé,
- dodatečná segmentace uživatelů umožňuje interpretovat data z různých perspektiv (postmaster může mít na věc jiný pohled než běžný uživatel).

Respondenti měli možnost přihlásit se k jednomu z pěti segmentů podle pozice v první otázce dotazníku. Byly zvoleny tyto segmenty:

- *Postmaster* – může poskytnout čistě technologický pohled zaměřený úzce na problematiku pošty.
- *IT manager* – stále technologický pohled, ale již v širším kontextu.
- *Manager* – pohled prizmatem efektivity, produktivity, nákladů a podobně.
- *Uživatel* – pohled běžného uživatele zaměřený zejména na konkrétní problémy (náklady za připojení, vyrušení, nutnost mazat spamy a podobně).
- *Jiná pozice* – pro respondenty, kteří se nechtějí nebo nemohou zařadit do předchozích segmentů.

Protože se otázky daly z různých pozic chápat různě, byly potenciálně sporné otázky upřesněny ve vysvětlujícím textu. Uživatelé, kteří si přáli odpovědět z více pozic, měli možnost odeslat otazník vícekrát. Nebylo povinné odpovídat na všechny otázky, neboť v určitých situacích z některých pozic nemusely dávat smysl, nebo nemohly být pro neodbornou veřejnost srozumitelné.

Oslovení uživatelů probíhalo v několika fázích. V první fázi bylo o dotazníku informováno několik desítek uživatelů pomocí komunitních webových stránek a IM ICQ a Jabber. Několik z nich dotazník vyplnilo a předalo připomínky, na základě kterých byly některé otázky v dotazníku upraveny. Vzhledem ke lhůtám byl odkaz na dotazník zadán do několika internetových vyhledávačů (Seznam, Atlas, Centrum). Jako klíčový nástroj oslovení uživatelů byly zvoleny internetové magazíny pro odbornou veřejnost, které se systematicky zabývají problematikou spamu. Za umístění odkazu jim bylo přislíbeno předání výsledků dotazníkového šetření. Několik magazínů odkaz na své stránky umístilo, některé dokonce samy od sebe.

Harmonogram

27. 2. 2003 – formulář zprovozněn, během prvních několika dní prováděny drobné úpravy na základě zpětné vazby od několika vybraných uživatelů

28. 2. 2003 – přidání odkazu na dotazník do vyhledávačů Seznam, Atlas a Centrum. Odeslán e-mail s žádostí o zveřejnění odkazu na pět českých serverů zabývajících se spamem.

1. 3. 2003 – zobrazení odkazu na serveru Underground.cz, ve vyhledávači Atlas

3. 3. 2003 – odkaz zobrazen na serverech Technet.cz, Lupa, POOH.CZ

5. 3. 2003 – odkaz zobrazen ve vyhledávači Centrum

18. 3. 2003 – odkaz zobrazen na stránkách České společnosti pro systémovou integraci

Na průzkum bylo také průběžně upozorňováno v diskusních skupinách (USENET) a v diskusích k článkům na téma spam na odborných serverech.

16. 4. 2003 – ve večerních hodinách byl dotazník odstraněn

Výsledky šetření

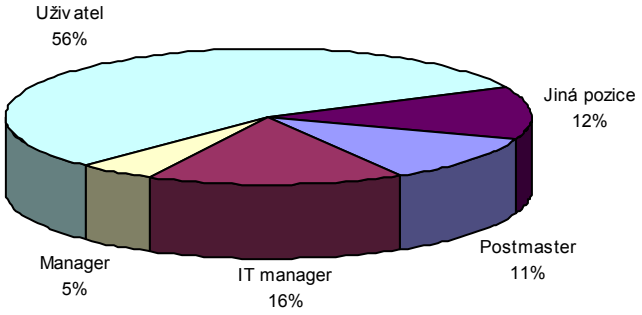
Po skončení průzkumu se v databázi nacházelo 697 řádek, tedy 697 odeslaných dotazníků. Po počáteční analýze byly dva řádky označeny jako velice podezřelé (velmi vysoké počty schránek, které tvořily více než polovinu celkového počtu, při naprosto minimálním počtu zpráv – necelé dvě zprávy na tisíc schránek a den), neobsahovaly e-mailovou adresu, která by umožňovala údaje ověřit, a proto byly z databáze odstraněny. Dále

byly odstraněny další dva řádky, které se ukázaly jako duplicitní. Po vyčištění tedy tabulka obsahovala 693 řádek.

Jeden ze sloupců tabulky obsahoval zakódované IP adresy, které měly sloužit zejména pro kontrolu duplicitních řádek. Dohromady bylo v tomto sloupci 633 unikátních IP adres. Rozdíl mezi celkovým počtem řádek a počtem IP adres se dá vysvětlit tím, že uživatelé měli možnost odesílat dotazník z různých pohledů vícekrát. Nelze však říci, že by to znamenalo přesně 633 respondentů, protože z některých IP adres byly prokazatelně (podle kombinace s uvedenou e-mailovou adresou) odesílány dotazníky více uživatelů. Dá se tedy předpokládat, že celkový počet osob, které na dotazník odpověděly, je necelých 640.

Základní informace

Blok nazvaný *základní informace* měl za cíl zjistit základní informace o respondentech a jejich postoji ke spamu. První otázka (Jaká je vaše pozice?) sloužila k segmentaci respondentů. Svou pozici uvedli všichni dotazovaní, rozdělení je zobrazeno v obrázku 1.



Obrázek 1 – Rozdělení segmentů

V absolutních číslech tedy na dotazník odpovědělo 77 postmasterů, 109 IT managerů, 33 managerů, 388 uživatelů a 86 respondentů zastává jinou pozici.

Následující otázky měly za úkol zjistit absolutní i relativní čísla týkající se objemu zpráv a podílu spamu. Celkový počet spravovaných schránek v průzkumu dosáhl hodnoty 404687. Mezi jednotlivé pozice byly počty a průměry schránek na respondenta rozloženy takto:

Tabulka 1 – Objemy a podíly spamu podle segmentů

	Součet	*Průměr*	*Medián*	*Směrodatná odchylka*
Postmaster	388057	504	100	39613
IT manager	13394	123	15	454
Manager	491	15	40	139
Uživatel	1353	1,49	3	2,46
Jiná pozice	1392	16	20	1070

Asi největší vypovídací hodnotu má z těchto statistik medián, neboť hodnoty u každé pozice obsahují poměrně značné extrémy, což potvrzují vysoké hodnoty směrodatných odchylek (například po odstranění řádku s nejvyšším počtem schránek u pozice postmaster by došlo ke snížení průměru na 494 a zejména směrodatné odchylky na 1861). Vzhledem k rozdílnosti charakteristik jednotlivých skupin a také k množství respondentů tedy lze za nejpřesnější statistiku považovat počet e-mailových schránek běžného uživatele, který se asi skutečně bude pohybovat někde kolem dvou nebo tří.

Celkový denní počet zpráv všech respondentů je 398696, což odpovídá celkovému průměru 0,99 zprávy na schránku a den. Tento počet se zdá být poměrně nízký, je však ovlivněn zejména tím, že postmasteři spravují

velké množství schránek, které však uživatelé příliš nepoužívají. Naopak lze očekávat, že na tento dotazník odpovídali ne úplně běžní uživatelé, ale spíše uživatelé, kteří mají o problematiku elektronické pošty zájem a jejích služeb využívají relativně častěji. Tuto teorii podporují i průměry rozdělené podle počtu schránek – průměr zpráv na schránku pro respondenty, kteří uvedli nejvíce deset schránek, byl 13,17, ale pro respondenty s více než 10 schránkami byl průměr pouhých 0,92. Potvrzením může být i rozložení mezi jednotlivé segmenty – postmaster 0,70, IT manager 7,12, manager 7,44, uživatel 10,24 a jiná pozice 10,82.

Jednou z nejdůležitějších statistik, která měla z tohoto průzkumu vyplynout je podíl spamu na celkovém množství zpráv. Ihned od počátku průzkumu se podíl ustálil velice blízko třiceti procentům, nakonec se zastavil na čísle 29,4517 % (směrodatná odchylka 24,3588). I rozdíly mezi jednotlivými segmenty jsou poměrně malé: postmaster 23,51 %, IT manager 26,88 %, manager 26,97 %, uživatel 31,88 % a jiná pozice 28,02 % (nižší číslo u postmasterů lze opět zdůvodnit značným množstvím neaktivních uživatelů). Zajímavý pohled přináší přepočet na absolutní čísla resp. denní průměry v tabulce 3.

Tabulka 3 – Podíl spamu a průměrný počet spamu za den

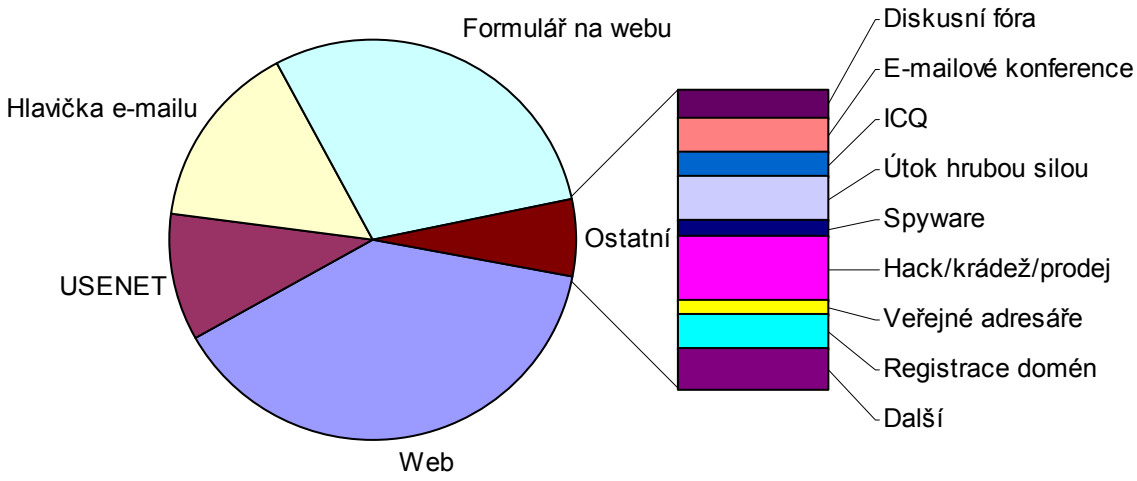
	Podíl spamu	*Průměrný počet spamu za den*
Postmaster	23,51 %	0,16
IT manager	26,88 %	1,91
Manager	26,97 %	2,01
Uživatel	31,88 %	3,26
Jiná pozice	28,02 %	3,03
Celkem průměr	29,45 %	0,29
Celkem absolutně	29,45 %	117422,75

I při celkovém průměru 0,29 spamové zprávy za den to pro průměrného uživatele znamená 106 zpráv ročně. Zajímavý by proto mohl být tento odhad – při jedné a půl schránce na jednoho českého uživatele internetu to znamená zhruba jeden spam na dva uživatele denně. Při počtu uživatelů českého internetu 2,75 miliónu 0 to znamená zhruba 291 miliard zpráv ročně. Podle [2] je průměrná velikost spamové zprávy asi 4,2 kB, což ročně znamená zbytečný přenos asi 1,14 TB dat.

I u podílu spamu je zajímavé porovnání průměrných podílů pro různé počty schránek na uživatele. Pro uživatele, kteří spravují méně než deset schránek tvoří podíl spamu 30,99 %, od deseti do padesáti schránek je to 26 % a pro vyšší počet schránek pouhých 22,09 %. I tento jev lze pravděpodobně vysvětlit přítomností neaktivních uživatelů.

Otázka „Jste si vědom, odkud získal spammer vaši e-mailovou adresu?“ nabízela respondentům možnost vybrat více možností a také doplnit další způsoby. Výsledky by měly být podkladem pro návrh prevence proti spamu.

Obrázek 4 ukazuje poměry jednotlivých způsobů získání e-mailové adresy. Koláčový graf obsahuje přednastavené možnosti, ve sloupcovém grafu jsou uvedeny další způsoby, které uživatelé uvedli slovně. Jako nejčastější způsoby úniku dat byly uváděny ty, které nějakým způsobem souvisí s webem – více než polovina respondentů (385, 56 %) uvedla jako zdroj zveřejnění na webových stránkách, poměrně značný počet se také vyjádřil ve prospěch formulářů na webu (296, 43 %). Ze způsobů, které uživatelé uváděli slovně, byla nejčastěji zmiňováno nelegální jednání (napadení serveru, krádež dat nebo jejich nelegální prodej, v grafu označeno jako „Hack/krádež/prodej“), skoro v polovině případů spojováno přímo s poskytovateli freemailových služeb (českých i zahraničních). Překvapivě častým zdrojem byl také útok hrubou silou (náhodné generování řetězců, které by mohly tvořit část e-mailové adresy před zavináčem).



Obrázek 2 – Způsoby získání e-mailové adresy

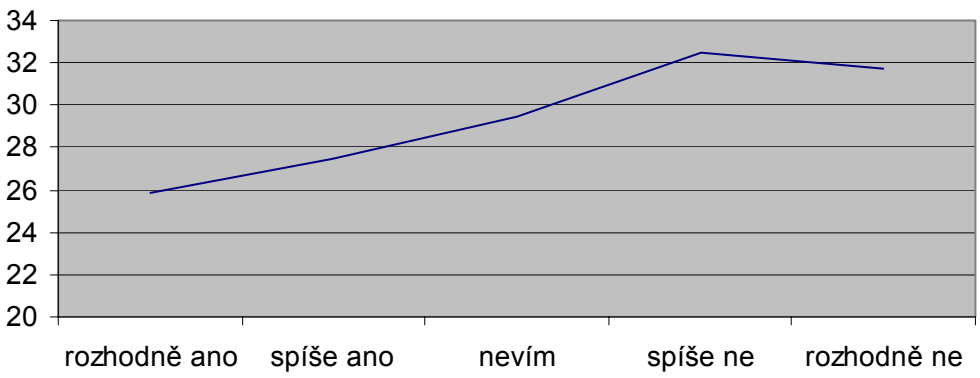
Na otázku „Bráníte se aktivně proti spamu?“ se bohužel vzhledem k chybě ve skriptu ukládajícím hodnoty do databáze nepodařilo získat relevantní odpovědi.

Poslední otázka tohoto bloku zkoumala postoj respondentů ke spamu – „Považujete spam za vážný problém?“ Stejně jako u dalších otázek, které uzavírají jednotlivé bloky, i zde měli dotazovaní možnost vybírat jednu z možností na škále od jedné do pěti (resp. vybírat slovně – rozhodně ano, spíše ano, nevím, spíše ne, rozhodně ne). Průměrné hodnocení této otázky bylo 1,62 (spíše ano, směrodatná odchylka 0,88; četnosti jednotlivých odpovědí [382,240,16,49,3]). Poměrně zajímavý pohled přineslo průměrné hodnocení jednotlivých segmentů – postmaster 1,53, IT manager 1,67, manager 1,91, uživatel 1,63 a jiné pozice 1,45. Postoje postmasterů a běžných uživatelů se daly odhadovat, zajímavá je přece jen vyšší tolerance zejména managerů a IT managerů, u kterých se dal kromě pohledu přímých nákladů také očekávat pohled prizmatem produktivity. Zajímavý je také velice kritický pohled respondentů z jiných pozic.

Prevence

Cílem bloku nazvaného *prevence* bylo získat informace o tom, zda uživatelé využívají některá z preventivních opatření a jak hodnotí jejich účinnost.

Příjemným překvapením byly určitě výsledky první otázky – „Máte k dispozici rady, jak se spamu vyhnout a jak s ním zacházet?“ 431 respondentů (62 %) uvedlo, že má k dispozici rady, jak se spamu vyhnout. Paradoxně však tito respondenti vykazovali vyšší podíl spamových zpráv než ti, kteří rady k dispozici neměli, rozdíl však byl na hranici významnosti a dal by se vysvětlit například tím, že poučení uživatelé jsou na internetu aktivnější. Rady, jak se spamem zacházet, má k dispozici 411 (59 %) respondentů. Alternativní e-mail pro zveřejnění kontaktních informací nebo jiné riskantní využití používá 552 uživatelů (tedy celých 80 %), 133 nikoliv.



Obrázek 3 – Závislost mezi názorem na účinnost prevence a podílem přijímaného

Důležitost prevence v boji proti spamu je průměrně hodnocena známkou 2,83 (nevím; směrodatná odchylka 1,16, četnosti jednotlivých odpovědí [67,269,76,245,29]). Mezi jednotlivými segmenty nebyly zpozorovány výraznější rozdíly – postmaster 2,87, IT manager 2,79, manager 2,85, uživatel 2,84 a jiné pozice 2,73.

Poměrně zřejmá (korelační koeficient 0,95) se ukázala závislost mezi účinností prevence a podílem přijímaného spamu. Respondenti, kteří si myslí, že prevence je dostatečně účinným nástrojem, vykazují až o šest procent nižší podíl spamu než ti, kteří prevenci nevěří (viz obrázek 3). Zda je tato závislost vyvolána tím, že důraz na prevenci snižuje podíl spamu nebo naopak tím, že nízký podíl spamu budí dojem, že prevence je účinná, však nelze rozhodnout.

Blokování

Blok nazvaný *blokování* měl zjistit přístup respondentů k eliminaci podezřelých zpráv podle různých kritérií. Blokování jako aktivní opatření již zdaleka není tak široce rozšířenou záležitostí jako prevence.

Blokování podle klíčových slov („Filtrujete e-maily podle klíčových slov?“) využívá pouhých 269 (39 %) respondentů, 419 nikoliv. Whitelist, seznam uživatelů, jejichž zprávy jsou bez problému přijaty, je poměrně extrémní metoda, využívá ji („Blokujete poštu od jiných než povolených uživatelů?“) pouze 46 (6,7 %) respondentů, 632 ji nevyužívá. Relativně rozšířené je naopak využívání blacklistů, seznamů zakázaných uživatelů („Blokujete poštu od konkrétních uživatelů/z konkrétních serverů?“). Vlastní blacklist si vytváří 346 (50 %) respondentů, některý z veřejných používá 39 (5,6 %) dotazovaných. Jako nejčastěji využívané veřejné blacklisty byly uváděny zejména tyto: Open Relay DataBase (ORDB), Mail Abuse Prevention System (MAPS), SpamCop a BrightMail (většinou zprostředkovaně, například na Yahoo!Mail nebo Hotmail). Uživatelé také často využívají služeb více blacklistů zároveň. Blacklisty nepoužívá 316 (46 %) respondentů.

Respondenti byli k blokování jako nástroji na obranu proti spamu poměrně skeptičtí – na otázku „Považujete blokování spamu za účinný prostředek?“ ohodnotili účinnost známkou 3,07 (nevím; směrodatná odchylka 1,12, četnosti jednotlivých odpovědí [31,228,104,279,42]).

Výraznější víru v účinnost blokování projevili ti, kteří se jím zabývají, tedy postmasteři (postmaster 2,64, IT manager 3,01, manager 3,09, uživatel 3,15, jiné pozice 3,15). Minimální rozdíl byl však mezi těmi, kteří některý ze způsobů blokování používají (2,98) a těmi, kteří příchozí poštu vůbec neblokují (3,09).

Software

Účinnost a využívání specializovaného software na obranu proti spamu sledoval blok nazvaný *software*. Ani software není příliš často využívaným prostředkem pro obranu proti spamu, neboť 465 respondentů uvedlo, že žádný software nepoužívá.

Nejčastějším modelem nasazení software je nasazení na straně serveru (142 respondentů). Na straně klienta používá software 83 respondentů, online služby využívá 28 dotazovaných. Jako nejčastěji využívané aplikace byly uživateli uvedeny Spam Assassin (69 respondentů) a Spam Killer (24), poměrně značné množství respondentů uvedlo, že využívají vlastní software (11). Častěji byly zmiňovány také samoučící filtr klienta Mozilla, Procmail, Postfix, CloudMark SpamNet, Kerio Mail Server a další – tedy často nikoliv specializovaný software na obranu proti spamu, ale spíše mailové servery nebo klienty s vestavěnými antispamovými funkcemi.

Ačkoliv je úspěšnost software („Považujete software za dostatečně účinný prostředek?“) respondenty z porovnávaných prostředků obrany hodnocena nejvýše, není nijak zvlášť vysoká – průměrné hodnocení je 2,73 (nevím; směrodatná odchylka 1,32, četnosti jednotlivých odpovědí [27,166,215,174,39]). Nejvyšší důvěru v software vkládají postmasteři a manageři (postmaster 2,40, IT manager 2,90, manager 2,52, uživatel 2,77, jiná pozice 2,73). Větší důvěru v software také vkládají ti, kteří některou z podob používají (2,59), zatímco ti, kteří software nepoužívají, hodnotí jeho účinnost na pouhých 2,87.

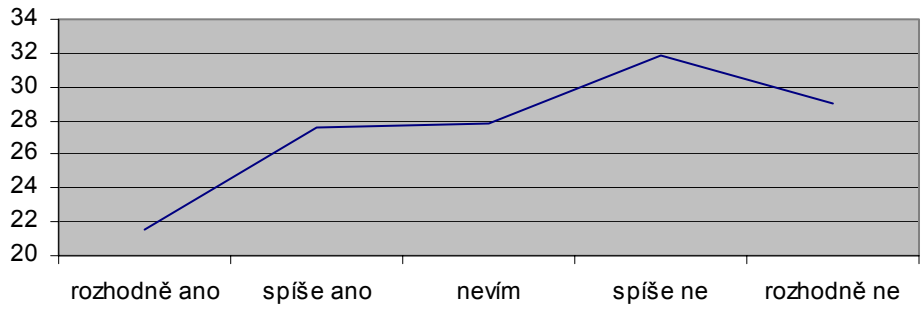
Přímý kontakt

V bloku nazvaném *přímý kontakt* odpovídali respondenti na otázky týkající se stížností přímo u spammera nebo u jeho poskytovatele připojení k internetu (ISP) nebo správce poštovního serveru. Zkoumána byla aktivita, úspěšnost a vnímaná úspěšnost tohoto počínání.

Přímo spammera se pokoušela kontaktovat („Snažil jste se někdy kontaktovat spammera?“) o něco více než polovina respondentů (345, 50 %), z toho 109 (32 %) úspěšně a 236 neúspěšně. Kontaktovat spammera se nikdy nepokoušelo 340 respondentů. Na ISP spammera se obracelo („Snažil jste se někdy kontaktovat poskytovatele připojení/správce pošty serveru, odkud vám přišel spam?“) menší množství respondentů – 199 (29 %), z toho 91 (45 %) úspěšně, 108 neúspěšně. Nikdy ISP spammera nekontaktovalo 488 dotazovaných.

Účinnost přímého kontaktu zúčastněných stran („Myslíte si, že má smysl kontaktovat spammera/jeho ISP?“) hodnotí dotazovaní průměrnou známkou 3,35 (nevím; směrodatná odchylka 1,24, četnosti jednotlivých odpovědí [53,119,99,312,97]). Z jednotlivých skupin jsou nejoptimističtější IT manageři a postmasteři, naopak znatelný pesimismus panuje zejména mezi uživateli – postmaster 3,09, IT manager 3,04, manager 3,27, uživatel 3,52, jiná pozice 3,23. Rozdíly jsou zřetelné také při rozdělení podle aktivity a úspěšnosti. Respondenti, kteří byli alespoň v jednom případě úspěšní, ohodnotili účinnost kontaktu průměrnou známkou 2,38, neúspěšní naopak pouze 3,34; ti, kteří se o kontakt nepokoušeli vůbec jsou ještě pesimističtější – 3,56.

Mezi postojem k účinnosti přímého kontaktu a podílem přijímaného spamu byla vypozorována poměrně silná závislost (korelační koeficient 0,80). Podíl spamu u respondentů, kteří věří přímému kontaktu byl nižší o šest a více procent než u těch, kteří nepovažují kontakt za účinnou metodu. (viz obrázek 4).



Obrázek 4 – Vztah mezi vnímanou účinností přímého kontaktu a podílem přijímaného spamu

Právo

Poslední blok nazvaný *právo* měl za úkol zkoumat možnosti obrany proti spammingu právní cestou. Zkušenosti respondentů s obranou právní cestou jsou minimální – pouze 25 (3,6%) respondentů má z právní obranou zkušenosti („Snažil jste se někdy bojovat proti spammerovi právní cestou?“), z toho 11 (44 %) bylo úspěšných a 14 neúspěšných. O právní obranu se nikdy nepokusilo 662 respondentů.

Právo je také hodnoceno jako nejméně účinný způsob boje proti spamu („Myslíte si, že je současné právo dostatečně účinné v boji proti spamu?“). Průměrná známka je 3,85 (spíše ne; směrodatná odchylka 1,03, četnosti jednotlivých odpovědí [11,41,167,258,209]). Důvěra v účinnost je mezi jednotlivé segmenty rozložena rovnoměrně – postmaster 3,81, IT manager 3,75, manager 3,88, uživatel 3,90 a jiná pozice 3,80. Vzhledem k malému počtu kladných odpovědí má průměrné hodnocení skupin podle aktivity a úspěšnosti poměrně nízkou vypovídací schopnost, rozdíly jsou však velké – úspěšní hodnotili účinnost známkou 2,73, neúspěšní 3,86 a ti, kteří se o právní obranu nikdy nepokoušeli, známkou 3,90. Mezi podílem přijímaného spamu a vnímanou účinností právní obrany nebyla zpozorována závislost (zejména vzhledem k malému podílu respondentů se zkušenostmi s právními protiakcemi).

Shrnutí

- Průměrný uživatel spravuje zhruba dvě nebo tři schránky, do každé z nich denně přijme zhruba jednu zprávu. Variabilita těchto hodnot je však velmi vysoká.
- Necelou třetinu zpráv tvoří spamy.
- Pro spammery je nejdůležitějším zdrojem adres web, zejména sbírání zveřejněných adres na webových stránkách a získávání adres pomocí formulářů.
- Respondenti poměrně jednoznačně považují spam za vážný problém.
- Zhruba dvě třetiny respondentů jsou preventivně poučeny o možnostech vyhýbání a zacházení se spamem. Přesto je postoj k účinnosti prevence jako metody obrany proti spamu pouze neutrální.
- Dotazovaní blokují zprávy zejména podle konkrétních uživatelů, kteří jim spamy zasílají, a podle klíčových slov. Žádný z těchto způsobů však nevyužívá více než polovina. Účinnost této metody je hodnocena neutrálně.
- Specializovaný software pro obranu proti spamu používá pouze třetina respondentů, nejčastěji na serveru. Software je považován za nejúčinnější metodu, přesto je jeho účinnost hodnocena jen mírně pozitivně.
- Zhruba polovina dotazovaných se někdy pokusila kontaktovat spammera, zhruba jedna třetina z nich byla úspěšná. Na ISP spammera se obracely necelá třetina, úspěšná však byla skoro polovina z nich. Kontakt je hodnocen mírně negativně.
- Zkušenosti s obranou pomocí právního systému jsou minimální. Respondenti považují účinnost právních protiopatření za nízkou.
- Přestože se dotazovaní cítí být spamem velmi obtěžováni, žádná z metod obrany si nezískala větší důvěru.

Použitá literatura

[1] DENÍKY BOHEMIA. Internet v prosinci navštívil rekordní počet uživatelů. 14.1.2003. Citováno 27.4.2003.

http://www.mojenoviny.cz/zpravy_z_cr/ekonomika/net030114.html

[2] WARD, Brian. Spam Statistics. Citováno 27.4.2003.

http://www.o--o.net/spam/index.php?exp=on

Kontaktní informace

Pavel Koběrský

xkobp03@vse.cz

http://diplomka.kobersky.com