

*spam, poczta elektroniczna, reklama, marketing,  
zwalczanie spamu, filtry anti-spamowe, spójność wiadomości*

Marek KOPEL\*

## **IDENTYFIKACJA SPAMU NA PODSTAWIE ANALIZY SPÓJNOŚCI WIADOMOŚCI**

Spam, czyli nie zamawiana wiadomość, to jedno z najbardziej uciążliwych zjawisk w dzisiejszym, online'owym świecie. Straty spowodowane wysyłaniem spamu liczone są w bilionach dolarów rocznie. W pracy omówione zostały współczesne formy spamowania oraz sposoby walki z nimi. Zaprezentowana została również metoda identyfikacji spamu na podstawie analizy spójności wiadomości. Spójność wiadomości wyznaczana jest na podstawie wektorów spójności nagłówek i treści wiadomości. Spójność nagłówek dwóch wiadomości obliczana jest ze względu na największe podobieństwo zawartości nagłówek. Do wyznaczania spójności treści została zaadoptowana metoda wyznaczania spójności dokumentów sieci WWW. Wybór tej metody wynika z założenia, że wiadomości zawierające odsyłacze hipertekstowe można potraktować jako szczególny przypadek dokumentów hipertekstowych.

### **1. WSTĘP**

Czasy kiedy Internet był słowem niezrozumiałym przez większość społeczeństwa, a dostęp do niego mieli jedynie „wybrańcy”, minęły bezpowrotnie. Obecnie Internet jest „słowem-wytrychem” dla wszelkiego rodzaju komercji. Wystarczy powiedzieć, że produkt jest „internetowy”, albo „łączy się z Internetem”, albo „jest dostępny za pośrednictwem Internetu” i od razu nabiera on wartości w oczach konsumentów. Nie można jednak powiedzieć, że przez to, określenie „internetowy” zostało przewartościowane, tak jak stało się to np. z określeniem „promocyjny”, które obecnie, dzięki ofertom hipermarketów, kojarzy się z tandetą. Słowo „internetowy” pewnie nigdy się nie przewartościuje, ponieważ musiałby zostać przewartościowany sam Internet.

Wraz z powszechnym dostępem to globalnej pajęczyny, stała się ona komercyjną inwestycją, medium porównywalnym, jeśli nie większym niż telewizja. Między telewizją i Internetem istnieje jednak zasadnicza różnica: w sieci każdy może być nie

---

\* Politechnika Wroclawska, Wyb. S. Wyspiańskiego 27, 50-370 Wrocław, [kopel@zsi.pwr.wroc.pl](mailto:kopel@zsi.pwr.wroc.pl)

tylko odbiorcą, ale i twórcą. Stąd wraz z wszystkimi pozytywami tego faktu, pojawiły się również problemy: wirusy, hacking, cracking, piractwo i, chyba najbardziej uciążliwy, spam.

## 2. SPAM

### 2.1. CZYM JEST SPAM?

Przyjmując podejście chronologiczne: od trzeciej dekady zeszłego stulecia, SPAM to zastrzeżony znak towarowy Hormel Foods Corporation - producenta puszkowanego mięsa [1][14]. Duży wpływ na nadanie temu terminowi negatywnego zabarwienia miał skecz Monty Python'a, w którym słowo spam pojawiało się tak często, że miało doprowadzać do irytacji. Jednak dzisiejsze znaczenie słowa spam miało swoją genezę w 1978 roku wraz z pierwszą wiadomością wysłaną w sieci Arpanet<sup>1</sup> do dużej ilości użytkowników, zawierającej reklamę firmy DEC. Od tego momentu słowo spam zaczęło używać do określania zjawiska coraz częściej zauważalnego w sieci Internet.

Spam to słowo używane obecnie tak często, że chyba każdy, kto się z nim zetknął, intuicyjnie rozumie jego znaczenie. Różne źródła podają różne definicje spamu. Najogólniej można powiedzieć, że spam to każda niezamawiana wiadomość. Stąd innymi nazwami określającymi spam są UBE (ang. *Unsolicited Bulk E-mail lub Unsolicited Broadcast E-mail*<sup>2</sup>) lub - ograniczająca się do reklam komercyjnych – UCE (ang. *Unsolicited Commercial E-mail*<sup>3</sup>). Nazwy te są rzadziej używane, ponieważ to właśnie interpretacja określenia „niezamawiana” powoduje najwięcej kontrowersji i rozbieżności dotyczących definicji spamu. Większość definicji „wymaga” również od spamu, aby była to wiadomość wysyłana do dużej liczby odbiorców. Trafianie do dużej liczby odbiorców jest typową cechą marketingu, dlatego często jego działania w Internecie określane są mianem *spamvertise* (od ang. *advertise via spam*<sup>4</sup>) [13].

Za spam najczęściej uznaje się: wszystkie wiadomości e-mail i posty grup dyskusyjnych, zawierające komercyjne i niekomercyjne reklamy, ogłoszenia i inne tzw. „zapychacze” skrzynek pocztowych. Niektóre źródła nazywają spamem legalne reklamy, które użytkownik akceptuje w zamian za darmowe używanie usług. Najlepszym przykładem takich reklam są e-maile wysyłane przez właścicieli serwerów darmowych kont pocztowych. W zamian za darmowe używanie usługi

---

<sup>1</sup> ang. *Advanced Research Projects Agency Network* – projekt rozwijany w latach 60-ych i 70-tych przez Departament Obrony Stanów Zjednoczonych, sieć będąca prekursorem Internetu

<sup>2</sup> ang. *Unsolicited Bulk E-mail lub Unsolicited Broadcast E-mail* – niezamawiany e-mail rozsyłany masowo

<sup>3</sup> ang. *Unsolicited Commercial E-mail* – niezamawiany e-mail reklamowy

<sup>4</sup> ang. *advertise via spam* - ogłaszanie przez spam

poczty użytkownik zgadza się na otrzymywanie wiadomości z reklamami wykupionymi u właściciela usługi.

Najczęściej przez spam rozumie się również tzw. *popup*'y<sup>5</sup>, czyli „wyskakujące” okienka towarzyszące przeglądaniem stron WWW. Analogicznie również bannery reklamowe na stronach WWW traktowane są przez użytkowników jako spam, ponieważ są to obrazki, których użytkownik nie chce oglądać, a jednak ładując wyszukaną stronę banner również się ściąga zajmując część łącza użytkownika.

Spamem można nazywać również tzw. *fake*'i<sup>6</sup>, czyli fałszywe pliki udostępniane głównie w sieciach P2P<sup>7</sup> udające pliki najczęściej wyszukiwane. Pliki takie najczęściej mają podobną nazwę i rozmiar do wyszukiwanych, ale zupełnie nierelevantną treść.

Kolejnym zjawiskiem podciągającym pod nazwę spam są reklamy wyświetlane przez oprogramowanie typu *adware*. *Adware* opiera się na podobnej zasadzie co darmowe skrzynki pocztowe: oprogramowanie jest udostępniane za darmo, ale w zamian w czasie używania programu wyświetlane są bannery reklamowe. Niestety najczęściej oprogramowanie *adware* jest jednocześnie *spyware*, czyli zawiera narzędzia szpiegujące użytkownika. Szpiegowanie polega na monitorowaniu działań użytkownika oraz udostępniające producentom oprogramowania prywatnych danych użytkownika. Takie działanie ma najczęściej na celu stworzenie profilu komercyjnego użytkownika.

Wraz z coraz powszechniejszym używaniem przez internautów komunikatorów internetowych spamery, czyli ludzie wysyłający spam, postanowili również ten obszar Internetu zaadaptować dla swojej działalności. Pojawiające się coraz częściej w komunikatorach ogłoszenia od nieznanymi nadawców zyskały już własną nazwę: *spim* (od złożenia „spam” i „IM”<sup>8</sup>). Oczywiście również legalny spam znalazł swoje miejsce w komunikatorach internetowych. Podobnie jak ma to miejsce na komercyjnych serwerach poczty, tak i na komercyjnych serwerach wiadomości błyskawicznych na konta IM przychodzą wiadomości reklamowe rozsyłane przez właścicieli serwerów. Dodatkowo, komercyjne komunikatory internetowe są najczęściej na licencji *adware*.

## 2.2. SKĄD SIĘ BIERZE SPAM?

Spam jest nielegalny. Przynajmniej spam wg definicji nie obejmującej: „reklama za usługę” i *adware*. Spamerzy są, w świetle prawa, przestępcami. Dlaczego? Ponieważ spam powoduje straty poprzez dodatkowe koszty: koszty łącza, w którym

<sup>5</sup> ang. *popup* – wyskakiwać

<sup>6</sup> ang. *fake* - udawać

<sup>7</sup> P2P – ang. *Peer to Peer*, architektura, w której każdy komputer jest jednocześnie serwerem i klientem

<sup>8</sup> IM – ang. *Instant Messaging* – wiadomości błyskawiczne wysyłane za pomocą komunikatorów internetowych

generowany jest niepotrzebny ruch, koszty sprzętu, który przechowuje niepożądane wiadomości i koszty czasu użytkownika, który aby dotrzeć do prawdziwych wiadomości w swojej skrzynce musi przebrnąć przez gąszcz spamu.

Aby móc rozsyłać reklamy legalnie, nadawca musi mieć prawnie udokumentowane źródło adresatów. Adresy najczęściej zbierane są w sposób "chcesz usługę za darmo? Zarejestruj się i podaj nam swój adres". Czasami jest to robione jawnie i użytkownik jest informowany, że musi zgodzić się na przetwarzanie swoich danych osobowych (w tym adresu e-mail) w celach marketingowych. Jednak w większości przypadków użytkownik, rejestrując się, nie jest świadomy, że właśnie dołącza swój adres do bazy spamerów.

Innym sposobem pozyskiwania adresów jest skanowanie sieci przez tzw. *address-scrapers*, czyli programy wydobywające adresy ze stron WWW. Programy te przeglądają kod HTML kolejnych stron WWW w poszukiwaniu znaku @, i po znalezieniu go dołączają odpowiedni fragment tekstu przed i po tym znaku do bazy. Legalność zdobytej w ten sposób bazy adresów jest kwestią sporną, jednak twórcy tego pomysłu twierdzą, że skoro ktoś umieszcza adres e-mail na stronie WWW, to znaczy, że chce aby każdy mógł wysłać do niego e-mail.

Aby nie musieć zbierać adresów spamery często obierają za cel grupy dyskusyjne (tzw. *usenet*) lub grupy mailowe (tzw. *opt-in*). O ile grupami dyskusyjnymi nie ma zbyt dużego problemu, ponieważ wszystkie wiadomości usenetowe są przechowywane na wspólnym serwerze, to grupy mailowe rozsyłające e-maile do wszystkich członków *opt-in*, zaatakowane spamem, mogą szybko zapchać skrzynkę pocztową.

Obecnie spam w wiadomościach to nie tylko reklamy. Handlowanie bazami adresów e-mail stało się na tyle opłacalne, że spam stosuje się do pozyskania lub potwierdzenia aktywności adresu e-mail w celu stworzenia własnej bazy adresowej. W ten sposób spam zaczął działać na zasadzie *perpetuum mobile*.

Innym, niekomercyjnym zastosowaniem spamu są samorozsyłające się wirusy, czy robaki internetowe, podszywające się pod zwykłe maile. Użytkownik otwierając wiadomość z wirusem zaraża swój komputer, dzięki czemu wirus może rozesłać się na wszystkie adresy z książki adresowej użytkownika. Warto nadmienić, że sam wirus lub robak również może służyć tworzeniu bazy adresowej.

Jeszcze inny zastosowaniem spamu jest promowanie stron WWW w indeksach wyszukiwarek internetowych. Jak wiadomo kolejność stron zwracanych w odpowiedzi przez wyszukiwarkę zależy od relewancji strony. Relewancja strony mierzona jest specjalnym wskaźnikiem. W przypadku Google, wskaźnikiem takim jest PageRank. Wartość PageRank dla strony mierzy się m.in. liczbą odwiedzin strony. Jeżeli wiadomość ze spamem przekierowuje użytkownika na daną stronę, to PageRank dla tej strony wzrasta. Korzyści płynące z pozycjonowania, czyli ustalania kolejności stron WWW w odpowiedzi wyszukiwarki wykraczają poza tematykę tego referatu, wystarczy jednak powiedzieć, że warto dla nich rozsyłać spam.

Ciekawym zastosowaniem spamu jest umieszczanie w sieciach P2P fake'ów. Po co udostępniać w P2P takie pliki? Za odpowiedź niech posłuży autentyczna historia sprzed 2 lat: Dwóch polskich muzyków, fanów Georga Michaela, nagrało własny utwór wzorując się na swoim idolu. Nagranie to było na tyle profesjonalne, że gdy ktoś udostępnił je w sieci P2P w postaci pliku mp3 z opisem „nowy singiel Georga Michaela”, fani wykonawcy z całego świata zaczęli ściągać ten plik. Głos na nagraniu był na tyle podobny, że słuchający wierzyli, że rzeczywiście słuchają Georga Michaela. Nagranie to bardzo szybko pojawiło się na pirackiej płycie CD, z wcześniejszymi utworami Michaela, jako utwór bonusowy. Całe zamieszanie zmusiło menadżera Michaela do oficjalnego oświadczenia na stronie WWW artysty, że wspomniany utwór nie został napisany, ani nagrany przez piosenkarza. Twórcy nagrania, z kolei, dzięki jednej piosence zostali wypromowani w sposób, który normalnie wiązałby się z kosztami przekraczającymi możliwości muzyków.

Zupełnie odwrotny cel stawiają sobie ludzie umieszczający w sieciach P2P fake'i pirackich kopii najnowszych filmów i nagrań. Celem jest tutaj zaśmianie czy “zaszumianie” niekomercyjnych sieci P2P, aby obniżyć ich efektywności i zniechęcić przyszłych *swapperów*<sup>9</sup>. Sami autorzy pomysłu nazywają to walką z piractwem, ale w rzeczywistości jest to, delikatnie mówiąc, promowanie systemów legalnej sprzedaży multimediiów w sieci, jakim jest np. iTunes.

### 2.3. JAK WALCZYĆ ZE SPAMEM?

Walkę ze spamem można prowadzić na trzech etapach jego rozsyłania: u źródła, przy przekazywaniu między serwerami i po stronie odbiorcy.

Pierwszym rozwiązaniem, jakie nasuwa się w walce ze spamem, jest namierzenie spamerów i uniemożliwienie ich dalszej działalności. Jednak, ponieważ spam jest nielegalny, spamerzy starają się pozostać w ukryciu, a dopóki im się to udaje pozostają bezkarni. Głównym dowodem winy spamera są nagłówki wiadomości zawierające adres zwrotny, oraz adresy serwerów przez które wiadomości została wysłana [4]. Jednak współczesne techniki spamowania bardzo dobrze radzą sobie z fałszowaniem tych danych. Wpisanie fałszywych danych nadawcy i adresu zwrotnego to nie problem. Nikt tego nie sprawdza. Większym problemem jest adres IP komputera i konto na serwerze z którego wysłano wiadomość. Wszystko to jest “do obejścia” dzięki tzw. serwerom *open relay* [11]. Serwery *open relay* to niezabezpieczone serwery pocztowe, dzięki którym każdy z dowolnego komputera w sieci może wysłać wiadomości do dowolnego adresata. Niestety serwerów takich ze względu na nieumiejętne administrowanie, jest bardzo dużo w sieci. Listy adresów IP tych serwerów można łatwo znaleźć w witrynach sprawdzających zabezpieczenia poczty lub namierzyć poprzez skanowanie portów [9][10].

---

<sup>9</sup> swapper – osoba wymieniająca pliki w sieciach P2P

Innym sposobem zatarcia śladów przez spamera jest użycie anonimowego serwera *proxy*. Serwer taki jest miejscem, w którym ślad po wykorzystującym do użytkownika się urywa. Wszystkie działania w sieci, użytkownika korzystającego z proxy, odbywają się przez ten serwer, który jest jednocześnie *firewall*'em dla połączeń do komputera użytkownika. Listy IP takich serwerów również są dostępne w WWW. A ich utrzymaniem zajmują się ludzie walczący o zachowanie anonimowości i prywatności w sieci.

Jeszcze inną techniką utrudniającą namierzenie spamatorów jest używanie połączeń *dial-up*. Połączenia te ze względu na swoją specyfikę pozwalają użytkownikowi przy każdym połączeniu uzyskać inny adres IP, przez co nie jest on już dowodem winy, jak w przypadku stałego łącza.

Ponieważ nie jesteśmy w stanie zabezpieczyć wszystkich serwerów open relay, proxy i prowajderów *dial-up*, istnieją systemy tworzące tzw. *blacklisty*, czyli adresy serwerów open relay i innych źródeł spamu, do których można zgłaszać podejrzone adresy. Przykładem takiego systemu jest SORBS [12]. Poza zgłaszaniem adresów, często stosuje się stawianie pułapek (ang. *honeypot*) [5] na spamatorów poszukujących serwerów, z których można by wysłać spam. Pułapki takie, to komputery działające jak serwery pocztowe, wykrywalne tylko poprzez skanowanie portów, a więc taktyki używane przez spamatorów. Serwery te przyjmują wiadomości, ale nie przekazują ich dalej. Spamer chcąc przetestować znaleziony serwer wysyła przez niego wiadomość i tym samym wpada w pułapkę.

Główna walka ze spamem u źródła ogranicza się jednak głównie do chronienia swojego adresu e-mail przed dostaniem się do bazy adresów spamatorów. Należy pamiętać, że jeżeli adres znajdzie się w jednej bazie, to bardzo szybko zostaje odsprzedany czy wymieniony za inne i trafia dużej liczby baz spamatorów. Jak chronić swój adres, skoro z jednej strony ma on być publicznie dostępny, a z drugiej jego opublikowanie nie oznacza, że zgadzamy się na otrzymywanie spamu?

Bazy adresów e-mail są tworzone głównie przez automaty i przechowywane w postaci elektronicznej, dlatego nie należy się martwić o umieszczeniu adresu na wizytówkach, czy papierze firmowym. Oczywiście zawsze może się zdarzyć, że ktoś ręcznie umieści taki adres w bazie, ale są to sytuacje wyjątkowe. Istotne jest aby nie nadużywać adresu w postaci elektronicznej, czyli nie podawać go, przy rejestracji usług, do których nie mamy zaufania, nie rozsyłać e-maili z jawną listą adresatów i najważniejsze zabezpieczyć adres przez robotami skanującymi WWW.

Techniki zabezpieczania publikowanego adresu przed address-scrap'er'ami są różne [15]. Najbardziej popularna polega na dopisywaniu lub zamienianiu tekstu w adresie np. `missi_usuń_2004#pwr(kropka)wroc(kropka)pl`. W ten sposób człowiek zrozumie, usunie i zamieni w adresie odpowiednie znaki otrzymując poprawny adres, natomiast robot w ogóle nie znajdzie tego adresu. Inna metodą maskowania adresu przez robotami to kodowanie go np. numerami encji ASCII. W HTMLu adres wyglądać będzie np. tak: `&#109;&#105;&#115;&#115;&#105;&#50;&#48;&#48;&#52;`

„64;112;119;114;46;119;114;111;99;46;112;108;”, a dopiero przy wyświetlaniu przeglądarka zamienia go na odczytywalny adres: missi2004@pwr.wroc.pl. Roboty jednak również ewoluują i są coraz skuteczniejsze w znajdowaniu adresu, dlatego najbezpieczniejsze obecnie wydaje się publikowanie adresu jako obrazka w formacie JPG czy GIF. Choć znane są już metody OCR, które adres nawet w tym formacie mogłyby przechwycić

Kolejnym etapem życia spamu, na którym można go zwalczać jest przekazywanie wiadomości pomiędzy kolejnymi serwerami pocztowymi. Najczęstszym rozwiązaniem są filtry oparte na regułach (np. „Jeśli w temat zawiera słowo ‘promocja’, to odrzuć wiadomość”) zintegrowane ze współtworzoną bazą blacklist. Najpopularniejszym rozwiązaniem tego typu jest SpamAssassin [7].

Nowszymi rozwiązaniami są metody obliczające sygnatury wiadomości, które ignorują niewielkie modyfikacje, wychwytyując kwintesencję treści. W ten sposób działają DCC [6] i Vipul's Razor [8]. Istnieją również prace angażujące do filtrowania spamu sieci Bayesa [2].

Największym polem walki ze spamem jest obecnie skrzynka odbiorcza użytkownika, czyli ostatni etap życia spamu. Istnieje bardzo wiele komercyjnych produktów filtrujących spam w klientach pocztowych. Poza tym, każdy portal oferujący usługę poczty posiada własny system antyspamowy. Należy jednak pamiętać, że takie systemy nie będą działać dla legalnych reklam. Ponieważ nie należy podcinać gałęzi, na której się siedzi, trudno szukać komercyjnego produktu zwalczającego podstawę komercji. Chyba, że będzie to produkt firmy niezależnej, ale wtedy można spodziewać się nawet procesów sądowych, jak miało to miejsce przy wprowadzeniu na rynek Replay TV 4000.

Replay TV 4000 to PVR<sup>10</sup> firmy SonicBlue, który pozwala na automatyczne pomijanie reklam podczas nagrywania programów telewizyjnych. Mimo, że technologia pomijania reklam znana jest już od kilkunastu lat, nie została wcześniej zastosowana w komercyjnym produkcie. Podobnie wygląda to w przypadku spamu. Dlatego użytkownik powinien sam zadbać o swoją ochronę przed spamem.

Ponieważ większość spamu obecnie nastawiona jest na pozyskanie lub potwierdzenie używania adresu e-mail dlatego nie należy nigdy otwierać wiadomości spamowej, jeśli do jej wyświetlenia istnieje potrzeba połączenia z Internetem. Ponieważ spamerzy często generują losowo adresy lub uzyskują je z nieautoryzowanych źródeł ich wiadomości mają na celu zweryfikowanie czy wiadomość dotarła do odbiorcy. Wystarczy, aby użytkownik otworzył odpowiednio spreparowaną wiadomość w HTMLu lub kliknął zawarty w wiadomości link, a spamer wie, że adres jest używany. Podobnie ma się sprawa z wypisywaniem się z grup, do których rzekomo jesteśmy zapisani. Próba wypisania się poprzez odpowiedź na adres zwrotny lub w inny wskazany sposób pozwala na potwierdzenie aktualności

---

<sup>10</sup> ang. *Personal Video Recorder* – nazwa używana w stosunku do *DVR (Digital Video Recorder)* i *VCR (Videocassette recorder)*, czyli cyfrowych i analogowych magnetowidów

adresu. Taki adres jest dołączany do bazy potwierdzonych adresów i jego właściciel staje się w krótkim czasie adresatem każdego spamu.

Z drugiej strony unikanie interakcji ze spamem nie powinno powstrzymywać użytkownika od wysyłania skarg na otrzymany spam. Zgłaszanie spamu pozwala na aktualizowanie *blacklist* oraz uniemożliwienie konkretnej wiadomości na dalsze rozprzestrzenianie się. Poza tym skarga może być podstawą do zamknięcia konta spamera, czy innej reakcji utrudniającej mu dalsze działanie [4].

Jak wiadomo najsłabszym ogniwem systemu jest zwykle człowiek, więc wszystkie powyższe zabezpieczenia nic nie dadzą w przypadku użytkownika bez świadomości zagrożeń spamu. Jediną ochroną może być wtedy zatrzymanie spamu przez niekomercyjny system, na jednym z wcześniejszych etapów, zanim odbiorca otworzy wiadomość. W tym celu można posłużyć się prezentowaną metodą badania spójności jako jednego z kryteriów identyfikujących spam.

### 3. METODA

Metoda polega na liczeniu spójności przychodzących wiadomości z wzorcowymi wiadomościami zawierającymi spam. Spójność może być jedną z miar pozwalającą systemowi antyspamowemu automatycznie zaklasyfikować przychodzącą wiadomość jako spam. Spójność liczona jest osobno dla nagłówków wiadomości i treści wiadomości. Dla każdej pary wiadomości (wzorcowej i sprawdzanej) tworzymy dwa wektory: wektor spójności nagłówków i wektor spójności treści, na podstawie których wyznaczamy spójności nagłówków i treści oraz całkowitą spójność wiadomości..

#### 3.1. WEKTOR SPÓJNOŚCI NAGŁÓWKÓW WIADOMOŚCI

Intuicja wyznaczania spójności nagłówków wiadomości opiera się na fakcie: spamery mogą wykorzystywać ograniczoną liczbę serwerów pozwalających rozsyłać spam oraz używać tych samych narzędzi, które fałszują nagłówki w podobny sposób. Dlatego nagłówki będą tym bardziej spójne im większe będzie współwystępowanie informacji w nich zawartych. Uwzględniając tą cechę tworzymy wektor spójności nagłówków wiadomości. Elementy wektora spójności nagłówków wiadomości wyznaczamy na podstawie najdłuższych jednakowych podciągów znaków w odpowiadających sobie nagłówkach.

*Elementem wektora spójności nagłówków wiadomości nazywamy liczbę:*

$$h_i = \frac{ml}{hl}, \quad (1)$$

gdzie  $ml$  to długość podciągu jednakowego w odpowiednim nagłówku w obu wiadomościach, a  $nl$  to długość dłuższego ciągu z odpowiednich nagłówków w obu wiadomościach

Przykładowo: nagłówek *subject* w wiadomości wzorcowej ma wartość "Re: important messege", a w wiadomości sprawdzanej - " Fw: Re: important notice". Długość jednakowych podciągów  $ml=15$ , a długość dłuższego nagłówka  $hl=24$ , więc element wektora spójności nagłówków odpowiadający nagłówkowi *subject*  $h=0,62$ .

*Wektorem spójności nagłówków wiadomości* nazywamy wektor:

$$H = \langle h_1, \dots, h_k \rangle, \quad (2)$$

gdzie  $h_1, \dots, h_k$  to elementy wektora spójności nagłówków wiadomości odpowiadające kolejnym nagłówkom.

### 3.2. WEKTOR SPÓJNOŚCI TREŚCI WIADOMOŚCI

Intuicja wyznaczania spójności treści wiadomości jest następująca: wiadomości zawierające spam są najczęściej w formacie HTML, a nawet jeśli nie, to zawierają hiperlinki odsyłające do stron reklamowanych usług czy produktów. Gdyby tak nie było, efektywność takiej reklamy byłaby niewielka. Zakładając więc, że każda wiadomość zawiera hiperlinki, możemy potraktować wiadomości jako dokumenty sieci WWW, do których nie prowadzą żadne hiperlinki, ale które odsyłają nas do innych dokumentów WWW. Przyjmując takie założenie, do liczenia spójności treści wiadomości możemy wykorzystać zmodyfikowaną wersję wyznaczania spójności dokumentów WWW na podstawie powiązań hiperlinkami, zaprezentowaną w [3].

Metoda polega na tworzeniu wektora spójności na podstawie ilości powtarzających się terminów w otoczeniu hiperlinka (czyli we fragmencie tekstu zawierającym odnośnik) i dokumencie docelowym (do którego hiperlink odsyła). Wiadomości nie mogą mieć bezpośrednich powiązań hiperlinkami. Jednak często zdarza się, że wiadomości zawierają hiperlinki odsyłające do tych samych adresów WWW. Dlatego modyfikacja metody będzie polegała na przyjęciu za podstawę liczenia powtarzających się terminów: otoczenia hiperlinków z wiadomości przychodzącej i otoczenia hiperlinków z wiadomości wzorcowej.

*Wektorem spójności treści wiadomości* nazywamy wektor:

$$B = \langle b_1, \dots, b_l \rangle, \quad (3)$$

gdzie  $b_i$  to suma liczby powtórzeń  $i$ -tego terminu w otoczeniach każdego hiperlinka w wiadomości przychodzącej z każdym hiperlinkiem wiadomości wzorcowej.

### 3.3. SPÓJNOŚĆ WIADOMOŚCI

Spójność nagłówków liczymy jako średnią ważoną elementów wektora spójności nagłówków wiadomości. Wagi przypisane kolejnym nagłówkom wyznaczone są empirycznie. Początkowo można przypisać im wartość 1, a następnie modyfikować.

*Spójnością nagłówków wiadomości* nazywamy liczbę:

$$hc = \frac{\sum_{i=1}^K h_i * w_i}{\sum_{i=1}^K w_i}, \quad (4)$$

gdzie  $w_i$  to waga  $i$ -tego nagłówka.

*Spójnością treści wiadomości* (zgodnie z [3]) nazywamy liczbę:

$$bc = \frac{1}{L-1} \sum_{i=1}^{L-1} \sum_{j>i}^L |b_i - b_j|. \quad (5)$$

Ostatecznie spójność wiadomości wyznacza się na podstawie spójności nagłówków i spójności treści regulując ich wpływ poprzez odpowiedni dobór parametrów.

*Spójnością wiadomości* nazywamy liczbę:

$$mc = x * hc + y * bc, \quad (6)$$

gdzie  $x, y \in \langle 0,1 \rangle$  oraz  $x+y=1$ . Wartości parametrów  $x$  i  $y$  można wyznaczyć empirycznie. Najczęściej wartość  $x$  nie powinna być większa od wartości  $y$ .

## 4. PODSUMOWANIE

Filtrowanie wiadomości wyłącznie za pomocą funkcji spójności może być mało wydajne, dlatego założenie jest takie, żeby używać miary spójności jako dodatkowego wskaźnika w systemie antyspamowym. Metoda analizy spójności może być użyta zarówno w rozwiązaniach działających na komputerach odbiorców jak i na serwerach pocztowych. Oczywiście efektywniejsze wydaje się stosowanie filtrów na serwerach, ponieważ zaoszczędza się na kosztach przesyłania i przechowania wiadomości spamowych. Jednak wtedy użytkownik nie miałby możliwości weryfikacji działania filtrów, a ryzyko tzw. *false positives*, czyli zwykłych wiadomości niewłaściwie

uznanych za spam, może okazać się zbyt duże, aby można je było podjąć. Zbadanie tego problemu może posłużyć za punkt wyjściowy do dalszej pracy w dziedzinie zwalczania spamu.

#### LITERATURA

- [1] FAHEY, C. P., *Spam: The Phenomenon*, [www.colinfahey.com](http://www.colinfahey.com), 2004
- [2] GRAHAM, P., *A Plan for Spam*, [www.paulgraham.com](http://www.paulgraham.com), 2002
- [3] KOPEL, M., DANIŁOWICZ, CZ., *Method of Completing the Consistency Graph of a Hyperlinked Document Collection*. W: Intelligent Technologies for Inconsistent Knowledge Processing. Advanced Knowledge International, Nguyen T. (red.), Adelaide, South Australia, 2004, 145-162.
- [4] KRAWCZYK, P., *Analiza nagłówek pocztowych*, [arch.ipsec.pl/prez/openmail-2001](http://arch.ipsec.pl/prez/openmail-2001), 2001
- [5] SPITZNER, L., *Definitions and Value of Honey pots*, [www.tracking-hackers.com](http://www.tracking-hackers.com), 2003  
*Filtry antyspamowe*
- [6] DCC - Distributed Checksum Clearinghouse, [www.rhyolite.com/anti-spam/dcc](http://www.rhyolite.com/anti-spam/dcc)
- [7] SpamAssassin, [spamassassin.org](http://spamassassin.org)
- [8] Vipul's Razor, [razor.sourceforge.net](http://razor.sourceforge.net)  
*Open Relay*
- [9] [www.3dmail.com/spam/](http://www.3dmail.com/spam/)
- [10] [www.abuse.net/relay.html](http://www.abuse.net/relay.html)
- [11] [www.ordb.org](http://www.ordb.org)
- [12] Spam and Open Relay Blocking System, [www.dnsbl.nl.sorbs.net](http://www.dnsbl.nl.sorbs.net)  
*Witryny WWW*
- [13] Fight Spam on the Internet!, [spam.abuse.net](http://spam.abuse.net)
- [14] Hormel Foods Corporation, Official SPAM Home Page, [www.spam.com](http://www.spam.com)
- [15] *Obrona przed spamem. Druga linia obrony - zabezpieczenie adresu mailowego*, [nosspam-pl.net/obrona2.php](http://nosspam-pl.net/obrona2.php), 2004

#### SPAM IDENTIFICATION BASED ON MESSAGE CONSISTENCY ANALYSIS

Spam, or Unsolicited Broadcast Mail, is today's most common Internet abuse. In this paper various types of spam and forms of fighting it are described. One of the forms of fighting spam is the presented method based on message consistency analysis. The consistency is established for each pair of messages: a template spam message and an incoming message. The consistency of message headers and body are calculated separately on the basis of headers' consistency vector and body's consistency vector. Eventually the two consistencies are used to establish the two messages' consistency.