Dr inż. Marian Wrzesień
Przemysłowy Instytut Automatyki i Pomiarów

SYSTEM POCZTOWY WYPOSAŻONY W FILTR ANTYWIRUSOWY I W FILTR ANTYSPAMOWY

Zaprezentowano implementację systemu poczty elektronicznej Postfix w systemie operacyjnym Fedora Linux. System poczty elektronicznej jest wyposażony w filtry antywirusowy i antyspamowy wbudowane w strukturę systemu pocztowego. Każdy z filtrów jest sterowany niezależnie z pomocą własnych narzędzi programowych definiowanych przez pliki konfiguracyjne. System pocztowy i filtry współpracują ze sobą przy przetwarzaniu wiadomości pocztowych tak odbieranych jak i wysyłanych. Wyszczególniono metodę integracji zastosowanych systemów, jak i działanie modułów programowych stosowanych do procesu przekazywania wiadomości pocztowych.

THE MAIL SYSTEM EQUIPPED WITH ANTIVIRUS AND ANTISPAM FILTERS

The implementation of the Postfix e-mail system based on the Fedora Linux operating system is presented. The e-mail system is equipped with antivirus and antispam filters embedded into the mail structure. Each of filters is controlled independently with the aid of its own software tools defined by configuration files. E-mail system and filters cooperate with each other while processing both received and sent mail messages. The method of integration applied systems is depicted as well as the actions of the software modules used to transfer process of the mail messages.

1. WSTĘP

Referat odnosi się do tematyki bezpieczeństwa systemów informatycznych; w szczególności prezentuje wdrożone w PIAP elementy ochrony poczty elektronicznej przed wirusami i spamem.

Przedstawiane zagadnienia mieszczą się w tematyce objętej polskimi normami dotyczącymi SZBI (System Zarządzania Bezpieczeństwem Informacji). I tak, najnowsza norma PN-ISO/IEC 27001 tej tematyki, wydana w styczniu 2007 r., wskazuje **wymagania** dla systemów zarządzania bezpieczeństwem informacji, a norma **P**N-ISO/IEC 17799, również ze stycznia 2007 r., prezentuje **praktyczne zasady wdrażania** tych systemów. Normy te obejmują:

- Techniki informatyczne oraz
- Techniki bezpieczeństwa.

Celem wdrożenia i stosowania zaleceń ww. norm jest zapewnienie podstawowych cech bezpieczeństwa przetwarzanych informacji, jakimi są:

- Poufność (**C**ONFIDENTIALITY) skutkująca ochroną przed nieautoryzowanym dostępem do informacji
- Integralność (**I**NTEGRALITY) chroniąca przed nieuprawnionym modyfikowaniem treści zawartej w informacji
- Autentyczność (**A**UTHENTICITY) uniemożliwiający podszywanie się pod autora informacji

Dzięki swej elastyczności, konstrukcja obu ww. norm PN-ISO/IEC 27001,17799 umożliwia zintegrowanie ich z normami ISO 9001:2000 oraz ISO 14001:2004 zapewniając w efekcie uzyskanie spójności procedur wspomagających zarządzanie firmą z uwzględnieniem bezpieczeństwa informacji.

W referacie zaprezentowano działanie systemów antywirusowego i antyspamowego, odpowiednio wkomponowanych w zaimplementowany wcześniej system poczty elektronicznej.

Główna ochrona przed wirusami i spamem zachodzi na etapie odbierania wiadomości w serwerze pocztowym, przed procesem przekazania tych wiadomości do skrzynek pocztowych, i przed odczytywaniem ich przez klienta pocztowego – narzędzie przetwarzające wiadomości wejściowe/wyjściowe każdego z użytkowników stacji lokalnych.

W wyniku takiego podejścia znacznie oszczędza się angażowanie serwera pocztowego odstępując od konieczności pełnego przetwarzania wiadomości zawirusowanych, bądź stanowiących spam. Również, ze względu na wcześniejsze filtrowanie przez serwer wiadomości pocztowych, czas poświęcany na analizę przeprowadzaną indywidualnie przez użytkownika ulega znacznemu skróceniu.

2. SYSTEM POCZTOWY W PIAP

Dostęp użytkowników do sieci informatycznej w PIAP jest realizowany za pomocą routera wyposażonego w system operacyjny Linux. Również do obsługi systemu pocztowego wykorzystuje się serwer funkcjonujący w tym samym systemie operacyjnym. Dzięki tej kompatybilności uzyskuje się duże możliwości konfiguracyjne w torze przekazywania informacji z WAN (Wide Area Network) do stanowisk komputerowych użytkowników końcowych w LAN (Local Area Network).

Z kilku dostępnych dzisiaj systemów pocztowych, stosowanych w różnorodnych wersjach dytrybucyjnych systemu operacyjnego Linux, a więc ***Sendmail***, ***Qmail***, ***Exim*** i ***Postfix***, w PIAP zaimplementowano ***Postfix,*** napisany przez Wietse Venema, a wcześniej znany jako ***VMailer***. O wyborze tym zadecydowały cechy tego systemu pocztowego takie jak: przejrzysta konfiguracja, proste zarządzanie, duże bezpieczeństwo, wysoka wydajność oraz konstrukcja modułowa umożliwiająca dołączanie elementów poszerzających podstawową funkcjonalność systemu pocztowego, a wśród nich m.in. filtr antywirusowy oraz filtr antyspamowy.

2.1. Funkcjonowanie systemu pocztowego Postfix

W przeciwieństwie do powszechnie stosowanego jeszcze niedawno systemu pocztowego ***Sendmail***, używającego jednego dużego programu do obsługi wiadomości pocztowych, ***Postfix*** przydziela do wykonania odrębnych zadań procesu obsługi wysyłania i odbierania poczty elektronicznej oddzielnym programom-demonom, działającym w systemie w tle. Każdy z demonów wykonuje określone zadanie, dla którego wykonywania jest stworzony.

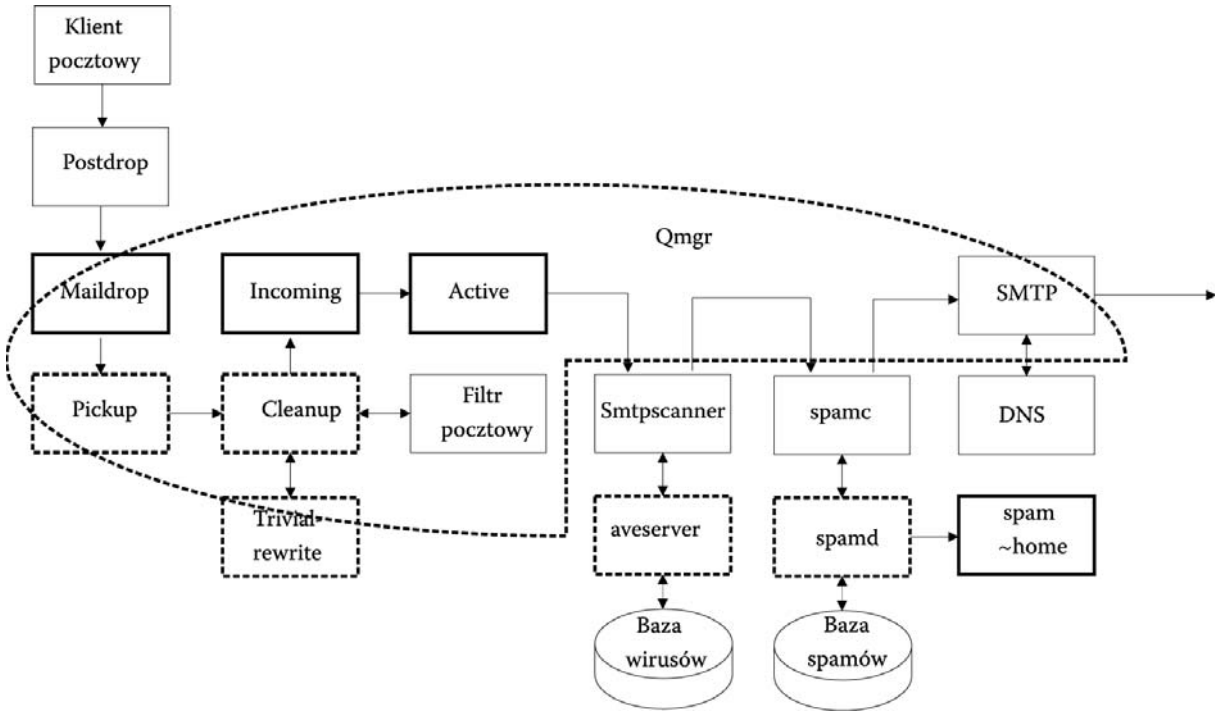
Wystartowanie systemu ***Postfix*** skutkuje uruchomieniem głównego demona ***master***, który w miarę potrzeb wywołuje większość innych procesów. Każdy z demonów wywoływanych przez ***master*** realizuje przydzielone mu zadanie, po wykonaniu którego przechodzi do stanu gotowości oczekując na kolejne polecenie.

Funkcjonalność demona ***master*** jest określone odpowiednimi zapisami zawartymi w dwóch plikach konfiguracyjnych ***main.cf*** i ***master.cf***.

Rys. 1. przedstawia ogólny schemat procesu realizowanego przez ***Postfix***. Najogólniej ujmując ***Postfix*** odbiera wiadomości pochodzące z systemu lokalnego lub zewnętrznego, kolejkuje je, a następnie doręcza - do lokalnej skrzynki pocztowej lub odległego serwera pocztowego.

2.1.1. Wysyłanie wiadomości

Wysyłanie wiadomości z serwera lokalnego na zewnątrz inicjuje lokalny klient pocztowy. Żądając wysłania wiadomości, powoduje zapisanie jej w katalogu ***maildrop.*** Wykorzystywany jest do tego program narzędziowy ***postdrop*** uruchamiany przez interfejs ***sendmail.postfix***. Katalog ***maildrop*** jest kontrolowany przez demona ***pickup.*** Po pojawieniu się informacji o nowej wiadomości w katalogu ***maildrop,*** demon ***pickup*** pobiera tę wiadomość z kolejki, sprawdza jej poprawność logiczną i przekazuje ją demonowi ***cleanup***, który finalizuje jej przetwarzanie. Gdyby klient pocztowy Nadawcy nie zawierał adresu From: lub nie użył pełnej nazwy hosta źródłowego w adresie, demon ***cleanup,*** we współpracy z demonem ***trivial-rewrite,*** wprowadziłby niezbędne poprawki w wiadomości. Po zakończeniu pracy demon ***cleanup*** umieszcza przetwarzaną wiadomość w kolejce ***incoming.*** Oczekujący na pojawienie się w tej kolejce wiadomości, demon-menedżer kolejek ***qmgr*** może teraz rozpocząć proces jej doręczania – zgodnie ze strategią trasowania wypracowaną przez demona ***trivial-rewrite*** (rys. 1).



Rys. 1. Tor przesyłu wiadomości pocztowych podczas ich wysyłania

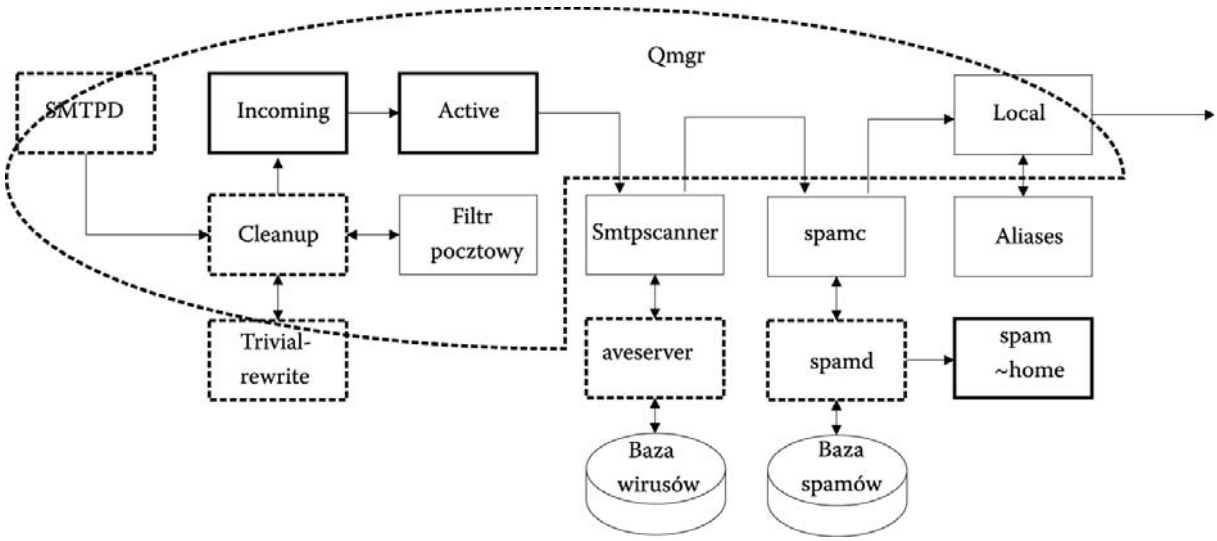
Poza kolejką ***incoming*** demon ***qmgr*** obsługuje kolejki wiadomości odpowiednio: ***active (aktywne)***, ***deferred (odroczone)***, ***corrupt (uszkodzone)*** i ***hold (zawieszone)***. Gdy nowa wiadomość z katalogu ***incoming*** jest gotowa do doręczenia, menedżer kolejek ***qmgr*** przenosi tę wiadomość do kolejki ***active***. Jeśli wiadomość jest przeznaczona dla użytkownika w systemie zewnętrznym, menedżer kolejek ***qmgr*** zgłasza żądanie doręczenia tej wiadomości agentowi ***smtp.***

Korzystając z usługi ***DNS***, agent ***smtp*** pobiera listę systemów pocztowych, które mogą przyjmować pocztę z domeny Adresata. Agent ***smtp*** wybiera najbardziej preferowanego hosta pocztowego MX z listy i kontaktuje się z nim, aby doręczyć wiadomość od Nadawcy. Agent ***smtp*** uaktualnia status Adresata zakolejkowanego pliku jako przetworzony, lub – wraz z jednym z agentów przetwarzania wiadomości, odpowiednio, ***bounce (odbite)***, ***defer (opóźnione)***, ***trace (informacja o statusie)*** – przygotowuje komunikat o stanie dostawy wiadomości, lub, jeśli doręczenie zakończone fiaskiem, o przyczynach tego stanu.

2.1.2. Odbiór wiadomości

Odbiór poczty od ***Nadawcy*** zewnętrznego jest realizowany przez demona ***smtpd***, współpracującego z agentem doręczającym ***smtp Nadawcy***. Gdy demon ***smtpd*** zweryfikuje, że powinien przyjąć tę wiadomość, przekazuje ją demonowi ***cleanup***, który – po przeprowadzeniu kontroli poprawności – zapisze wiadomość w kolejce ***incoming***. Gdy nowa wiadomość z katalogu ***incoming*** jest gotowa do doręczenia, menedżer kolejek ***qmgr*** przenosi tę wiadomość do kolejki ***active***.

Po ustaleniu przez ***qmgr,*** że wiadomość może być dalej przetwarzana, menedżer ***qmgr*** wywołuje agenta ***local***, aby sfinalizować doręczenie wiadomości. Jeśli agent doręczający ***local*** ustali, że Nazwa-Adresata jest jego aliasem, wtedy ponownie zgłasza wiadomość poprzez demona ***cleanup*** do doręczenia wiadomości pod nowy adres.



Rys. 2. Tor przesyłu wiadomości pocztowych podczas ich odbierania

Zarówno ***cleanup***, jak i menedżer kolejek ***qmgr*** wywołują przy przetwarzaniu wiadomości demona ***trivial-rewrite***. Pomaga on konwertować adresy e-mail do standardowego formatu i ustalać trasowanie wiadomości (typ transportu oraz następne miejsce na trasie doręczenia).

Gdy zamierzone jest doręczenie nowej wiadomości do innej sieci, to wtedy menedżer kolejek wywołuje agenta ***smtp***. Ten szuka w DNS-ie serwerów pocztowych, które mogą przyjąć pocztę w domenie innej sieci.

W rezultacie wiadomość jest przekazana agentowi doręczającemu ***local*** tej nowej sieci, który zapisuje ją w magazynie wiadomości w tym systemie (skrzynki pocztowe).

Jeśli z jakieś przejściowej przyczyny doręczenie wiadomości nie było możliwe, to o takim fakcie agent doręczający powiadamia menedżera kolejek, który umieszcza wiadomość w kolejce ***deferred*** i ponawia próbę doręczenia jej w późniejszym terminie.

Tu rola ***Postfixa*** się kończy. Pobranie i odczytanie wiadomości z lokalnego magazynu wiadomości może zostać przeprowadzone z wykorzystaniem własnego klienta pocztowego. W PIAP może nim być klient POP3, IMAP, bądź pakiet pocztowy ***SquirrelMail*** bazujący na PHP i obsługujący IMAP oraz SMTP w szyfrowanym trybie dostępu (https).

3. ANTYWIRUSOWY FILTR POCZTOWY

Funkcjonująca poczta elektroniczna została uzupełniona o filtrowanie wiadomości. Jako filtr antywirusowy zastosowano w PIAP oprogramowanie poza systemowe: ***KASPERSKY ANTI-VIRUS® 5.5 FOR LINUX, FREEBSD AND OPENBSD MAIL SERVER (KAV)***. Na podstawie analizy intensywności ruchu pocztowego w PIAP oszacowano, że filtr pocztowy nie wymaga stosowania oddzielnego serwera; może być on zainstalowany na serwerze, na którym funkcjonuje system web, intranet, DNS, e-mail, podstawowy backup, oraz szeroko rozumiana obsługa sieciowa. Dzięki temu stosuje się jedno rozwiązanie sprzętowe i zintegrowany software (konwergencja).

Włączenie demona filtrowania antywirusowego ***aveserver*** ma miejsce podczas uruchamiania systemu operacyjnego. Parametry pracy filtra antywirusowego określono z zapisami w pliku konfiguracyjnym ***/etc/kav/5.5/kav4mailservers.conf.*** We współpracy z demonem ***aveserver*** i antywirusową bazą danych przeprowadzane jest badanie obecności wirusów w poczcie. Procesem tym zarządza program ***smtpscanner***, wchodzący w zestaw narzędzi ***KAV***.

Podczas przetwarzania przekazów pocztowych przez ***Postfix***, przed przejęciem kolejnej wiadomości pocztowej przez agenta ***local*** (przy przesyłkach do użytkownika) lub przez agenta ***smtp*** (przy przesyłkach od użytkownika), następuje przekierowanie ruchu pocztowego za pomocą protokołu ***lmtp*** do antywirusowego filtra pocztowego celem przeskanowania kolejnych wiadomości. Wiadomości są testowane przez ***aveserver***, a następnie, w oparciu o uzyskany wynik, są kierowane przez ***smtpscanner***, za pomocą protokołu ***smtp***, do menedżera ***qmgr***, a w przypadku zawirusowania mogą być poddane próbie leczenia, kwarantannie, lub mogą zostać usunięte, przy czym każdej z tej akcji może towarzyszyć, w zależności od konfiguracji, przekazanie odpowiedniej informacji o statusie testu do nadawcy i/lub odbiorcy wiadomości oraz do administratora systemu.

3.1. Modyfikacja Postfix, integrująca z filtrem pocztowym

Instalacja filtru antywirusowego o opisanej funkcjonalności wymaga poniższych modyfikacji w systemie Postfix:

1. Założenie grupy i użytkownika ***filter*** oraz ich skonfigurowanie:
   - ***mkdir /var/spool/filter***
   - ***groupadd filter***
   - ***useradd filter –s /bin/nologin –d /var/spool/filter –g filter***
   - ***chown filter.filter /var/spool/filter***
2. Skonfigurowanie w pliku konfiguracyjnym ***master.cf*** wymogu skanowania antywirusowego:

smtp inet n - n - - smtpd
-o content_filter=lmtp:127.0.0.1:10030

oraz zdefiniowanie funkcjonowania skanera smtp:

```
127.0.0.1:10030 inet n n n - 20 spawn
        user=kluser argv=/opt/kav/5.5/kav4mailservers/bin/smtpscanner
127.0.0.1:10031 inet n - n - 21 smtpd
        -o content_filter=
        -o local_recipient_maps=
        -o relay_recipient_maps=
        -o smtpd_restriction_classes=
        -o smtpd_client_restrictions=
        -o smtpd_helo_restrictions=
        -o smtpd_sender_restrictions=
        -o mynetworks=127.0.0.0/8
        -o strict_rfc821_envelopes=yes
        -o smtpd_error_sleep_time=0
        -o smtpd_soft_error_limit=1001
        -o smtpd_hard_error_limit=1000
        -o myhostname=ni.piap.pl
kluser:x:401:401::/var/db/kav:/sbin/nologin
```

4. ANTYSPAMOWY FILTR POCZTOWY

SPAM to nazwa mielonki ("SPiced HAM luncheon meat") produkwanej przez firmę Hormel Foods. W roku 1987 powstał skecz pt. „Monthy Python Viking Scene": małżeństwo odwiedza restaurację, w której wszystkie potrawy zawierają właśnie dodatek SPAM'u, a na dodatek stołuje się tam grupa wikingów, coraz bardziej natarczywie wychwalających zalety tej potrawy. Hormel Foods Corporation przyjęły SPAM jako swój znak handlowy, dlatego też do określania niechcianej korespondencji stosuje się ***spam (a nie SPAM)***, co podkreśla definicja słowa ***spam*** zawarta w RFC 2828.

Wykorzystując strukturę modułową systemu pocztowego Postfix uzupełniono go o inteligentny filtr antyspamowy – ***spamassassin*** - będący częścią wersji dystrybucyjnej systemu operacyjnego ***Fedora Linux*** zaimplementowanego w serwerze sieciowym w PIAP. Implementację przeprowadzono dwutorowo: uruchomienie systemu ***spamassassin*** oraz zintegrowanie go z zaimplementowanym wcześniej systemem pocztowym ***Postfix***.

Wywołanie ***spamassassin*** uruchamia demona ***spamd***, z opcjami zawartymi w pliku konfiguracyjnym ***/etc/sysconfig/spamassassin***. Demon ten będzie skanował wiadomość pocztową na żądanie klienta ***spamc*** tworzącego parę z demonem ***spamd***. Każdorazowo po przekazaniu wiadomości do przeskanowania przez menadżera kolejki ***qmgr***, zgodnie z zapisem w pliku konfiguracyjnym Postfixa ***master.cf*** uaktywnia się klient ***spamc*** nadzorując testowanie. Wynik testu ***spamc*** przekazuje z powrotem do kolejki pocztowej ***Postfix***. Do tego etapu toru przetwarzania e-maili docierają te wiadomości, które nie posiadają cech spamu, oraz które są pozbawione wirusów. Przetwarzane wiadomości są w tym momencie przekazywane agentom ***local*** bądź ***smtp***, w zależności od kierunku przepływu poczty.

4.1. Modyfikacja Postfix, integrująca z filtrem antyspamowym

Instalacja filtru antyspamowego o opisanej funkcjonalności wymaga poniższych modyfikacji w systemie Postfix:

1. Założenie grupy i użytkownika ***spamfilter*** oraz ich skonfigurowanie:
   - ***mkdir /home/Special/spamfilter***
   - ***groupadd filter***
   - ***useradd spamfilter –s /bin/nologin –d /home/Special/spamfilter –g filter***
   - ***chown filter.filter /home/Special/spamfilter***
2. Skonfigurowanie w pliku konfiguracyjnym ***master.cf*** wymogu skanowania antyspamowego:

smtp inet n - n - - smtpd
-o content_filter=spamfilter

oraz zdefiniowanie funkcjonowania skanera antyspamowego:

spamfilter unix - n n - - pipe
flags=Rq user=spamfilter argv=/usr/bin/postfixfilter -f ${sender} -- ${recipient}

3. Zdefiniowanie działania klienta ***spamc*** zapisem w pliku wykonawczym ***/usr/bin/postfixfilter:***

/usr/bin/spamc | /usr/sbin/sendmail -i "$@"
exit $?

5. INTERAKCJA FILTRÓW Z UŻYTKOWNIKIEM

O ile filtr antywirusowy jest ustawiony tak, że zawirusowana wiadomość jest odrzucana obligatoryjnie, bez powiadamiania tak Nadawcy jak i Adresata, to filtr antyspamowy, wspomagany dodatkowo oprogramowaniem autorskim, zachowuje się w bardziej wyrafinowany sposób. I tak:

Filtr antyspamowy powoduje, że każda wiadomość wskazana jako spam:

- ma przypisaną wycenę punktową, w zależności od tego, jak spełnia ona określone wcześniej kryteria, szczegółowo określone zapisami w plikach konfiguracyjnych.
- ma przeprowadzoną ocenę przyczyn tego rezultatu,
- jest umieszczana w katalogu będącym własnością użytkownika, umiejscowionym w serwerze pocztowym

Oprogramowanie autorskie zapewnia, że

- zbiory smamów są cotygodniowo usuwane automatycznie z serwera, przy czym użytkownik ma dostęp do swojego prywatnego zbioru spamów oraz jest cotygodniowo, automatycznie, informowany o podjętych działaniach,
- zbiory spamów są analizowane, przetwarzane i udostępniane użytkownikom w celu łatwego przeglądu przez użytkownika będącego ich właścicielem,

Spamassassin m.in.

- przystosowuje się do e-maili danego użytkownika i uczy się na bieżąco na podstawie nadchodzącej poczty jak traktować nadchodzące wiadomości,
- tworzy indywidualne, prywatne słownikowe bazy danych wykorzystywane przez jeden centralny system do kwalifikacji e-maili poszczególnych użytkowników

systemu pocztowego pod kątem przynależności tych wiadomości do spamu. Te bazy danych zawarte są w plikach: ***~/user/.spamassassin/bayes_seen, .../bayes_toks, .../bayes_journal***

- umożliwia ustalenie użytkownikowi indywidualnego poziomu restrykcyjności przy testach na spam (konfiguracja pliku ***~/user/.spamassassin/user_prefs***)
- pozwala użytkownikowi zestawiać tzw. Białą i/lub Czarną Listę Nadawców, których wiadomości są traktowane zawsze jako prawidłowe, bądź zawsze jako spam.
- Dostarcza narzędzi do indywidualnego zarządzania i ręcznego uczenia zestawu generowanego przez filtr Bayesa za pomocą narzędzia:
  (***sa-learn —spam PlikZeSpamem***)

5.1. Funkcjonowanie SpamAssassin

SpamAssassin jest filtrem pocztowym, lub klasyfikatorem, którego rola polega na zidentyfikowaniu spamu przy użyciu różnorodnych mechanizmów takich jak analizatory tekstu, filtrowanie Bayesa, filtrowanie heurystyczne, listy blokujące DNS oraz współpracujące bazy danych wspomagające filtrowanie.

5.1.1. Filtry Bayesa

Filtr Bayesa porównuje częstość występowania określonych słów w zwykłej i niepożądanej poczcie, i na tej podstawie określa prawdopodobieństwo czy e- mail jest spamem. Filtr tworzy rodzaj słownika, w którym każde ze słów ma przypisane prawdopodobieństwo jego wystąpienia w spamie. Następnie w momencie gdy do użytkownika przychodzi nowy mail, filtr dzieli tekst wiadomości na pojedyncze wyrazy i przypisuje im odpowiednie prawdopodobieństwa otrzymane ze słownika.

Do oceny dokumentu, czy jest on spamem, czy nie, na podstawie jego zawartości, stosuje się tu metodę ***Naiwnego Klasyfikatora Bayesa***. Jako pierwszy metodę tę zastosował Paul Graham opisując wyniki w swoim artykule „A Plan for spam" [4]. Istota pomysłu jest następująca. Jeśli założymy, że dokumenty są przypisane do pewnej liczby klas dokumentów, które mogą być modelowane jako komplety słów, oraz, że i-te słowo w_i danego dokumentu zdarza się w dokumencie klasy C z (niezależnym!) prawdopodobieństwem, zapisywanym jako:

$$p(w_i \mid C)$$

oraz założymy, że prawdopodobieństwo wystąpienia słowa w dokumencie jest niezależne od długości dokumentu lub też, że wszystkie dokumenty mają tę samą długość, to wtedy prawdopodobieństwo, że dany dokument "D" należy do klasy "C" zapiszemy jako:

$$p(D \mid C) = p(D \mid C) = \prod_i p(w_i \mid C)$$

Przyjmijmy dalej, że istnieją tylko dwie klasy C: ***S*** (spam) i ***¬S*** (nie-spam). Określając jakie jest prawdopodobieństwo, że dany dokument *D* należy do danej klasy *C oraz* korzystając z definicji prawdopodobieństwa warunkowego i z Bayesianu, można po przekształceniach określić metodę klasyfikowania dokumentu jako spam, gdy:

$$\ln \frac{p(S \mid D)}{p(\neg S \mid D)} > 0$$

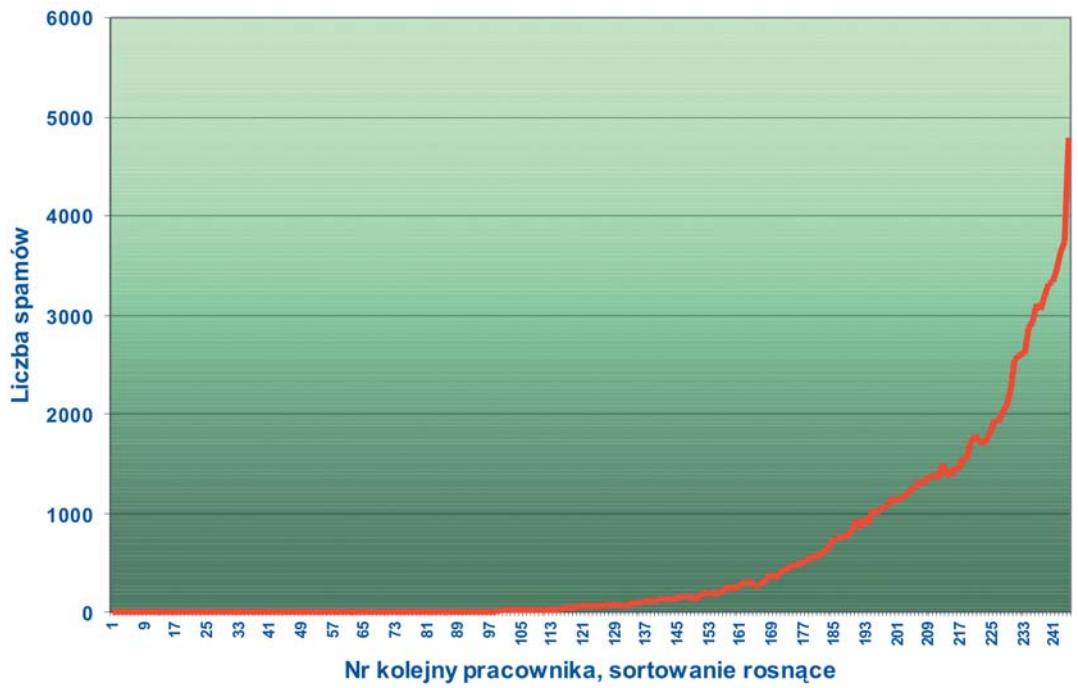
przy czym prawdopodobieństwa $p(S \mid D)$ i $p(\neg S \mid D)$ są obliczane z wykorzystaniem indywidualnych, prywatnych słownikowych baz danych konstruowanych przez SpamAssassin niezależnie dla każdego użytkownika poczty elektronicznej w systemie.

5.1.2. Filtry heurystyczne (regułowe)

Filtry oparte na regułach wyszukują wzorce charakterystyczne dla spam, a konkretnie szukają pewnych wyrazów i fraz, ciągów wielkich liter czy wykrzykników. Znalezionym ciągom odpowiadającym wzorcom jest przypisywana punktacja. Suma punktów klasyfikuje stopień „siły" spamu. W porównaniu z filtrami Bayesowskimi są one prostsze w implementacji, nie wymagają uczenia, ale ich statyczny charakter wymaga wprowadzania zmian dostosowujących je do nowych sztuczek spamerów.

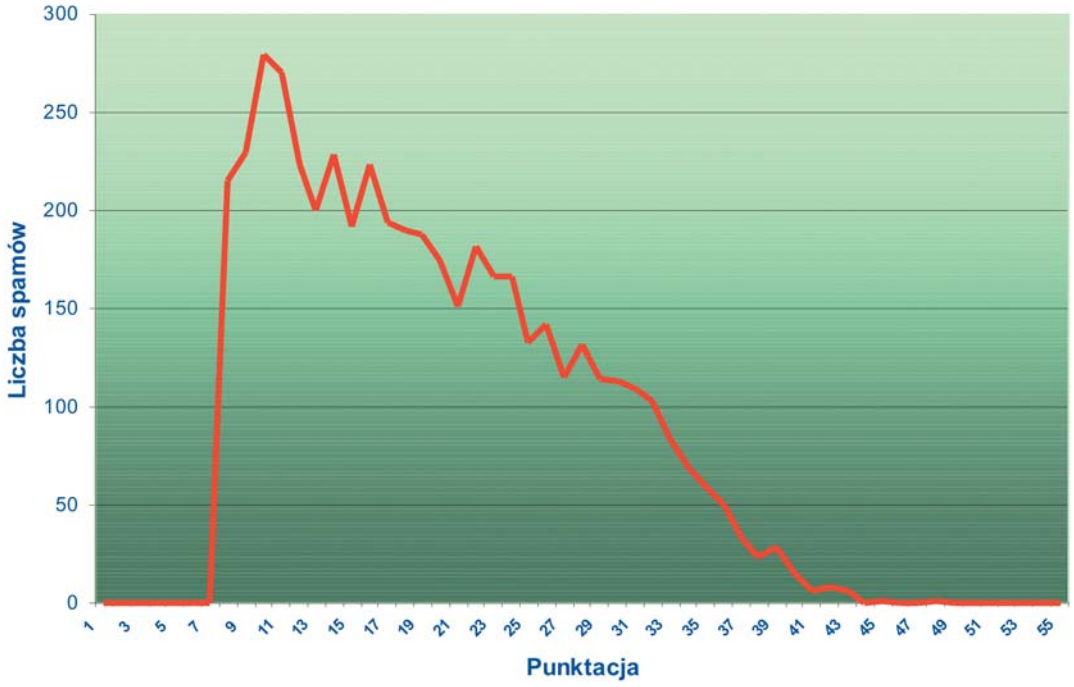
6. PODSUMOWANIE

System antyspamowy działa w PIAP od III kwartału 2007 r. Roczny okres obserwacji pozwala sporządzić pewną analizę funkcjonowania systemu i uogólnić przeprowadzone obserwacje. W PIAP przyjęto globalnie poziom restrykcyjności 7. Oznacza to, że 7 punktów przypisanych do określonej wiadomości elektronicznej stanowi o tym, że jest ona spamem. (Niezależnie od ustawień globalnych każdy z użytkowników może ustawić własny poziom restrykcyjności. Własne ustawienia zastosowano w ok. 5% przypadków. Doświadczenie wskazuje, że poziom 5 lub 4 nie jest przyczyną „gubienia wiadomości"). Przy ww. ustawieniu globalnym liczba e-maili klasyfikowanych jako spamy osiąga wartość ok. 35000 tygodniowo, co daje średnią ok. 150 spamów na osobę. Jednakże zbadany rozkład spamów (rys. 3) wskazuje, że liczba przypadających spamów na pracownika jest silnie nieliniowa. Bliższa analiza wskazuje, że rozkład ten jest powodowany aktywnością w wymianie wiadomości e-mail oraz brakiem higieny w propagowaniu swojego własnego adresu pocztowego.



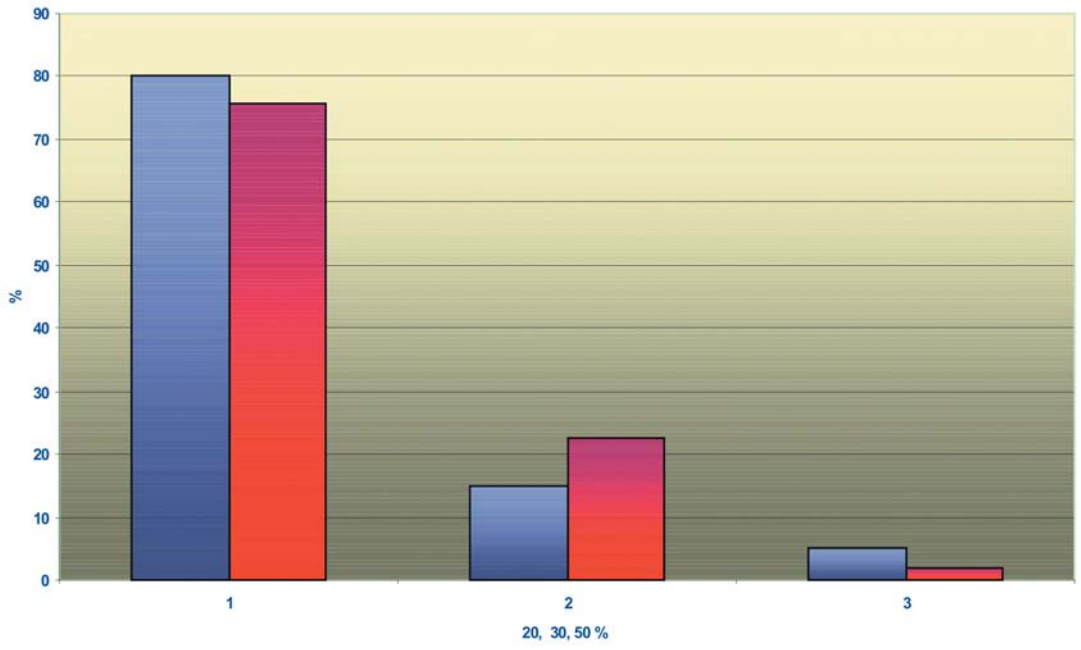
Rys. 3. Liczba spamów przypadająca użytkownikom poczty elektronicznej

Analiza jednego pracownika – rekordzisty spamowego (5000 spamów tygodniowo) – pokazuje rozkład punktów kwalifikujących przesyłaną wiadomość jako spam (rys. 4).



Rys. 4. Liczba spamów przypadająca określonym wartościom punktów karnych

Na podstawie informacji uzyskanych z tej obserwacji można aproksymować liczbę spamów w funkcji przypisywanych punktów karnych (malejąco).



Rys. 5. Wykres Pareto liczby spamów pracowników PIAP

Na uwagę zasługuje również analiza wg prawa ***Pareto*** rozkładu spamów przypadających na użytkownika poczty w PIAP (rys. 5). Na podstawie porównania z wykresem teoretycznym można wnioskować, że spamy atakujące pracowników PIAP da się opisać zgodnie z prawem zwanym zamiennie ***ABC, Pareto lub 20-80***

Poniżej przedstawiono komentarz systemu antyspamowego dotyczącego jednej przesyłki.

From Raber@archieeurope.info Wed Mar 21 08:24:33 2007
Return-Path: <Raber@archieeurope.info>
X-Spam-Flag: YES
X-Spam-Checker-Version: SpamAssassin 3.1.8-rules070225mw (2007-02-13) on ni.piap.pl X-Spam-Level:
****
X-Spam-Status: Yes, score=30.8 required=8.0 tests=BAYES_99,EXTRA_MPART_TYPE,
HTML_50_60,HTML_IMAGE_ONLY_20,HTML_MESSAGE,HTML_SHORT_LINK_IMG_3,
RATWARE_RCVD_PF,RCVD_IN_NJABL_DUL,RCVD_IN_SORBS_DUL,RCVD_IN_WHOIS_BOGONS,
URIBL_AB_SURBL,URIBL_JP_SURBL,URIBL_SBL,URIBL_SC_SURBL autolearn=spam version=3.1.8-
rules070225mw
X-Spam-Report: =?ISO-8859-1?Q?
•3.6 RATWARE_RCVD_PF Bulk email fingerprint (Received PF) found
•* 1.1 EXTRA_MPART_TYPE Nagłówek zawiera nadmiarowy wpis
•* Content- type:...type=
•* 1.2 HTML_IMAGE_ONLY_20 BODY: HTML: images with 1600-2000 bytes of words
•* 0.1 HTML_50_60 BODY: Wiadomość zawiera 50 do 60% HTML
•* 0.0 HTML_MESSAGE BODY: Wiadomość zawiera kod HTML
•* 3.5 BAYES_99 BODY: Bayesowskie prawdopodobieństwo spamu wynosi * 99 do 100%
•* [score: 1.0000]
•* 2.4 RCVD_IN_WHOIS_BOGONS RBL: CompleteWhois: sender on bogons IP block
•* [110.100.162.153 listed in combined-HIB.dnsiplists.completewhois.com]
•* 2.0 RCVD_IN_SORBS_DUL RBL: SORBS: sent directly from dynamic IP address [85.180.62.42 listed in
dnsbl.sorbs.net]
•* 1.9 RCVD_IN_NJABL_DUL RBL: NJABL: nadawca dial-up używa nie lokalnego
•* SMTP
•* [85.180.62.42 listed in combined.njabl.org]
•* 1.6 URIBL_SBL Contains an URL listed in the SBL blocklist
•* [URIs: gamelive.hk]
•* 3.8 URIBL_AB_SURBL Contains an URL listed in the AB SURBL blocklist
•* [URIs: gamelive.hk]
•* 4.1 URIBL_JP_SURBL Contains an URL listed in the JP SURBL blocklist
•* [URIs: gamelive.hk]
•* 4.5 URIBL_SC_SURBL Contains an URL listed in the SC SURBL blocklist
•* [URIs: gamelive.hk]
•* 0.9 HTML_SHORT_LINK_IMG_3 HTML is very short with a linked
• image?=
•X-Original-To: xxxxxxxxxxx@piap.pl
•Delivered-To: xxxxxxxxxxx@piap.pl

7. LITERATURA

1. Kyle D. Dent, 2004 r., Postfix. Przewodnik encyklopedyczny, Helion
2. Ralf Hildebrandt, Patrick Koetter, 2006 r., Postfix nowoczesny system przesyłania wiadomości, Helion
3. Zdziarski Jonathan A., 2006 r., Spamowi STOP! Bayesowskie filtrowanie zawartości i sztuka statystycznej klasyfikacji języka, PWN
4. Paul Graham, A plan for spam, 2002 r., Harvard University