Tomasz Andrzej Nidecki
tomasz.nidecki@students.mimuw.edu.pl
nr indeksu 136413
Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Metody ochrony przed spamem po przyjęciu listu przez serwer oraz nowatorskie rozwiązania

Seminarium: Protokoły komunikacyjne
dr Sławomir Lasota, dr hab. Jerzy Tyszkiewicz
[1000-2D02PK], SOCRATES: 11304

2006-03-07

Plan ogólny referatu

- Ochrona przed spamem po odebraniu listu.
  - Wady i zalety metod ochrony przed spamem po odebraniu listy przez serwer.
  - Przegląd metod:
    - analiza heurystyczna,
    - analiza statystyczna (bayesowska),
    - sieci antyspamowe (systemy sum kontrolnych),
    - systemy wyzwanie-odpowiedź (challenge-response),
- Nowatorskie rozwiązania antyspamowe.
  - koszty po stronie spamera.
  - uwierzytelnianie przez firmy trzecie.
  - całkowicie nowa infrastruktura pocztowa.

Ochrona przed spamem po odebraniu listu

Wady i zalety ochrony po odebraniu listu

- Zalety:
  - O wiele większa ilość danych, na których można oprzeć analizę — większa dokładność.
  - Możliwość wybiórczego stosowania analizy w zależności od odbiorcy.
  - Możliwość dostosowania analizy do indywidualnych życzeń użytkownika.
- Wady:
  - Konieczność przyjęcia każdego listu — większe zużycie zasobów (łącza, pamięć, dysk).
  - Większa złożoność analizy — również zużycie zasobów (procesor, pamięć, dysk).
  - W przypadku niektórych metod — zwiększenie ruchu na łączach.

Przegląd metod — analiza heurystyczna

- Analiza heurystyczna polega na wykrywaniu wystąpienia charakterystycznych elementów w nagłówkach i treści listu, a następnie przypisaniu im określonych wag (pozytywnych lub negatywnych).
- Przykłady wag negatywnych:
  - Brak adresu odbiorcy w nagłówku ***to:***.
  - Wystąpienie określonych słów, np. *viagra*, *v1agr4* itp.
  - Data wiadomości z przyszłości.
  - Nieprawidłowy format nagłówka ***msgid:***.
- Przykłady wag pozytywnych:
  - Prawdziwy adres odbiorcy w nagłówku ***to:***.
  - Zweryfikowana tożsamość nadawcy (SPF, DKIM itp.).
- Administrator określa poziom wag, przy których następuje rozgraniczenie między hamem i spamem.

Przegląd metod — analiza heurystyczna

- Zalety analizy heurystycznej:
  - Możliwość scentralizowanej dystrybucji aktualizacji zasad.
  - Możliwość dostosowania do indywidualnych potrzeb użytkownika (np. korzystanie z wybranych zestawów zasad).
  - Niewielkie zaangażowanie użytkownika i administratora .
- Wady analizy heurystycznej:
  - Duże zużycie zasobów (analiza jest czasochłonna, często wiąże się z koniecznością korzystania z innych metod).
  - Średnia skuteczność (niższa od np. analizy statystycznej).
  - Konieczność częstego aktualizowania zasad i uzależnienie od dostawcy rozwiązania.
- Najlepiej sprawdza się na serwerach firmowych charakteryzujących się dużą dostępnością zasobów, średnim ruchem oraz koniecznością oszczędzania czasu administratora i użytkowników.

Przegląd metod — analiza statystyczna

- Analiza statystyczna (bayesowska) polega na wskazaniu systemowi, jakie charakterystyczne elementy listu świadczą o tym, że jest to spam, a jakie, że ham.
- Nie jest to koncepcja nowa (M. Sahami, S. Dumais, D. Heckerman, E. Horvitz: *A Bayesian approach to filtering junk e-mail*, 1998.), ale zdobyła popularność dopiero dzięki artykułowi Paula Grahama w 2002 r.
- Największym problemem w wykorzystaniu analizy bayesowskiej do walki ze spamem był sposób dzielenia listu na elementy (tokeny) poddawane analizie. Ma on ogromny wpływ na skuteczność.
  - *http://www.paulgraham.com/spam.html*
  - *http://www.paulgraham.com/better.html*

Przegląd metod — analiza statystyczna

- Zalety analizy statystycznej:
  - Wbrew pozorom, jest o wiele mniej zasobochłonna niż analiza heurystyczna.
  - Umożliwia stosowanie zasad indywidualnie (dla każdego użytkownika) lub globalnie (dla całego systemu).
  - Ma najwyższą praktyczną skuteczność spośród stosowanych obecnie metod.
- Wady analizy statystycznej:
  - Wymaga dużego zaangażowania administratora lub użytkownika (uczenie analizatora bayesowskiego).
  - Stosowanie gotowych kolekcji spamu i hamu powoduje znaczne zmniejszenie praktycznej skuteczności.
  - Staje się skuteczna dopiero po kilku tygodniach uczenia filtra.
- Najlepiej sprawdza się w praktyce, gdy nie jest stosowana na serwerze, lecz w programie klienckim.

Przegląd metod — sieci antyspamowe

- Polegają na tworzeniu sum kontrolnych w oparciu o wybrane nagłówki oraz treść otrzymywanych listów i dzielenie się sumami z innymi użytkownikami (przy użyciu zautomatyzowanych systemów).
- Wersja interaktywna (Razor, Pyzor):
  - Przyjęty list jest oznaczany przez odbiorcę jako spam.
  - System oblicza sumę kontrolną i wysyła do serwera sieci antyspamowej.
  - U innych użytkowników sieci antyspamowej, system sprawdza czy nachodzący list ma sumę kontrolną, która występuje w sieci.
- Wersja zautomatyzowana (DCC):
  - Suma kontrolna każdego odebranego listu jest przesyłana do serwera sieci antyspamowej.
  - Serwer oprócz samej sumy przechowuje również liczbę zgłoszeń danej sumy.
  - U innych użytkowników, system sprawdza sumę nadchodzącego listu i liczbę zgłoszeń. Po przekroczeniu limitu zgłoszeń traktuje jako spam.

Przegląd metod — sieci antyspamowe

- Zalety sieci antyspamowych:
  - Niewielkie zużycie zasobów.
  - W wersji zautomatyzowanej: brak zaangażowania użytkowników i administratora.
- Wady sieci antyspamowych:
  - Bardzo niska skuteczność (zaledwie około 50%).
  - W wersji interaktywnej:
    - uzależnienie od opinii i szybkości reakcji innych użytkowników.
  - W wersji zautomatyzowanej:
    - nierozróżnianie rodzajów wysyłek masowych (listy dyskusyjne są traktowane jako spam).
- Sieci antyspamowe są obecnie stosowane sporadycznie, najczęściej jako element wspomagający w analizatorach heurystycznych.

Przegląd metod — systemy wyzwanie-odpowiedź

- Zasada działania podobna do mechanizmów używanych w listach dyskusyjnych — nadawca musi potwierdzić swoją tożsamość, nim list trafi do odbiorcy.
- Przykład:
  - List trafia do systemu odbiorcy.
  - System wyzwanie-odpowiedź wysyła do nadawcy list z prośbą o potwierdzenie tożsamości (na podstawie adresu nadawcy — nagłówka ***from:***).
  - Nadawca potwierdza (wysyłając list pod podany w wyzwaniu adres, klikając na link itp.)
  - Oryginalny list trafia do odbiorcy.
  - W zależności od rodzaju systemu, adres nadawcy jest dodawany do białej listy (jednorazowe potwierdzanie) lub nie (każdorazowe potwierdzanie).

Przegląd metod — systemy wyzwanie-odpowiedź

- Zalety systemów wyzwanie-odpowiedź:
  - Ogromna skuteczność — praktycznie sto procent odfiltrowywanego spamu.
  - Brak zaangażowania użytkownika i administratora (ew. przeglądanie kolejki).
- Wady systemów wyzwanie-odpowiedź:
  - Duża ilość odrzuconego hamu (osoby, które nie chciały potwierdzić, nie zrozumiały prośby, systemy zautomatyzowane itp.).
  - Niebezpieczeństwo generowania nadmiarowego ruchu (fałszywe adresy nadawców w wirusach i spamie).
- Nadają się do stosowania praktycznie tylko na kontach prywatnych, przez osoby, które otrzymują bardzo dużo spamu i które nie obawiają się ewentualnej utraty części listów (lub też regularnie przeglądają kolejkę).

Nowatorskie rozwiązania antyspamowe

Nowatorskie rozwiązania — koszty po stronie spamera

- Tarpitting i spam throttling
  - Tarpitting to opóźnienie przetwarzania przy połączeniu z serwerem pocztowym, kiedy przekroczona została określona liczba odbiorców (***RCPT TO***).
  - Ponieważ spamerzy obchodzą to ograniczenie generując wiele jednoczesnych połączeń, powstało *spam throttling*, które dodatkowo bierze pod uwagę liczbę połączeń z danego adresu IP.
- Znaczki pocztowe
  - Generowanie cyfrowego znaczka pocztowego wymaga zaangażowania stosownych zasobów przez nadawcę. Odbiorca może zweryfikować jego poprawność. Znaczek jest uzależniony od docelowego odbiorcy, więc znacznie utrudnia życie spamerom.
  - Mechanizm generowania: Hashcash — *http://www.hashcash.org/*.
  - Zastosowanie w praktyce: Camram — *http://www.camram.org/*.
  - Badania: Penny Black Project — *http://research.microsoft.com/research/sv/PennyBlack/*.

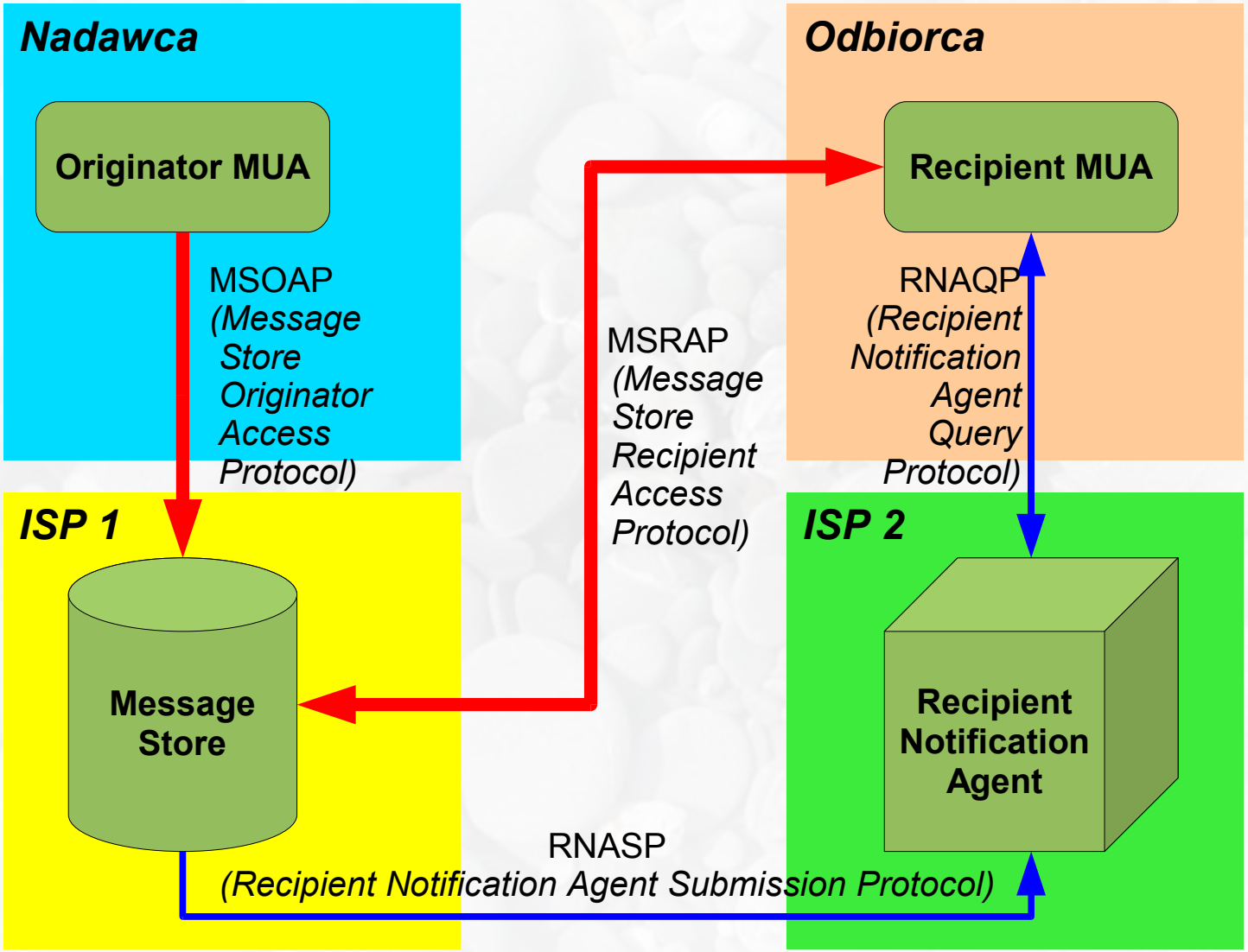
Nowatorskie rozwiązania — uwierzytelnianie

- Habeas
  - Firmy, które przejdą proces autoryzacji (i zapłacą...), mogą wstawiać do listów specjalny nagłówek, gwarantujący że wiadomość nie jest spamem (w rzeczywistości... wiersz haiku podlegający prawu autorskiemu, zawierający nazwę która jest zarejestrowanym znakiem handlowym firmy Habeas).
  - Odbiorca może dla dodatkowej weryfikacji skorzystać z prowadzonych przez Habeas białych i czarnych list adresów IP nadawców stosujących nagłówek Habeas.
  - Jeśli spamer zastosuje nagłówek Habeas, łamie prawa autorskie, a Habeas dysponuje efektywnym zespołem adwokatów i specjalistów sieciowych (do wykrycia winnego).
- Bonded Sender
  - Odpowiednik list DNSBL (RBL), zawierający whitelistę adresów IP, które przeszły weryfikację firmy jako niewysyłające spamu.
- We Can Stop Spam
  - *http://web.archive.org/web/20040614012209/wecanstopspam.org/jsp/Wiki?WeCanStopSpam*
  - *http://wecanstopspam.org/*
- Przemyślenia Brada Templetona (prezesa EFF)
  - *http://www.templetons.com/brad/spam/endspam.html*

Nowatorskie rozwiązania — nowa infrastruktura

- Internet Mail 2000 (IM2000)
  - Całkowicie nowa propozycja infrastruktury sieciowej.
  - Opiera się na założeniu, że za przechowywanie wiadomości odpowiada nadawca.
  - Oryginalna propozycja: Dan J. Bernstein (twórca qmaila, djbdns itp.)
  - Więcej informacji:
    - *http://www.im2000.org/*
    - *http://cr.yp.to/im2000.html*

Nowatorskie rozwiązania — IM2000



Nowatorskie rozwiązania... ?!

- A może jedynym wyjściem z sytuacji jest zastosowanie metody a'la Lcamtuf?...

  http://eprovisia.coredump.cx/

  As soon as you subscribe to our programme, you will be provided with a unique, personal e-mail address on our server. We will also send you a package of detailed, easy to follow instructions on how to redirect your existing mail accounts to this new address.

  *The minute your mail starts flowing, a dedicated team of over a hundred trained Screening and Preselection Specialists, working 24 hours a day**, will begin manually reviewing, hand-picking and approving important correspondence, vigilantly discarding all junk mail.*