Illustration Willam Duke

Hayato Kunigo
Markus Stuhm

Übersicht

- Einführung
- Anti-Spam-Verfahren
- Spam-Vorsorge

Einführung (1)

- Usenet-SPAMs (1)
  - Excessive multi-posting (eigentl. SPAM)
    - Mehrfach versendete Nachrichten
  - Excessive cross-posting
    - An mehrere Empfänger versendete Nachrichten
  - Spew
    - Mehrfach versenden durch falsche Konfiguration

Einführung (2)

- Usenet-SPAMs (2)
  - Off-topic postings
    - Nachrichten, deren Themen mit der Newsgruppe nichts zu tun haben
  - Commercial postings
    - Artikel bzw. Nachrichten, die für etwas Werben

Einführung (3)

- Email-SPAMs (1)
  - Unsolicited commercial email
    - Unerwünschte Werbemail
  - Unsolicited bulk email
    - Unerwünschte Massenmail
  - Make money fast
    - Werbemail in der der schnelle € bzw. $ versprochen wird

Einführung (4)

- Email-SPAMs (2)
  - Reputation attacks
    - Rufmord / -schädigung
      - ... Durch Fälschung bzw. Vortäuschen einer Absenderadresse
      - ... Durch Ausstreuen von Gerüchten
      - Besonders bei Firmen kann dies zu einer wirtschaftlichen Katastrophe führen

Anti-Spam Allgemeines

- Zahlreiche verschiedene Verfahren
- Einsatzgebiet
  - Lokal (Rechner des Benutzers)
  - Zentral (auf dem Mailserver)
  - Extern (z.B. Spamcop, DCC, SpamNet)
- Praxis
  - Kombination der Anti-Spam-Verfahren
  - Stufenweise Prüfung aller E-Mails

Anti-Spam Verfahren (1)

- Verwendung von Listen
  - BlackList
  - WhiteList
- Durchführen einer Textanalyse
  - Heuristische Textanalyse
  - Lexikalische Textanalyse
  - Statische Textanalyse

Anti-Spam Verfahren (2)

- Namensauflösung durch DNS
- Distributed Checksum Clearinghouse
- SpamNet
- Jamspam Konsortium

Verwendung von Listen

- BlackList
  - Liste aus „auffälligen“ E-Mail- und IP-Adressen, die blockiert werden
  - Realtime Blackhole List
    - Enthält ungeschützte Mailserver
- WhiteList
  - Liste der gewünschten E-Mail- und IP-Adressen, die zugestellt werden
  - Andere E-Mail-Adressen werden blockiert

Textanalyse (1)

- Heuristische Textanalyse
  - E-Mail-Kopfzeilenanalyse
    - Auswertung des E-Mail-Header (Routing-Informationen, verwendeter E-Mail-Client des E-Mail-Senders)
    - Gewichtung der analysierten Merkmale
  - Strukturanalyse
    - Auswertung des E-Mail-Body
    - Suche nach bestimmten Strukturen z.B. verstümmelte Wörter, kodierter Text

Textanalyse (2)

- Lexikalische Textanalyse
  - Analyse der E-Mail mit Hilfe von Wortlisten
    - Zu untersuchende Wörter (z.B. free)
    - Gewichtung der Wörter
    - Schwellwert der Wörter
  - Berechnen der Werte der Wörter
  - Vergleich der Werte mit den Schwellwerten
    - Wert › Schwellwert → Blockieren der E-Mail
    - Ansonsten zustellen der E-Mail

Textanalyse (3)

- Statische Textanalyse (1)
  - Aufteilen der E-Mail in Teilstücke (Token)
  - Unterteilung in 2 Phasen
    - Vorgelagerte Lernphase und Betriebsphase
  - Vorgelagerte Lernphase
    - Erfassen von Spam und normaler E-Mail
    - Bestimmen der Auftrittshäufigkeit der Token für Spam und normaler E-Mail
    - Speichern der ermittelten Häufigkeitswerte

Textanalyse (4)

- Statische Textanalyse (2)
  - Betriebsphase
    - Berechnen der Wahrscheinlichkeit jedes Token
    - Berechnen der Gesamtwahrscheinlichkeit
    - Gesamtwahrscheinlichkeit > Schwellwert
      - ➔Blockieren der E-Mail, sonst Zustellen der E-Mail
  - Vertreter
    - Content Recognition Engine (CORE) oder Bayesiam Filter

Namensauflösung

- Verwendung eines DNS
- Auflösung des Domain Teils der E-Mail Adresse des Senders in die IP-Adresse
- Testen auf Übereinstimmung
  - Positiv → Zustellen der E-Mail
  - Negativ → Blockieren der E-Mail
- In Praxis:
  - Umgehung durch Nutzung realer, fremder oder eigener Mail-Systeme

DCC

- Distributed Checksum Clearinghouse
  - Verteilte Prüfsummenverrechnungszentrale
- Wozu? - Zweck
- Funktionsweise
- Sicherheit

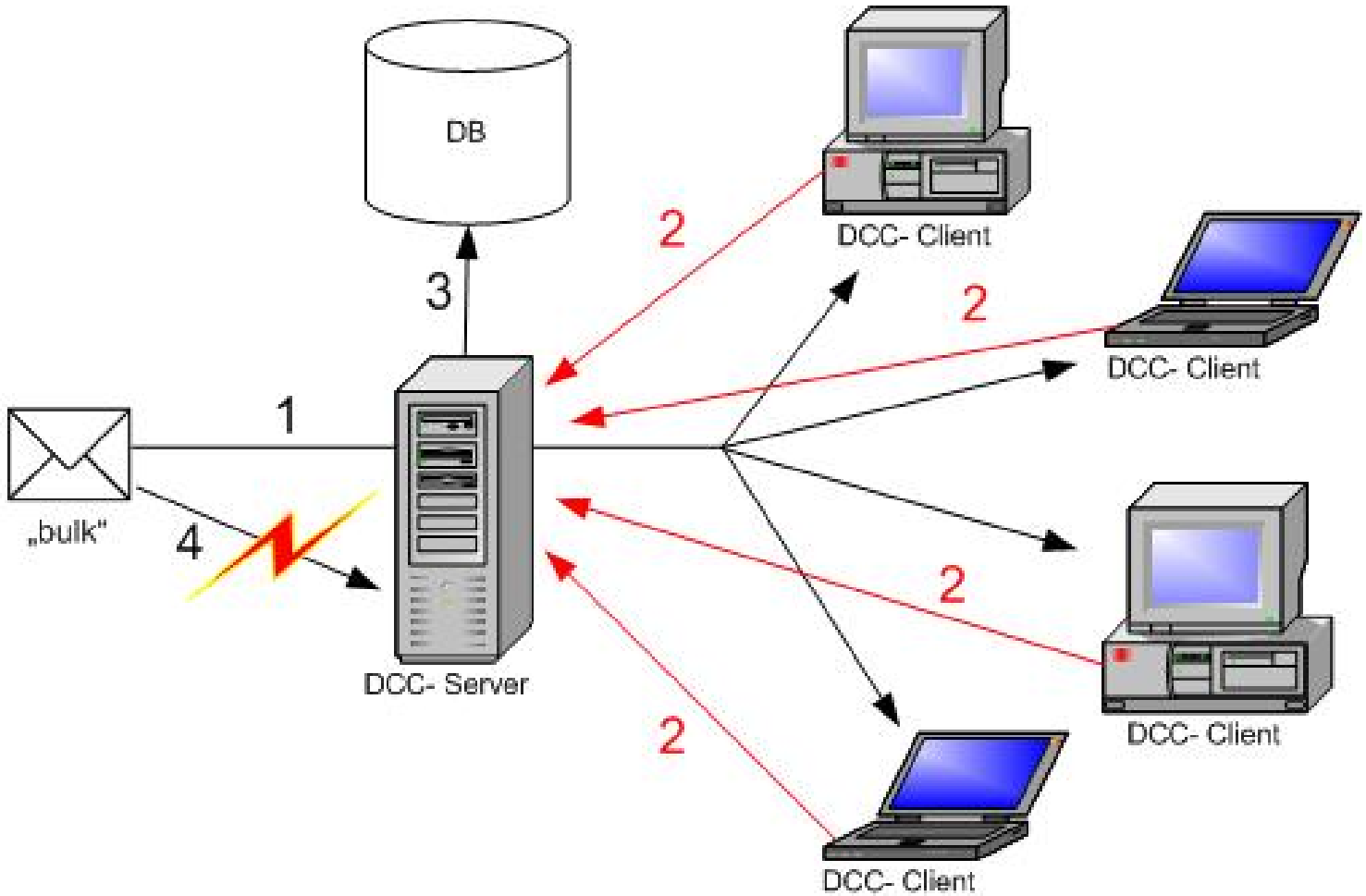
DCC - Zweck

1. SPAMs erkennen
2. SPAMs abblocken
3. Erwünschte „Massenpost“ und Werbung soll die Empfänger erreichen
4. Effizienz !

DCC – Funktionsweise (1)

- Über 100 Server
- Tausende Clients
- Austausch von Prüfsummen / Reports
- DCC-Server tauschen auch unter-einander Reports aus, um „Erfahrungs-werte“ zu verbreiten.

DCC – Funktionsweise (2)



DCC – Funktionsweise (3)

- Prüfsummen sind „fuzzy“
  - Nicht-relevante Unterschiede werden ignoriert
- Report:
  - Prüfsumme vom Absender-Adresse
  - Prüfsumme vom Body
  - Anzahl der Empfänger
    - Einstufung in VIEL und WENIG

DCC – Sicherheit

- Whitelist und Blacklist
- Graylist
  - Temporäres Verwerfen
- Datenaustausch in Klartext
- Gesicherter Empfang

SpamNet

- Kommerzielles Verfahren
- Monatlicher Beitrag 3,99 US$
- Hersteller Cloudmark
- Starke Ähnlichkeit mit DCC
- Meldung neuer Spams durch Benutzer
- Server aktualisiert durch Meldungen den zugehörigen Schwellwert

Jamspam Konsortium

- Zahlreiche Teilnehmer unter anderem AOL, Dell, IBM, Microsoft, Nokia, Oracle
- Zur Zeit im Vorschlagsstadium
- Neue Wege beim E-Mail Versand
  - Deutlich schwerer Spam Versand
  - Keine wirtschaftliche Attraktiv
  - Sichere Authentifizierung des Senders

Spam-Vorsorge (1)

- Überlegte E-Mail Adressweitergabe
- Verschiedene E-Mail Adressen für
  - Newsgroups, Mailinglisten
  - Private E-Mails
- Schützen der E-Mail Adresse bei Veröffentlichung auf der Website
- Verwendung von Dummy E-Mail Adressen beim Download von Dateien

Spam-Vorsorge (2)

- BCC-Adressierung bei untereinander unbekannten Empfängern
- Löschen/ Blockieren von Spams
- In keinem Fall auf Spam antworten
- Links in Spams nicht besuchen
- Boykott der in den Spam Mails beworbenen Produkte

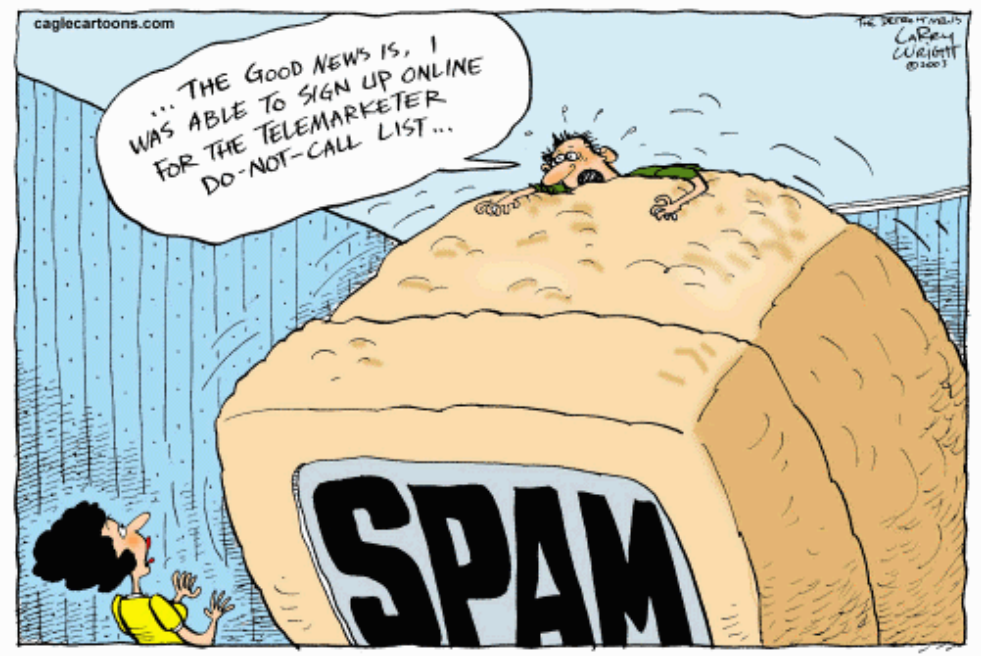
Quellenangaben (1)

- Peter Klau: Perfekt geschützt vor Spam & Spy
  - bhv, 2003
  - ISBN: 3-8266-7316-6
- Alan Schwartz & Simson Garfinkel: Stoppt Spam
  - O'Reilly, 2000
  - ISBN: 3-89721-221-8
- DCC Dokumentation
  - Softwareversion 1.2.14 vom 19.10.2003
  - URL: http://www.dcc-servers.net/dcc/

Quellenangaben (2)

- Prof. Dr. N. Pohlmann
  Vorläufiges Paper Anti Spam Technolog.
- Zeitschrift LANline
- Zeitschrift Monitor
- Weitere siehe Ausarbeitung

Vielen Dank für Ihre Aufmerksamkeit



Fragen ???