Types of Spam Filters

- Community Filters
  - Work on the principal of "communal knowledge" of spam
  - These types of filters communicate with a central server.
- Bayesian Filters
  - Statistical email filtering
  - Uses Naïve Bayes classifier

Introduction

- Federal Regulations (CAN-SPAM act of 2003)
  - Controlling the Assault of Non-Solicited Pornography And Marketing (CAN-SPAM) Act of 2003 – signed on 12/16/2003
  - Commercial email must adhere to three types of compliance
    - Unsubscribe Compliance
    - Content Compliance
    - Sender Behavior Compliance

Types of Spam Filters

- Content Filters
  - Scan the text content of emails
  - Use fuzzy logic
- Permission Filters
  - Based on Challenge /Response system
- White list/blacklist Filters
  - Will only accept emails from list of "good email addresses"
  - Will block emails from "bad email addresses"

Computing the Probability

3. Combining individual probabilities:

- The Bayesian spam filtering software makes the "naïve" assumption that the words present in the message are independent events
- With that assumption,

$$p = \frac{p_1 p_2 \cdots p_N}{p_1 p_2 \cdots p_N + (1 - p_1)(1 - p_2) \cdots (1 - p_N)}$$

where:

- p is the probability that the suspect message is spam
- p_i is the probability $p(S|W_i)$

Some Statistics

- Spam Averages about 88% of all emails sent
  - Barracuda Networks:

Spam Data

Email Processed (last 2 days)

	Yesterday 05/29		Today 05/30 (so far)	
Blocked: Spam	426,479,573	88.72%	356,240,825	88.75%
Blocked: Virus	92,685	0.02%	57,563	0.01%
Quarantined	3,570,495	0.74%	2,951,550	0.74%
Allowed: Tagged	3,572,008	0.74%	2,688,442	0.67%
Allowed	46,979,422	9.77%	39,447,928	9.83%
Total Received	480,694,183	100.00%	401,386,308	100.00%

Last Updated 05/30/11 02:45:51 GMT

Introduction

- Federal Regulations (CAN-SPAM act of 2003)
  - Controlling the Assault of Non-Solicited Pornography And Marketing (CAN-SPAM) Act of 2003 – signed on 12/16/2003
  - Commercial email must adhere to three types of compliance
    - Unsubscribe Compliance
    - Content Compliance
    - Sender Behavior Compliance

Computing the Probability

2. Spam or ham:

$$\Pr(S|W) = \frac{\Pr(W|S)}{\Pr(W|S) + \Pr(W|H)}$$

- This quantity is called "spamicity" (or "spaminess") of the word "*replica*"
  - **Pr(*W*|*S*)** is approximated to the frequency of messages containing "*replica*" in the messages identified as **spam** during the learning phase
  - **Pr(*W*|*H*)** is approximated to the frequency of messages containing "*replica*" in the messages identified as **ham** during the learning phase

Introduction

- Spam as a problem
  - Consumes computing resources and time
  - Reduces the effectiveness of legitimate advertising
  - Cost Shifting
  - Fraud
  - Identity Theft
  - Consumer Perception
  - Global Implications
- John Borgan [ReplyNet] – "Junk email is not just annoying anymore. It's eating into productivity. It's eating into time"

Some Statistics

Business Email Users, 2005-2010

	2005	2006	2007	2008	2009	2010
North America	125.2	128.7	132.4	136.0	139.8	143.6
Europe	162.6	179.8	196.5	212.8	228.6	244.1
Other Americas	179.1	191.9	204.7	217.4	230.2	243.0
Africa	16.0	19.5	23.0	26.6	30.1	33.7
Asia (incl. Mid-East)	182.8	198.3	213.6	229.0	244.3	259.6
Oceania	8.7	9.1	9.5	9.9	10.4	10.8
Total	674.2	727.3	779.7	831.7	883.3	934.8

Source: Ferris Research, The Email Security Market, 2005-2010

Figures are in millions of users, rounded to the nearest 100,000.

Some Statistics

- Cost of Spam 2009:
  - $130 billion worldwide
  - $42 billion in US alone
  - 30% increase from 2007 estimates
  - 100% increase in 2007 from 2005
- Main components of cost:
  - Productivity loss from inspecting and deleting spam missed by spam control products (False Negatives)
  - Productivity loss from searching for legitimate mails deleted in error by spam control products (False Positives)
  - Operations and helpdesk costs (Filters and Firewalls – installment and maintenance)

Spam Filtering Techniques

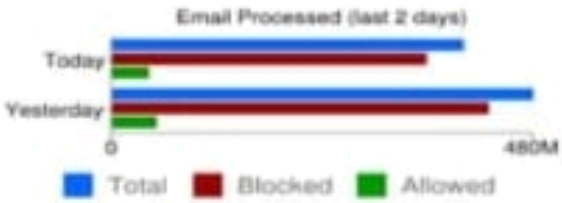
UMAR M. A. ALHARAKY, B.SC.

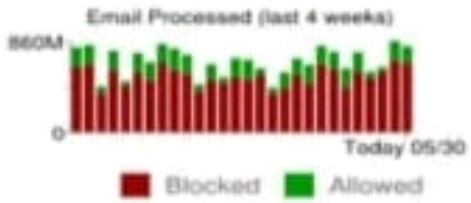
RAKAN RAZOUK, PH.D.

Some Statistics

- Spam Averages about 88% of all emails sent
  - Barracuda Networks:

Spam Data





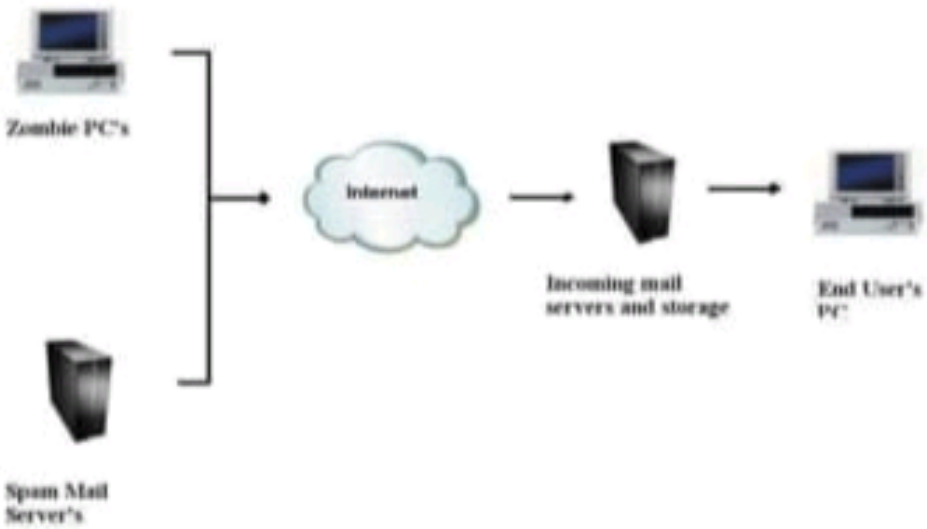
Email Processed (last 2 days)

	Yesterday 05/29		Today 05/30 (so far)	
Blocked: Spam	426,479,573	88.72%	356,240,825	88.75%
Blocked: Virus	92,685	0.02%	57,563	0.01%
Quarantined	3,570,495	0.74%	2,951,550	0.74%
Allowed: Tagged	3,572,008	0.74%	2,688,442	0.67%
Allowed	46,979,422	9.77%	39,447,928	9.83%
Total Received	480,694,183	100.00%	401,386,308	100.00%

Last Updated 05/30/11 02:45:51 GMT

Spam Life Cycle

Life Cycle of Spam



Introduction

- Email Address Harvesting - Process of obtaining email addresses through various methods
  - Purchasing /Trading lists with other spammers
  - Bots
  - Directory harvest attack
  - Free Product or Services requiring valid email address
  - News bulletins /Forums

Types of Spam Filters

- Header Filters
  - Look at email headers to judge if forged or not
  - Contain more information in addition to recipient , sender and subject fields
- Language Filters
  - filters based on email body language
  - Can be used to filter out spam written in foreign languages

Introduction

- **Purpose of Spam**
  - Advertisements
  - Pyramid schemes (Multi-Level Marketing)
  - Giveaways
  - Chain letters
  - Political email
  - Stock market advice

Conclusions

- Spam emails not only consume computing resources, but can also be frustrating
- Numerous detection techniques exist, but none is a "good for all scenarios" technique
- Data Mining approaches for content based spam filtering seem promising

Thank You / Questions