UNIVERSITY OF CALIFORNIA
RIVERSIDE

Fighting Spam, Phishing and Email Fraud

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Computer Science

by

Shalendra Chhabra

December 2005

Thesis Committee:

Professor Dimitrios Gunopulos, Chairperson
Professor Vana Kalogeraki
Professor Eamonn Keogh
Professor Mart Molle
Dr. William S. Yerazunis

The Thesis of Shalendra Chhabra is approved:

_____

_____

_____

_____

_____

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Fighting Spam, Phishing and Email Fraud

by

Shalendra Chhabra

Master of Science, Graduate Program in Computer Science
University of California, Riverside, December 2005
Professor Dimitrios Gunopulos, Chairperson

Spamming in the electronic communications medium is the action of sending unsolicited commercial messages in bulk without the explicit permission or desire of the recipients. The most common form of spam is email spam. Phishing is a particular type of spam characterized by attempts to masquerade as a reputed business. The objective is to trick recipients into divulging sensitive information such as bank account numbers, passwords, and credit card details. Spam and phishing cause billions of dollars' worth of losses to businesses.

Many initiatives on technical and legal levels are currently underway for fighting this challenge. In this thesis we will examine these issues from the aspects of network protocols, filtering, reputation, human psychology, scalability and corporate alliances.

We will present a comprehensive description of technical initiatives to fight spam which include server-side and client-side filtering (using statistical and collaborative techniques); lists (blacklist, whitelist, greylist and brownlist (in CAMRAM)); email authentication standards such as Identified Internet Mail (IIM) from Cisco Systems, Domain Keys (DK) from Yahoo!, Domain Keys Identified Mail (DKIM), Sender Policy Framework (SPF) from Pobox, Sender ID Framework (SIDF) from Microsoft Corporation; and emerging sender reputation and accreditation services (from Habeas, Return Path, IronPort Systems, and CipherTrust).

We will touch upon specific anti-spam techniques used in the popular spam-filtering appliances such as IronMail Connection Control from CiphterTrust, MailHurdle (RazorGate) from Mirapoint, Mail Security 8160 from Symantec, and MailGate Edge from Tumbleweed Communications.

We will investigate various tricks that spammers use to fool spam filters, and will also analyze a spammer's *to do* list by looking through Jeremy Jaynes's court transcripts. Jaynes was the world's eighth most prolific spammer until he was convicted and sentenced to nine years in prison under Virginia statute. We will illustrate malicious stages in a phishing attack lifecycle. We will also discuss the anatomy of a phishing email and various other tricks that fraudsters use in the spoofed web sites.

We will first explain *sender-pays* model for email and will then describe *The Penny Black Project*, a *challenge-response* system with bankable tokens developed by Microsoft Researchers.

We will focus on a special class of Human Interactive Proofs (HIPs) known as Completely Automatic Public Turing test to tell Computers and Humans Apart (CAPTCHA). CAPTCHAs are used by Hotmail, Yahoo!, Google, and other companies to prevent automatic registration of email accounts by bots and to prevent bulk mailing by the spammers.

We will explain a mechanism for throttling Internet resource abuse based on the cryptographic technique Proof of Work (PoW) known as Hashcash. We will then describe in detail a *hybrid sender-pays* email system based on Hashcash, known as the Campaign for Real Mail or CAMRAM.

We will also focus on the machine learning approach for spam filters. A large number of spam filters and other mail communication systems have been proposed and implemented in the past. We will describe a possible unification of these filters, allowing their technology to be described in a uniform way. In particular, describing these filters reveals a large commonality of designs and explains their similar performance. We will present results of our experiments with the Markov Random Field (MRF) model and Winnow-based approaches for spam-filtering in the open source spam filter CRM114 Discriminator Framework. Our results indicate that such models outperform the Naïve Bayesian approach for spam-filtering. We will illustrate CRM114 usage for small-, medium-, and large-scale enterprises (for filtering up to one million client email accounts). We will also investigate the significance of reputation-based protocols for the email communication flow.

We will highlight some problems with the current spam-fighting techniques. We conclude that with the combination of better spam-fighting techniques, legal actions, awareness among Internet users, and cooperation within the industry, the spammers' business model can be disrupted in the next few years.

In addition to work at the Department of Computer Science and Engineering, University of California Riverside, this thesis is a product of collaborative work at the Mitsubishi Electric Research Laboratories, Cambridge, MA (MERL), and enlightening interactions with the MSN Safety Team of Microsoft Corporation, Network and Spam Solutions Team of Cisco Systems, Gmail Team of Google, and Anti-Spam Team of Yahoo!. Preliminary results of this thesis have appeared on Slashdot and presentations at the *MIT Spam Conference (2005, 2004, 2003)*; Cisco Systems (2005); *Second Conference on Email and Anti-Spam (CEAS 2005), Stanford University*; *The Fourth IEEE International Conference on Data Mining, Brighton, UK (ICDM04)*; *8th European Conference on Principles and Practice of Knowledge Discovery in Databases, Pisa, Italy (PKDD 2004)* and *1st International Workshop on Peer2Peer Data Management, Security and Trust, Zaragoza, Spain, (PDMST04).*