VIRGINIA JOURNAL OF LAW & TECHNOLOGY

FALL 2006 UNIVERSITY OF VIRGINIA VOL. 11, NO. 10

An Empirical Study of the Patent Prospect Theory: *An Evaluation of Anti-Spam Patents*

MARTIN CAMPBELL-KELLY†
PATRICK VALDURIEZ‡

ABSTRACT

This article is an empirical study of Edmund Kitch's patent prospect theory, based on an evaluation of anti-spam patents. The article seeks to contribute to the wider debates about the benefits and costs of patents in general, and of software patents in particular. The anti-spam patent prospect has been selected for the study because the anti-spam sector of the software industry was formed in the last decade, when software patents were routinely available, and the sector is sufficiently small such that all known anti-spam patents have been individually evaluated and a majority of the key players in the industry examined, including open-source firms.

The article concludes that the anti-spam patent prospect has facilitated the orderly development of the industry by encouraging diverse anti-spam solutions, discouraging technology appropriation by reverse engineering, and fostering a market for anti-spam technologies. The article also argues that the patent environment has helped, rather than hindered, small, innovative, first-mover firms in protecting and marketing their technologies. The article also shows that the proprietary and open-source modes of software delivery are not mutually exclusive, that collaboration takes place between these modes, and both can derive benefit from the patent system.

© 2006 Virginia Journal of Law & Technology Association, *at* http://www.vjolt.net. Use paragraph numbers for pinpoint citations.

† Martin Campbell-Kelly is a professor in the Department of Computer Science, Warwick University. He is an historian and computer scientist with a special interest in the history of information processing. His most recent book is From Airline Reservations to Sonic the Hedgehog: A History of the Software Industry (MIT Press 2003).

‡ Patrick Valduriez is a director of research at INRIA, the national center for computer science in France, working on data management in distributed systems. His most recent book is Principles of Distributed Database Systems (co-authored with Professor Tamer Özsu) (Prectice Hall 2d ed. 1999). The authors thank Anne Layne-Farrar and Albert L. Nichols of LECG for helpful comments and suggestions, and acknowledge the research support of LECG.

TABLE OF CONTENTS

I. INTRODUCTION

¶ 1 In the last 15 years there have been many legal, policy, and technical articles about whether or not software patents should be permitted. Today, however, there are well over 100,000 software patents in existence, so it is perhaps time to move the debate

forward.[1] In this article, we accept software patents as a fact of life, and try to evaluate whether they are in practice helping or hindering the industry.

¶ 2 In order to frame our discussion, we evaluate anti-spam patents as a technology "prospect." The concept of a technology prospect was first proposed by Edmund Kitch in 1977.[2] At the time that Kitch was writing, the "reward theory" had dominated economic discussions of the patent system for many years[3]. The reward theory posited that a patent served to motivate inventors by rewarding them with a temporary monopoly on an invention. This, *inter alia*, would enable the inventor to commercialize the invention without fear of rapid imitation, providing the inventor "breathing space" to assemble the resources needed for commercialization, as well as a tradable instrument in the form of a patent that would facilitate negotiations for financial and other resources. There are, of course, arguments both for and against patents in the context of the reward theory. For example, some commentators have argued that society can lose out because an inventor with a monopoly may not be able to fully exploit the opportunity—thereby restricting output and increasing prices.[4]

¶ 3 In their discussions about the various theories of patents, Mazzoleni and Nelson have argued that an implicit feature of reward-type theories is that they apply to narrow domains of invention where "there is basically one commercial product at the end of the rainbow."[5] Nice, clear-cut examples of such inventions would include King Gillette's safety razor and Lazlo Biro's ballpoint pen. This is not to say that these inventions cannot be improved upon (there are plenty of patented safety razor and ballpoint pen improvements); rather, these inventions are narrow developments that do not open up major areas of innovative activity.

¶ 4 The prospect theory addresses the situation where "an initial discovery or invention is seen as opening up a whole range of follow-on developments or inventions."[6] Inventions such as antibiotics, semiconductors, or speech recognition technologies are different in degree than safety razors or ballpoint pens.[7] The former are

---

1 Seth Shulman *Software Patents Tangle the Web*, TECH. REV. , Mar./Apr. 2000, at 71, *available at* http://www.technologyreview.com/BizTech/wtr_12074,311,p1.html. For a detailed analysis of the issuance of software patents, see Stuart J. H. Graham & David C. Mowery, *Intellectual Property Protection in the U.S. Software Industry*, *in* NAT'L ACADEMIES PRESS, PATENTS IN THE KNOWLEDGE-BASED ECONOMY, 219-58 (Wesley M. Cohen ed., 2003).

2 Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977).

3 *Id.* at 266.

4 Roberto Mazzoleni & Richard R. Nelson, *Economic Theories about the Benefits and Costs of Patents*, 32 J. ECON. ISSUES 1031, 1042 (1998). *See also* Roberto Mazzoleni and Richard R. Nelson, *The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate*, 27 RES. POL'Y 273, 273-84 (1998).

5 Mazzoleni & Nelson, *Economic Theories about the Benefits and Costs of Patents*, *supra* note 4, at 1042.

6 *Id.*

7 ROSS KNOX BASSETT, TO THE DIGITAL AGE: RESEARCH LABS, START-UP COMPANIES, AND THE RISE OF MOS TECHNOLOGY (Johns Hopkins Univ. Press 2002) (describing the development of semiconductors); STUART B. LEVY, THE ANTIBIOTIC PARADOX: HOW THE MISUSE OF ANTIBIOTICS

technological prospects of the greatest importance to society and so broad that they could not be fully exploited by a single inventor or even by a single firm. Adherents of the prospect theory believe that the patent system "permits the development of the full range of possibilities to proceed in an orderly fashion."[8]

¶ 5 In explaining the prospect theory, Kitch analogized the patent system with the mineral claim system developed in the American West in the second half of the nineteenth century.[9] This system enabled a person who discovered mineralization on public land to file a claim which gave him exclusive mining rights. Thus, in the words of Kitch, the claim system created "incentives for prospectors to pack their burros and walk off into the desert in search of mineralization."[10] Kitch noted that, far from restricting output, the claim system "tended to generate the socially optimum level of investment in prospecting."[11] Kitch urged students of the patent system to see it as a form of claim system for an invention prospect, rather than as a monopoly conferred on an individual inventor that restricted output.

¶ 6 Kitch argued the case for granting broad patents on a new prospect. Unless a broad patent was granted on the prospect, there would be "races for specific targets of opportunity and general over-fishing in the prospect pond."[12] However, Mazzoleni and Nelson—no fans of the patent system—have pointed out that such broad patents could have the adverse social cost of reducing the number of diverse inventors working the prospect.[13] Thus, from the perspective of prospect theory, the patent system should award patents that are sufficiently broad so as to discourage overlapping inventions (which would result in wasteful over-fishing), but narrow enough so as to encourage diverse inventors to work the prospect.

¶ 7 In their article, Mazzoleni and Nelson noted that there is a lack of empirical studies on which to base fruitful discussions about the competing theories of the benefits and costs of patents (outside of the reward theory).[14] We therefore offer our study as a modest contribution to the wider debate about the benefits and costs of patents.

¶ 8 Our article takes advantage of what might be called a "natural experiment." By focusing on anti-spam technology, we believe we can discount some of the uncertainties that can render discussions about software patents inconclusive. Anti-spam technologies

---

DESTROYS THEIR CURATIVE POWERS (Perseus Publishing 2d ed. 2002) (discussing the evolution of antibiotics); Savitha Srinivasan & Eric Brown, *Is Speech Recognition Becoming Mainstream?*, IEEE COMPUTER, Apr. 2002, at 38-41 (providing "a comprehensive view of speech recognition research together with a practical perspective on speech recognition applications.").

[8] Mazzoleni & Nelson, *Economic Theories about the Benefits and Costs of Patents*, *supra* note 4, at 1042.

[9] Kitch, *supra* note 2, at 271-75.

[10] *Id.* at 274.

[11] *Id.*

[12] Mazzoleni and Nelson, *Economic Theories about the Benefits and Costs of Patents*, *supra* note 4, at 1042.

[13] *Id.*

[14] *Id.* at 1044-46.

are very new because the spam problem itself only surfaced in the first half of the 1990s. Hence, the anti-spam industry developed when patents were an established feature of the software landscape, rather than being imposed on a pre-existing industry structure shaped in the non-patent era. There are not many sectors of the fifty year old software industry for which this would be true. Also, developing anti-spam software is a sizable activity; there are dozens of firms, of which more than forty are analyzed in this article. Thus we can draw some conclusions that go well beyond anecdotes and particular cases. At the same time, the anti-spam industry is small enough that we have been quite exhaustive in our research. We have examined in detail the entire cannon of the 100-plus issued anti-spam patents that we could locate, and we have tracked technology transfer at the level of the individual firm.

¶ 9 The article is organized as follows. In section II, we establish our context with a brief history of the spam problem. In section III, we describe a taxonomy for the principal anti-spam technologies and argue that these technologies constitute a technology prospect.

¶ 10 In the next three sections, we describe in detail the "mining" of the technology prospect. Section IV provides an overview of the three principal sectors of the industry. Section V explores the licensing activity between major firms within each sector and across sectors. In section VI, we look at open-source anti-spam solutions.

¶ 11 In section VII, we conclude by directly addressing the following questions (as they relate to the anti-spam prospect), several of which are of concern to the communities hostile to software patents:

1. Is the patent system achieving an orderly development of the anti-spam prospect?
2. Are big firms with strong patents excluding small firms?
3. Are broad patents blocking competitors?
4. Are there too many narrow patents, resulting in over-fishing of the prospect?
5. Is the patent system facilitating a market for anti-spam technology with reasonable transaction costs?
6. Is open-source activity being threatened by the patent system?

¶ 12 In answering these questions, first, we find that the anti-spam industry is competitive and flourishing, and that patents play a positive role in this enterprise. Second, there is no evidence that small firms are being excluded by big firms with strong patents; firms generally show a propensity to license technology to competitors. Third, we found only two examples of broad patents with blocking potential, and in both cases the long-term impact has been minimal. Fourth, because of the rapid evolution of spam and anti-spam technologies, the prospect constantly expands and there is no evidence of over-fishing. Fifth, there is a strong market for anti-spam technology, and technology delivery mechanisms serve to minimize patent and other transaction costs. Finally—and surprisingly in light of the reported hostility between the advocates of open-source

software and proprietary vendors—we find there is in practice a productive coexistence between the two sectors.

II. A BRIEF HISTORY OF SPAM

¶ 13 Spam is a phenomenon of Internet e-mail. It did not exist with commercial e-mail services or consumer networks in the 1980s. Although e-mail as a technology has existed since the mid-1960s, it was not until personal computers arrived on people's desktops in the mid-1980s that e-mail services took off.[15] The first mover was MCI, which offered an e-mail service in 1983.[16] In the next two or three years it was imitated by the other major telecommunications companies—AT&T, ITT, Sprint, and Western Union.[17] E-mail services were also provided by consumer networks—such as CompuServe, Delphi, The Source, and Prodigy—which also took off in the second half of the 1980s.[18]

¶ 14 Spam was not possible in the aforementioned networks for two reasons. First, e-mail was expensive. For example, MCI Mail charged a minimum of 45 cents a message and AT&T Mail charged a minimum of 40 cents a message; other providers charged by connect time rather than individual messages, but the effective costs were similar.[19] In consumer networks, users were usually given an e-mail allowance in their monthly subscription, with an incremental charge for additional messages. For example, CompuServe subscribers were allowed 60 free messages per month, with extra messages being charged at 15 cents each;[20] Prodigy allowed 30 free e-mails and charged 25 cents for additional messages.[21] Certainly, some unsolicited commercial e-mail existed on private networks—often from the service provider or a trading partner—but its impact on users was minimal.[22] The second reason that spam did not exist was that all the commercial networks were centrally controlled, so it was easy to enforce acceptable use policies by disciplining or terminating subscribers who broke the rules.

¶ 15 In the 1980s, it was generally only possible for a subscriber to communicate with members of the same network. In the early 1990s, however, providers gradually integrated their services by providing gateways and protocol translations between services. For example, by 1994, CompuServe users could communicate with MCI Mail, AT&T Mail, SprintMail, Western Union, and the Internet.[23] It was with the connection to

---

[15] Anthony Ralston, Edwin D. Reilly & David Hemmendger, *Electronic Mail*, in *Encyclopedia of Computer Science*, 637, 637-42 (4th ed. 2000).

[16] PHILIP L. CANTELON, THE HISTORY OF MCI, 1968-1988: THE EARLY YEARS 368-83 (Heritage Press 1993).

[17] *Easy E-mail: Modern Electronic Mail Products and Services*, NETWORK WORLD, Dec. 5, 1988, at 35-36, 45, 47.

[18] Alfred Glossbrenner, *The Little Online Book A Gentle Introduction to Modems, Online Services, Electronic Bulletin Boards, and the Internet* (Peachpit Press 1995).

[19] *Id.* at 36.

[20] CHARLES BOWEN, COMPUSERVE FROM A TO Z 315-16 (Random House Info. Group 2d ed. 1994).

[21] JOHN L. VIESCAS, THE OFFICIAL GUIDE TO THE PRODIGY SERVICE 183 (Microsoft Press 1991)

[22] *Id.* at 182.

[23] Bowen, *supra* note 20, at 309-15.

the Internet that the flood of spam began.

¶ 16 In the 1980s, the Internet had been a non-commercial academic and research network, initially funded by the Department of Defense and later by the National Science Foundation.[24] The Internet used an e-mail protocol known as SMTP (Simple Mail Transfer Protocol), which had been designed in the context of a decentralized, cooperating community of like-minded users. Consequently the protocol had security flaws that only became obvious with the commercialization of the Internet in the early 1990s.

¶ 17 One of the most serious security flaws in SMTP was that it permitted a mail server to operate as an "open relay," so that any user could insert mail into any server on the Internet.[25] This unrestricted access to mail servers was analogous to the availability mailboxes on suburban street corners, with the important distinction that it is not possible to drop a million letters into a mailbox free of charge. Although well run mail servers no longer operate as open relays, at any one time there are enough poorly configured servers to provide a conduit for spam.

¶ 18 Spam arose on the Internet for reasons complementary to those which had made it virtually non-existent on the early commercial networks. First, Internet e-mail is free—or at least it is unmetered—so that once a user has access to the network, it is possible to disgorge thousands or millions of e-mails with no incremental cost. Second, the Internet has no centralized authority, so the main sanctions against spammers are moral reprobation and various forms of vigilantism such as "denial-of-service attacks" or "e-mail bombs."[26] It is true that most Internet service providers (ISPs) now have an acceptable use policy, but the decentralized architecture of the Internet makes it impossible to enforce such practices on all rogue operators.

¶ 19 In the early 1990s, before the "Information Superhighway" had morphed into the Internet, unsolicited commercial e-mail appeared to be a legitimate business opportunity to people unfamiliar with "netiquette". Most famously, two lawyers, Laurence Canter and Martha Siegel, sparked a controversy by the unsolicited advertising of their services and by publishing a book entitled *How to Make a Fortune on the Information Superhighway*, which espoused spamming and described their techniques.[27] They also formed a

---

[24] JANET ABBATE, INVENTING THE INTERNET (MIT Press 1999).

[25] SPAMMER-X, THE SPAM CARTEL: TRADE SECRETS FROM THE DARK SIDE 36-39 (Jeffrey Posluns ed., Syngress 1st ed. 2004) [hereinafter SPAMMER-X].

[26] A denial-of-service attack is action taken by hackers that causes a computing resource to be unavailable to legitimate users, for example by accessing a website with such a frequency that the response time degrades unacceptably. *See* MARK EGAN, EXECUTIVE GUIDE TO INFORMATION SECURITY: THREATS, CHALLENGES, AND SOLUTIONS 196 (Addison-Wesley Prof'l 2004). An e-mail bomb deluges an e-mail server with so much mail that it ceases to function. *See* ALAN SCHWARTZ & SIMSON GARFINKEL, STOPPING SPAM: STAMPING OUT UNWANTED EMAIL AND NEWS POSTINGS 104 (O'Reilly 1998).

[27] LAURENCE CANTER & MARTHA SIEGEL, HOW TO MAKE A FORTUNE ON THE INFORMATION SUPERHIGHWAY: EVERYONE'S GUERRILLA GUIDE TO MARKETING ON THE INTERNET AND OTHER ON-LINE SERVICES (HarperCollins 1st ed. 1994).

company, Cybersell, in order "to help others do the same thing."[28]

¶ 20 The best documented early spamming operation was Cyber Promotions Inc. which was established in 1994 by an Internet "newbie," Stafford Wallace, who saw unsolicited commercial e-mail as a legitimate business.[29] It is likely that Cyber Promotions' rise and fall was similar to, if not more dramatic than, many other spamming operations of the 1990s. In its heyday, Cyber Promotions was reportedly sending 25 million e-mail messages a day on behalf of 11,000 clients.[30] However, Cyber Promotions' ISP experienced floods of complaining e-mail, so the account was terminated. This happened again with other providers, so that Cyber Promotions eventually established its own Internet presence. The Cyber Promotions Internet service was then subjected to denial-of-service attacks from Internet vigilantes. The firm also became embroiled in lawsuits with AOL, CompuServe, Sprint, Prodigy, and EarthLink, over alleged activities such as theft of service. Apparently exhausted from dealing with such opposition, Wallace announced his retirement from Cyber Promotions in 1988 and offered his services as an expert witness in anti-spamming cases.[31]

¶ 21 By about 2000, the great majority of ISPs had introduced no-spamming acceptable use policies, so that open, arguably legitimate, spam services such as CyberSell and Cyber Promotions became unworkable.[32] As a result, spamming operations became clandestine, sometimes illegal, and increasingly operated offshore. Today, the spam "industry" is extremely fragmented, consisting of individuals or very small partnerships either actively spamming or providing spamming services, software tools, or e-mail address lists.[33] A spam operation needs four things to make it viable. First, the *sine qua non* is Internet access. Although spamming is no longer legal in the United States,[34] there are overseas operations that provide bulk e-mail facilities.[35] In any case, an accomplished hacker can find open relays or readily abuse an ISP for a few hours before their account is terminated. Second, spammers need a financial incentive. Spammers typically market products (such as medications, pirated software, or pornography) directly to consumers, or they act as commission agents for third parties offering such products (the Internet bristles with such opportunities). Third, spammers need a specialized e-mail program for bulk e-mailing. Inexpensive bulk e-mail programs,

---

28 *Id.* at 31.

29 For a history of Cyber Promotions, see SCHWARTZ & GARFINKEL, *supra* note 26, at 25-29, 177-83; BRIAN MCWILLIAMS, SPAM KINGS: THE REAL STORY BEHIND THE HIGH-ROLLING HUCKSTERS PUSHING PORN, PILLS, AND @*#?% ENLARGEMENTS 21-26 (O'Reilly 2005).

30 MCWILLIAMS, *supra* note 29, at 21; SCHWARTZ & GARFINKEL, *supra* note 26, at 180.

31 SCHWARTZ & GARFINKEL, *supra* note 26, at 182.

32 *Id.* at 136-138, 175.

33 For the best accounts of the contemporary spam industry, see SPAMMER-X, *supra* note 25, and MCWILLIAMS, *supra* note 29, at 21-26.

34 CAN-SPAM Act of 2003, Pub. L. No. 108-187, 117 Stat. 2699 (2003).

35 *See, e.g.*,SPAMMER-X, *supra* note 25, at 39-40 (citing Black Box Web Hosting as an example of an overseas bulk e-mail hosting facility). Due to denial-of-service attacks, the Internet storefronts of these operations are difficult to access.

which possibly have legitimate uses, are routinely available on the Internet.[36] Lastly, the spammer needs a list of e-mail addresses. These can be obtained from legitimate vendors that specialize in targeted permission-based e-mail lists, but more likely from illegitimate sources—often themselves advertising their services by spam. Typically, e-mail lists for spamming contain millions of e-mail addresses and are completely untargeted.[37] It is also possible for spammers to use a "harvesting" program to scan web sites for e-mail addresses; these are routinely available on the Internet.[38]

¶ 22 Most recently, "Trojan horses" or "spambots" have added a new twist to the ever-evolving spam story. A spambot is a computer virus or worm that resides on a PC and sends spam through the user's own mail server. As this technology has matured, like any other, its price has dropped:

¶ 23 In the beginning the cost was high. For a 200-client Botnet you could expect to pay up to $1,000.00, but as more worms propagated, the price dropped. Soon, "exclusive" control over 1,000 hosts could be bought for as little as $500.00. Now, exclusive control over a single zombie can sell for as little as 10 cents![39]

¶ 24 Anti-virus software on a PC is effective against this form of infestation, but there are still many unprotected machines.

¶ 25 In January 2004, the Controlling the Assault of Non-Solicited Pornography and Marketing (CAN-SPAM) Act came into force in the United States.[40] Many other countries have enacted similar laws.[41] The CAN-SPAM Act is a well-meaning statute that has already diffused good practice in the use of commercial e-mail by legitimate firms.[42] For example, the Act requires that potential e-mail recipients "opt in" to receive e-mail, their addresses cannot be passed to third parties without express consent, and each e-mail must have a simple method for unsubscribing from the mailing list. In addition, commercial e-mail cannot be sent anonymously, and the falsification of e-mail "headers" is an offense.[43] The content cannot be deceptive or misleading, and it cannot be designed to bypass anti-spam filters (for example, by spelling "Viagra" as "V1agra"). The penalties are severe: up to a $250 fine for each e-mail message, with a maximum liability

---

[36] A popular shareware bulk e-mail program is DarkMailer—which offers "anonymous bulk e-mail software for marketing"— downloadable from many shareware sites. *See* SPAMMER-X, *supra* note 25, at 22-23. *See also* Send-Safe.com – Professional Tools for Bulk Mailers, http://www.send-safe.com (last visited Apr. 18, 2006), which offers several such products.

[37] SPAMMER-X, *supra* note 25, at 17.

[38] *Id.* at 76-86.

[39] *Id.* at 43.

[40] CAN-SPAM Act of 2003, Pub. L. No. 108-187, 117 Stat. 2699 (2003).

[41] David E. Sorkin, *Spam Laws*, *available at* http://www.spamlaws.com (last visited Dec. 28, 2006).

[42] *See* ARIAL SOFTWARE, 2004 CAN-SPAM B2C COMPLIANCE AUDIT, (2004), *available at* http://www.arialsoftware.com/whitearticles/CANSPAMComplianceAudit2004.pdf (reporting results of a secret audit of 1,057 organizations for CAN-SPAM Act compliance).

[43] Federal Trade Commission, *The CAN-SPAM Act: Requirements for Commercial Emailers*, Apr. 2004, *available at* http://www.ftc.gov/bcp/conline/pubs/buspubs/canspam.pdf.

of $2 million, and aggravated damages that can triple the fine.[44] Numerous spammers have been prosecuted since the Act was passed.[45] However, although compliance by legitimate operations has been good, spam has not diminished in volume.[46] It has simply been driven further underground. The decentralized architecture of the Internet and the security flaws in SMTP make it impossible to enforce the provisions of the CAN-SPAM Act. And in any case, spam can still prosper legally beyond the jurisdiction of anti-spam laws.

¶ 26 Thus, the current solutions to the spam problem do not lie in the realms of law, moral admonishment from the Internet community, or vigilantism. These solutions are technological, but still provide only temporary fixes.

III. PATENTS AND THE ANTI-SPAM PROSPECT

¶ 27 In this section we seek to describe the extent and nature of the anti-spam prospect. The prospect covers a considerable range: it is a broad inventive space of more than a hundred—and potentially several hundred—distinct patented inventions (and many non-patented inventions). We do not propose to describe these inventions individually, but instead to use a taxonomic organization as a way of viewing the invention space. We have mapped the 100-plus issued spam patents onto this taxonomy and thereby shown that there is a comprehensive colonization of the invention space and a strong correspondence between real innovations and patents.

¶ 28 Such a static analysis does not fully capture the nature of the constantly evolving anti-spam prospect. As soon as one form of spam becomes vulnerable to an anti-spam technique, it mutates in order to circumvent detection; in turn, anti-spam techniques evolve to address this new form of spam. It is an endless cycle of innovation. This is perhaps a more fluid situation than Kitch envisaged in his classic article, but we think that some other technology prospects may share this dynamic character.[47]

A. Taxonomy of Anti-spam Technology

¶ 29 E-mail systems use a "client-server" architecture. An e-mail client (or mail client) is an application that runs on a desktop PC that lets the user send, receive, and organize e-mail within various folders. An e-mail server (or mail server) is a computer that handles the storage and the transfer of messages for local mail clients (within the organization), and the exchange of messages with other mail servers (outside the organization). Spam eradication can take place inside the mail server or in the desktop client.

¶ 30 Classifications of anti-spam techniques are widely used in the professional anti-

---

[44] CAN-SPAM Act, Sec. 7f3.

[45] Federal Trade Commission, *Effectiveness and Enforcement of the CAN-SPAM Act: A Report to Congress*, Dec. 2005, at www.ftc.gov/reports/canspam05/051220canspamrpt.pdf

[46] Tom Zeller Jr., *Law Barring Junk E-Mail Allows a Flood Instead*, NEW YORK TIMES, Feb. 1, 2005, *available at* http://www.nytimes.com/2005/02/01/technology/01spam.html?ex=1265000400&%2338;en=7f39918007d9ca0f&%2338;ei=5088.

[47] *See, e.g.*, LEVY, *supra* note 7 (discussing antibiotics).

spam and security literature.[48] Similar classifications are also used in "white papers" from the major anti-spam vendors.[49] However, we cannot apply these classifications directly to anti-spam patents because these patents frequently use several techniques or deal with general security issues in addition to spam. Therefore, we construct a taxonomy that reflects the fundamental anti-spam techniques while also capturing the majority of anti-spam patents. We define four classes. Each of these classes can be further subdivided, as shown in Figure 1.

¶ 31 The first three classes correspond to the fundamental anti-spam techniques: Access Control (AC), Content Filtering (CF), and Sender Verification (SV). We also have a fourth class, Mail Management (MM), which deals with the management of the resources involved in an anti-spam solution. We define the four classes as follows:

¶ 32 ***Access Control*** is a set of techniques that determines whether an e-mail message is legitimate, for example, by identifying whether it comes from a known spammer or whether the identity of the sender appears to be forged.

¶ 33 ***Content Filtering*** consists of inspecting the text or other attributes of a message to determine if it is likely to be spam.

¶ 34 ***Sender Verification*** is a method whereby the receiver of a message requires the sender to perform some action to prove that the message is not spam. Most methods require human intervention so that an automated system cannot provide a legitimate response.

¶ 35 ***E-mail Management*** includes the technologies that deal with the efficient and automated management of the resources involved in an anti-spam solution. These techniques integrate the anti-spam techniques described above in a more complete system and in a variety of ways.

B. The Anti-spam Prospect

¶ 36 We have analyzed the 100-plus issued anti-spam patents and have assigned them to the four classes.[50] Our results are illustrated in Figure 2. We highlight two observations from this analysis. First, there appears to be a good spread of patents in each of the four taxonomic classes. This distribution suggests that one effect of patents may be to foster a diversity of approaches to the spam problem, because in order to secure a patent, an invention must occupy a distinct part of the invention prospect. Second, the distribution of patents has a temporal dimension. Thus, while Content Filtering has been, and remains, the most important category of anti-spam patents, since 2000, much

---

[48] *See, e.g.*, ROBERT HASKINS & DALE NIELSEN, SLAMMING SPAM: A GUIDE FOR SYSTEM ADMINISTRATORS 3-7 (Addison-Wesley Prof'l 2004); DANNY GOODMAN, SPAM WARS: OUR LAST BEST CHANCE TO DEFEAT SPAMMERS, SCAMMERS AND HACKERS 203-25 (Select Books 2004).

[49] See, *e.g.*, PARIS TRUDEAU ET AL., SURF CONTROL, INC., MAJOR TECHNIQUES FOR CLASSIFYING SPAM (2004), *available at* http://www.surfcontrol.com/uploadedfiles/general/white_articles/4ClssfySpm_Apr03.pdf.

[50] See Appendix A for our detailed analysis.

attention has been given to Access Control and Mail Management, as reflected in the evolving anti-spam industry discussed in the following sections. We also note that Sender Verification appears to have blossomed and died; the rise and fall of vendors using this technique is reflected in the industry's evolution. These findings indicate that anti-spam patents correspond to real industry trends and real research activities.

¶ 37 As noted above, one shortcoming of Kitch's mineral claim analogy with respect to the anti-spam prospect is that the latter is neither static nor finite like a mineral prospect. We argue that there are actually two co-evolving prospects—spam and anti-spam. Although spamming is now illegal in the United States, like organized crime, it is nonetheless a significant economic activity.[51] Just as there is a literature on anti-spam, there is a subversive pro-spam literature on the Internet,[52] and there is even an openly published how-to manual for spammers.[53] There is, in effect, an arms race between anti-spammers and spammers:

¶ 38 It's all a race against time—spammers versus anti-spam groups. For every technique spammers come up with to send spam, anti-spam groups come up with a way to block it. And for every technique anti-spam groups create to block spam, spammers come up with a way to bypass it. In the end, no one really wins. So much spam is sent daily that if filters caught 99 percent of it there would still be millions of dollars made from the 1 percent that is delivered. . . . Spam has become an odorless, tasteless gas—undetectable, untraceable, and penetrating every inch of the cyber-connected world. For a spammer, it is all about sending the spam at any cost; there is no room for guilt or remorse in how you send it.[54]

¶ 39 Thus, as soon as a particular solution to a spam problem is found, spammers either invent a counter technology or move on to another area of their prospect. Anti-spam (and anti-virus) technologies bear an intriguing similarity with antibiotic therapies. In the case of antibiotics, the constant Darwinian evolution of bacteria promises that the battle will never be over.[55] With spam and anti-spam, the two co-evolving prospects suggest that this battle will also never be over.

¶ 40 The development of Content Filtering illustrates how spam and anti-spam co-evolve. In the dawn of the anti-spam era, the "lexical analysis" of spam was one of the earliest techniques.[56] In this technique, incoming mail containing a suspicious word such as "Viagra" would be quarantined in a spam folder. Spammers quickly learned to conceal such keywords by disguising them as "V1agra" or "V_I_A_G_R_A."[57] It was impossible

---

51 MCWILLIAMS, *supra* note 29, at xi.

52 *See, e.g.*, SPAM LINKS, PRO-SPAM?, *available at* http://spamlinks.net/prospam.htm (last visited Apr. 22, 2006) (contains a list of pro-spam links).

53 SPAMMER-X, *supra* note 25.

54 *Id.* at 30.

55 *See* LEVY, *supra* note 7, at 71-114 (describing microbial adaptation and evolution).

56 *See* SCHWARTZ & GARFINKEL, supra note 26, at 74-80 (discussing the use of "filters" for stopping spam). *See also* U.S. Patent No. 5,377,354 (issued Dec. 27, 1994) (describing a patent for a "[m]ethod and system for sorting and prioritizing electronic mail messages").

57 GOODMAN, *supra* note 48, at 206-07.

to list all the possible spelling variations by which a word could be disguised. In the next generation of filters, the spam problem was tackled by considering whether or not the mail contained words that looked as though they might be intended to bypass a spam filter (for example, by containing words that contained unusual combinations of alphabetic and non-alphabetic symbols). Again, spammers quickly adapted to such filters; for example, by sending an advertisement in the form of an image, which filters cannot detect.[58] Additionally, the use of the HTML authoring language in e-mail (the same language that is used to create web pages) opened a completely new area of the spam prospect, by enabling techniques such as "web bugs."[59]

¶ 41 The result of all these resistance measures against anti-spam was to make lexical techniques largely obsolete. Instead, probabilistic or Bayesian analysis offered a new approach. This technique is relatively new, the first patent being issued to Microsoft in 2000.[60] In Appendix B, we define three distinct approaches to Bayesian analysis; in principle, this could introduce another tier to the taxonomy of Figure 1. Since these patents have not actually been issued, we cannot comment on their validity. However, there is no reason to suppose that the great majority of them will not ultimately be issued.

IV. STRUCTURE OF THE ANTI-SPAM SOLUTIONS INDUSTRY

¶ 42 In this and the following section, we describe the mechanisms of technology transfer and licensing in the anti-spam solutions industry. We begin with an overview of the structure of the industry.

¶ 43 There are three main sectors in the industry: software products, appliances, and hosted services. In 2003, the worldwide market for anti-spam solutions was $300 million, of which about two-thirds was for software products, and the remainder divided approximately equally between appliances and hosted services.[61] IDC predicted that the market would grow at a rate of 50 to 60% for the following two years, slowing substantially as the market saturated toward the end of the decade.[62] Within the three sectors, the relatively new markets for appliances and hosted services were predicted to have much higher growth rates than that for software products.[63]

A. Software Products

¶ 44 Anti-spam software products are programs supplied for installation on mail servers and clients. On a mail server, the anti-spam software filters the incoming e-mail stream, rejecting known spam and flagging suspected spam (for example by inserting a

---

58 GOODMAN, *supra* note 48, at 153-54.

59 *Id.* at 149-50 ("Web bugs" are programs that "[i]nvisibly [track] an e-mail recipient's viewing or previewing of a message, and [verify] your e-mail address as being active.").

60 U.S. Patent No. 6,161,130 (issued Dec. 12, 2000)

61 INTERNATIONAL DATA CORP., WORLDWIDE ANTI-SPAM SOLUTIONS 2004-2008 FORECAST AND 2003 VENDOR SHARES 7 (2004).

62 *Id.*

63 *Id.*

marker such as **SPAM** in the subject line). For organizations and ISPs with high volumes of e-mail, spam filtering at the server is increasingly a necessity. Spam often arrives in bursts, with the same message being sent to many users. Stopping broadcast spam at the server, rather than at the client, simultaneously reduces bandwidth consumption and the time spent by individuals weeding spam from their e-mail.

¶ 45 The single biggest problem with spam filtering is the "false positive"—the rejection of legitimate mail as spam. There is an obvious trade-off between aggressive filtering and the risk of producing false positives. A majority of anti-spam patents address this problem in novel ways. Firms compete on two spam elimination metrics: the percentage of spam detected and the number of false positives. For example, Mail-Filters, a prominent supplier, guarantees to eliminate at least 95% of spam with less than 1 in 100,000 false positives.[64] Firms also compete on the processing load placed on a server, because some filtering algorithms have the disadvantage of being computationally intensive. They likewise compete on the ease and flexibility with which the software can be tuned to the local environment (some environments, such as universities, have a much higher propensity for spam e-mail than others).

¶ 46 Because *any* false positives are unacceptable to many users, server-side filtering can only be as aggressive as the consensus in an organization allows. Hence, there is also a need for anti-spam solutions on the client side. Typically, these programs divert suspected spam from the user's inbox to a "quarantine" folder. Programs allow for a good deal of customization by individual users, for example by the creation of personal blacklists and whitelists. There are roughly a score of established brand names in client-side anti-spam software, such as McAfee's SpamKiller, MailShell's SpamCatcher, and MailFrontier's Matador. Consumer PC magazines regularly feature product comparisons.[65] Client-side anti-spam software is the least capital intensive sub-sector of the industry, and there are at least a hundred shareware products that users can download for a free trial.[66] It is very difficult to evaluate these products, and most are unlikely to survive long term. It is possible that some of these products infringe patents, but since shareware solutions are an insignificant part of the market, patent owners are unlikely to either know or care about such infringement.

¶ 47 By contrast, the supply of server-side anti-spam software is much more capital intensive, as it requires active selling to corporate users, customer support, and much higher R&D investments in novel anti-spam technologies, either through internal development or by licensing. Corporate buyers are kept well informed about competing products through features in professional magazines such as *Internet Security* and from

---

[64] MAIL-FILTERS.COM, INC., SPAM FILTERS NEED THE HUMAN TOUCH 6 (October 2004) *at* http://www.mail-filters.com/Technology/Human%20Analysis.pdf. Incidentally, as indicated by the title of this white article, Mail-Filters uses human spam recognition to augment automated systems. Although firms compete on the basis of proprietary or patented technology, it is by no means the only factor.

[65] *See, e.g.*, Cade Metz, *Personal Anti-spam Tools*, PC MAGAZINE., Feb. 25, 2003, http://www.pcmag.com/article2/0,4149,844251,00.asp.

[66] *See, e.g.*, Shareware List, http://www.shareware-list.com/category-7-3-1.html (last visited Sept. 23, 2007); Soft Hypermarket, http://www.softhypermarket.com/Anti-spam-category_160_1.html (last visited Sept. 23, 2007).

reports by market analysts such as IDC, Ostermann Research, and the Radicati Group. The market for corporate anti-spam products is moderately concentrated; in 2003, a dozen firms supplied 80% of the market, and the top six firms accounted for over half of the market.[67]

B. Appliances

¶ 48 The newest and fastest growing anti-spam sector is appliances. In 2003, appliances accounted for $50 million of the $300 million worldwide market for anti-spam solutions, and the sector is growing at twice the rate of that for anti-spam software products.[68] An anti-spam appliance is sometimes known as an e-mail firewall, and it resides between a mail server and a conventional Internet firewall. Appliances are relatively expensive, priced at around $10,000 for a medium-sized enterprise, and up to $100,000 for a major organization with a high volume of e-mail.[69]

¶ 49 A major advantage of an appliance is that it requires minimal set up—like an Internet firewall, it is physically inserted, plug-and-play, into the network. One manufacturer claims (perhaps optimistically) that its devices can be installed in 15 minutes.[70] An appliance consists of a computer (typically an Intel-based server) loaded with anti-spam software. Because appliances are specialized to the single task of spam detection, the real-time processing of a very high volume of e-mail is possible. By contrast, anti-spam software on a conventional server can create an unsupportable processor load, ultimately requiring the purchase of additional servers. The latest appliances also include anti-virus and other Internet security solutions. By rejecting spam at the network edge, appliances reduce the volume of network traffic in the enterprise. Appliances make use of both proprietary anti-spam techniques and bought-in anti-spam software products installed in the appliance; in effect, many appliance manufacturers are conventional "turnkey" systems integrators.[71]

¶ 50 The appliance sector is more capital intensive and more concentrated than the software-products sector. Unlike the anti-spam software vendor, the appliance maker has additional burdens of manufacture and on-site maintenance. In 2003, the top four firms supplied 70% of the market.[72] However, at the time of writing, several of the larger anti-spam software products suppliers—among them Symantec, Tumbleweed, Network

---

67 Calculations based on INTERNATIONAL DATA CORP, *supra* note 61, at 5-6.

68 *Id.*

69 *See, e.g.*, Logan G. Harbaugh, *Exclusive: CipherTrust, Corvigo, and MessageLabs Lighten the Spam Load*, INFOWORLD, Feb. 13, 2004, http://www.infoworld.com/article/04/02/13/07TCspam_1.html.

70 Todd R. Weiss, *Barracuda Networks Launches Anti-spam Appliance Line*, COMPUTERWORLD, Oct.13, 2003, http://www.computerworld.com/softwaretopics/software/groupware/story/0,10801,86007,00.html.

71 The turnkey concept was established in the 1970s. Turnkey suppliers sold or leased a computer loaded with application software as a self-contained system that required little or no systems administration. The most important early turnkey markets were for office automation and computer aided design. *See* MARTIN CAMPBELL-KELLY, FROM AIRLINE RESERVATIONS TO SONIC THE HEDGEHOG: A HISTORY OF THE SOFTWARE INDUSTRY 128-29 (M.I.T. Press 2003).

72 *See* INTERNATIONAL DATA CORP, *supra* note 61.

Associates, and NetIQ—are diversifying into appliances, so it is likely that concentration in the appliances sector will lessen over time.

C. Hosted Services

¶ 51 For small- and medium-sized enterprises, a hosted solution is an attractive alternative to a costly appliance or an appliance that is costly to manage, server-side software. In a hosted service, the enterprise's incoming e-mail is diverted to a remote site where it is filtered by the service provider, before being returned to the enterprise. Spam e-mail is typically flagged in the subject line or quarantined on the service provider's site. For the user, a hosted service is at least as easy to set up as an appliance, and has the lowest up-front cost of any solution; the user typically pays a service charge based on the e-mail volume.

¶ 52 Hosted services, like anti-spam appliances, can make use of both proprietary techniques and regular anti-spam software products. Hosted services firms also compete on the basis of response times, so that there is not an unacceptable processing delay, and have therefore invested in complex, geographically-distributed networks located close to their major markets. Patented innovations have also been made for the efficient redirection of e-mail streams.[73]

¶ 53 Hosted services accounted for $46 million of the $300 million anti-spam solutions market in 2003, and the sector is growing nearly as fast as appliances.[74] Hosted services are capital intensive because they require a significant infrastructure of servers, and the sector is the most concentrated in the anti-spam industry, with the top 3 vendors supplying 80% of the market in 2003.[75]

¶ 54 A variant of the hosted solution is the "verification service." This technology was one of the first anti-spam solutions in the market. In a verification service, sometimes known as a challenge-response system, incoming e-mail is redirected to a host computer where it is held in quarantine until it has been verified by the sender. Typically this involves the sender having to answer a question such as "What color is an orange?" or "How many wheels are on a car?" Only a human being could answer such a question and this would defeat an automated spamming system.[76] Challenge-response technology emerged very early in anti-spam history, and has presented significant shortcomings. For example, such technology can fail to discriminate between spam and valid automated e-mail (such as an airplane e-ticket).

¶ 55 The early market leader in challenge-response was Mailblocks Inc.,which acquired two key patents from two independent inventors for an undisclosed sum.[77]

---

73 *See, e.g.*, U.S. Patent No. 6,650,890 (issued Nov. 18, 2003) (patenting a hosting process).

74 *See* INTERNATIONAL DATA CORP, *supra* note 61.

75 *Id.*

76 These questions appear in the specification of Mailblocks-owned U.S. Patent No. 6,199,102 (issued Mar. 6, 2001) at col. 5.

77 U.S. Patent No. 6,112,227 (issued Aug. 29, 2000); U.S. Patent No. 6,199,102 (issued March 6, 2001). *See also* Declan McCullagh, *Promising Spam Blocker Stuck in Court?*, CNET NEWS.COM, May 19,

Mailblocks defended the patents vigorously, filing infringement suits in May 2003 against a direct competitor, Spam Arrest, and an ISP, EarthLink.[78] Mailblock's dispute with Spam Arrest was settled for an undisclosed sum; in the dispute with EarthLink neither company has made a statement on the outcome, although EarthLink continues to use a challenge-response system.[79] Mailblocks was acquired by AOL in August 2004, and its technology is now incorporated in AOL's e-mail service.[80]

¶ 56 Although Mailblocks and Spam Arrest continue in business, they are not major players in the anti-spam solutions sector. Verification services are a declining market, and will not be discussed further in this article. However, it is worth noting that although Mailblocks had a strong patent position, it was essentially a one-golf-club solution that was unable to deal with the rapidly evolving spam nuisance. Had it not been bought by AOL, Mailblocks would undoubtedly have had to incorporate further anti-spam technologies in order to remain a market leader. It would have been forced into technology sharing with other innovators, trading its patents for the best deal it could make.

V. TECHNOLOGY LICENSING IN THE ANTI-SPAM INDUSTRY

¶ 57 Technology licensing practices differ among the three sectors of the anti-spam industry, and therefore they need to be considered separately. In short, software product vendors compete largely on the basis of filtering efficiency, as measured by the proportion of spam detected and the number of false positives. Appliances and hosted services, on the other hand, compete only partly on the efficiency of spam elimination; factors such as ease of installation, processing speed, and response times are at least as important.

A. Software Products

¶ 58 The spam problem began to surface in the second half of the 1990s and became epidemic in the early 2000s. In the late 1990s a number of specialist firms, primarily start-ups, began to offer software solutions to eliminate spam, in a manner analogous to the anti-virus solutions that had come onto the market a few years earlier.

¶ 59 Table 1 lists the leading specialist anti-spam firms in order of their 2003 revenues. There were two main market opportunities for these firms: the development of software products for corporate and consumer end users, and the licensing of anti-spam technology

---

2003, http://news.zdnet.com/2100-9595_22-1007581.html (describing Mailblocks' claim to complete ownership of the challenge-response concept).

[78] Saul Hansell, *EarthLink is Sued by Holder of Anti-spam Patents*, N. Y. TIMES, May 8, 2003, *available at* http://oceanpark.com/webmuseum/bullshit_patents.html.

[79] Press Release, Spam Arrest, Statement by Spam Arrest LLC Concerning Mailblocks Litigation (Aug. 13, 2004), *available at* http://www.spamarrest.com/pr/releases/20040813.jsp; EarthLink SpamBlocker, http://www.earthlink.net/software/free/spamblocker (last visited Sept. 23, 2007).

[80] Scarlet Pruitt, *AOL Buys Mailblocks in Spam-Fighting Effort*, INFOWORLD, Aug. 4, 2004, http://www.infoworld.com/article/04/08/04/HNaolmailblocks_1.html.

to existing Internet-security vendors. The sale of dedicated anti-spam products is declining at the time of writing, because users are demanding more comprehensive solutions—for example, in addition to the elimination of spam from their in-boxes, users also need immunity from spam-borne viruses and other "malware." This has led firms either to develop wide-spectrum security products, or to specialize in developing anti-spam technology for licensing to larger Internet-security firms.

¶ 60 Long before the spam problem became serious, there was a significant Internet-security industry that provided integrated anti-virus solutions, security against hackers, and secure communications.[81] With the rise of the spam problem, these vendors had to incorporate anti-spam technologies into their existing products. Although it is possible for end users to buy security solutions for different threats from individual suppliers, integrating such products is notoriously difficult. While the best solution may be to acquire "best of breed" solutions from individual suppliers, most end users prefer to avoid having to deal with the inevitable system incompatibilities between products and the overhead of dealing with multiple suppliers.[82] The need for wide-spectrum Internet security products is currently driving consolidation in the industry. Table 2 lists the leading Internet-security suppliers that have integrated anti-spam technologies into their products. Vendors are ranked by their 2003 anti-spam revenues (their total revenues would, of course, be much greater). These incumbent suppliers have acquired anti-spam technologies from three sources: the acquisition of specialist anti-spam firms; by licensing technology from specialist anti-spam firms; or by developing proprietary solutions in-house.

¶ 61 The most significant firm acquisition has been that of Brightmail by Symantec in 2004; their combined revenues would put them at the head of Table 2. Other firms in Table 2 that have made anti-spam acquisitions include Clearswift, NetIQ, Network Associates, Sophos, and ZixCorp. In all of these cases, the acquiring firm gained much more than raw technology (whether patented or not). For example, the U.K. firm Clearswift, in acquiring Content Technologies in 2002, gained a successful product (MIMEsweeper), the workforce that developed and marketed it, a large customer base, and a beach-head into the U.S. market.[83] Network Associates, in acquiring Deersoft, gained a product (SpamAssassin), a development team, and a listening post into the open-source community.[84] Much the same was true for the U.K.-based Sophos, which acquired open-source ActiveState in 2004.[85]

¶ 62 Responding to the demand from Internet-security incumbents for anti-spam technologies, several specialist vendors—including Mail-Filters, MailFrontier, MailShell and Commtouch in Table 1—have made "OEM sales" their primary market. The most

---

81 *See generally* EGAN, *supra* note 26; PETER SZOR, THE ART OF VIRUS RESEARCH AND DEFENSE (Addison-Wesley Prof'l 2005).

82 EGAN, *supra* note 26, at 143-44.

83 CLEARSWIFT, CLEARSWIFT CORPORATE FACT SHEET 2005, http://www.clearswift.com/company/200504_CorporateFactsheetUK.pdf (last visited Aug. 23, 2005).

84 SpamAssassin is an open-source product. *See* discussion *infra* section VI(A).

85 *See* discussion *infra* section VI(A).

common form of technology delivery is a software development kit (SDK). The SDK consists of a program module with a set of application program interfaces (APIs) that can be easily integrated into an existing product. The market for OEM sales goes well beyond the Internet-security industry, and also includes vendors of messaging products that are vulnerable to spam-borne nuisances. For example, MailShell supplies its technology to Oracle and Stalker Software for inclusion in their collaborative software.[86]

¶ 63 Some OEM solutions include patented technology. A criticism often voiced by opponents of software patents is that the need to license many patents will make software writing infeasible. It is true that multiple software patents could interfere with developing some kinds of monolithic software, but the slowly emerging component approach to software—of which anti-spam solutions are an example—has the effect of reducing the transaction costs of patent licensing.[87] The OEM licensee pays a single price that includes software, support, and any patent royalties—whether the patents are owned by the OEM vendor, or by component suppliers further up the supply chain. In this regard, software patents may prove a stimulus for growth in the software component industry because they protect computer algorithms from appropriation by reverse engineering or cloning.[88]

¶ 64 Not surprisingly, many Internet-security vendors have highly effective R&D capabilities (attributable to their growth from ranks of anti-virus and secure communications specialists) and have developed their own proprietary solutions. Examples of firms in Table 2 following this route include SurfControl, GROUP Technologies, Trend Micro, and F-Secure.

¶ 65 Patents are widely used in the Internet-security industry, and the great majority of leading firms have a portfolio of anti-virus and secure communication patents.[89] For these firms, besides serving their normal function of IP protection, patents also facilitate technology sharing and negotiation. For example, Trend Micro has entered into cross-license agreements with Symantec, IBM, and Network Associates for anti-virus technologies.[90] Gradually, Internet-security firms are including anti-spam patents in their

---

[86] *See generally* Mailshell, About MailShell, *at* http://www.mailshell.com/datasheet_about.pdf (last visited Aug. 23, 2005) (listing some of Mailshell's OEM partners).

[87] *See* Michael S. Guntersdorfer & David G. Kay, *How Software Patents Can Support COTS Components Business*, IEEE SOFTWARE, May–June 2002, at 81-82 (addressing arguments against software patents). *Cf.* HOWARD BAETJER, JR., SOFTWARE AS CAPITAL: AN ECONOMIC PERSPECTIVE ON SOFTWARE ENGINEERING 131-35 (Wiley-IEEE Computer Society Press 1998) (discussing challenges to the component market)

[88] The software component industry has been slow to take off partly because prior IP regimens made it difficult to protect a component from appropriation. If a component was protected by trade secrecy, this was difficult to enforce; it was also legitimate for competitors to reverse engineer software of equivalent function. If code was protected by copyright, this protected only the particular implementation and the underlying concepts could be freely appropriated. Guntersdorfer & Kay, *supra* note 87, at 78-79.

[89] The exceptions tend to be open source vendors or vendors from European countries where software patents are not yet routinely available.

[90] Press Release, Trend Micro, Inc., IBM and Trend Micro Sign Patent Cross-Licensing Agreement (Dec. 10, 1997), http://www.trendmicro.com/en/about/news/pr/archive/1997/pr121097.htm; Press Release, Trend Micro, Inc., Symantec and Trend Micro Cooperate to Benefit Customers (Apr. 8,

portfolios, either by acquiring firms or by developing their own technologies. Of particular interest are recent patents that protect novel ways of integrating multiple security technologies.[91]

¶ 66 For the first movers in anti-spam technologies, a number of fairly broad patents were an important factor in protecting their business at the start-up stage. Among these were patents issued to Brightmail,[92] Tumbleweed,[93] and Postini[94]. These patents may have discouraged Internet-security incumbents from reverse engineering their anti-spam technologies, and Table 2 reflects the extent to which incumbents have acquired technologies by firm acquisition and licensing. In this regard, it would seem patents may serve to give innovators breathing space to develop into significant, wide-spectrum suppliers. However, broad patents have very much been the exception. The great majority of patents are narrow, and do not block competitors from major competitive areas. A significant minority of the players in Table 1 have no patents at all, and yet this does not appear to have inhibited their product development or licensing activity. Where firms do have patents, however, such patents may enable vendors to gain some small advantage in the market place, though this tends to be short-lived as a result of the rapid evolution of spam and anti-spam technologies. Probably a more significant factor is that the ownership of a patent signals to interested parties—such as potential acquirers, licensees, venture capitalists, and end users—that the patent-possessing firm has technology that is novel and cannot be freely appropriated.[95] For example, for venture capital firms, patents are one factor—taken along with several others, such as rational business plans and management depth—which affect investment decisions.[96] The majority of U.S. firms in Table 1 have received venture capital investments of tens of millions of dollars.[97]

B. Appliances

¶ 67 The market for anti-spam appliances has exploded since the first product was released by CipherTrust in 2001. An appliance offers three primary benefits to the corporate buyer (on which firms compete). First, an appliance requires minimal configuration and customization: appliances can be plugged into the network with little or no configuration and are essentially plug-play-and-forget; they are automatically updated

---

1998), http://www.trendmicro.com/en/about/news/pr/archive/1998/pr040898.htm; Press Release, Trend Micro, Inc., Network Associates and Trend Micro Settle Anti-virus Software Patent Dispute (Jun. 1, 2000), http://www.trendmicro.com/en/about/news/pr/archive/2000/pr060100.htm.

91 *See, e.g.*, U.S. Patent No. 6,836,792 (issued Dec. 28, 2004).

92 U.S. Patent No. 6,052,709 (issued Apr. 18, 2000); U.S. Patent No. 5,999,932 (issued Dec. 7, 1999).

93 U.S. Patent No. 6,609,196 (issued Aug. 19, 2003) (Tumbleweed e-mail firewall).

94 U.S. Patent No. 6,650,890 (issued Nov. 18, 2003).

95 Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625 (2002).

96 *Id.* at 653.

97 *See, e.g.*, Alistair Goldfisher, *VCs Host Big Spam Sale—Get It While It's Hot!*, VENTURE CAP. J., May 1, 2004, at 3-5 (detailing investments in ActiveState, BrightMail, Corvigo, and MessageFire); Dan Primack, *VCs Get Piggy with Spam Smorgasbord*, VENTURE CAP. J. Oct. 1, 2003, at 15-17 (detailing investments in Corvigo, Postini, FrontBridge, MessageGate, and MailFrontier).

with new software releases and patches; and adjust automatically to the characteristics of incoming e-mail. Second, appliances offer very high throughputs, up to several hundred thousand e-mails per hour, which software-product solutions on conventional servers cannot match. Third, appliances offer extremely high levels of availability, with close to zero outages.

¶ 68 The leading firm in the appliance industry, by far, is CipherTrust, which supplied nearly half the market in 2003.[98] The firm was founded in March 2000 and its IronMail appliance was released in August 2001.[99] CipherTrust has developed its own anti-spam technologies for which patents are pending.[100]

¶ 69 There is no such thing as a "pure" anti-spam appliance; appliances additionally incorporate protection against e-mail-borne viruses and other Internet threats such as worms, Trojan horses, and spambots. All of the anti-spam appliance vendors in Table 3, including CipherTrust, license anti-virus technology from vendors such as Sophos, Network Associates, Kapersky, and F-Secure. Some appliances also incorporate encryption technology from vendors such as PostX, PGP, and RSA. In this respect, anti-spam appliance vendors, by integrating software from several suppliers into a computer system, are operating as classic turnkey suppliers.[101]

¶ 70 There is no evidence of any reluctance on the part of vendors to license their anti-spam or anti-virus technologies, even to close competitors. For example, several vendors in Table 3 have licenses with Symantec-Brightmail, Tumbleweed, and Network Associates, even though all of these vendors are in the process of entering the appliance market.[102] There are many suppliers of anti-spam technology, and because substitutes are so generally available, an individual vendors' best economic option is to license its technology. This allows a vendor to garner some favorable publicity through cross-marketing (e.g., "spam detection with Symantec BrightMail"[103]) and royalty income, even at the risk of cannibalizing its own sales. The licensing of anti-spam patents prevents technology appropriation by reverse engineering rather than using the patents to block competitors.

---

98 *See* INTERNATIONAL DATA CORP, *supra* note 61.

99 *CipherTrust Launches the Industry's First E-mail Specific, Security Appliance – IronMail*, SECURE COMPUTING, Aug. 27, 2001, http://www.ciphertrust.com/company/press_and_events/article.php?id=0000139.

100 U.S. Patent Application No. 20030172302 (filed Mar. 8, 2002); U.S. Patent Application No. 20030172301 (filed Mar. 8, 2002) ; U.S. Patent Application No. 20030172294 (filed Feb. 24, 2003); U.S. Patent Application No. 20030172292 (filed Feb. 7, 2003); U.S. Patent Application No. 20030172291 (filed Feb. 7, 2003); U.S. Patent Application No. 20030172167 (filed Mar. 6, 2003) ; U.S. Patent Application No. 20030172166 (filed Mar. 8, 2002).

101 *See supra* note 55.

102 Symantec acquired the anti-spam appliance vendor TurnTide in July 2004; Tumbleweed acquired the appliance vendor Corvigo in March 2004. *See* Symantec Acquisitions, http://www.symantec.com/about/profile/development/acquisitions/index.jsp (last visited Sept. 23, 2004); Press Release, Tumbleweed, Tumbleweed Acquires Anti-spam Appliance Vendor Corvigo (Mar. 18, 2004), *available at* http://www.tumbleweed.com/news/press_releases/2004/2004-03-18.html.

103 *See, e.g.*, IronPort C-Series Overview, http://www.ironport.com/pdf/ironport_cseries_brochure.pdf (last visited Jan. 3, 2006).

¶ 71 The majority of the appliance vendors in Table 3 compete primarily on the basis of reliability and throughput, and make use of licensed, and somewhat commoditized, anti-spam technology. Reliability and throughput have been achieved by the development of proprietary operating systems specialized to the task of message transmission and filtering, usually based on Unix or Linux. For example, IronPort Systems has developed its operating system AsyncOS, which was "[f]ounded on a rock-solid UNIX-based kernel stripped of all non-essential components."[104] Borderware—an Internet-security appliance vendor of ten years standing—has a proprietary hardened operating system, S-Core, designed for immunity from all classes of internet threats; the firm claims that in "over 10 years of field testing S-Core has never been compromised."[105] Borderware licenses its patented technology to competitor 3Com.[106] MiraPoint extols the speed of its trademarked "Messaging Operating System," claiming an industry speed record of 700,000 e-mails per hour on its high-end Message Director MD450 appliance.[107] Nokia—which entered the Internet-security appliance market in 2000—has a hardened operating system that incorporates patented, "IP clustering technology," which enables the integration of multiple appliances for greater resilience and throughput.[108] Borderware, MiraPoint, and Nokia all have patented features of their appliance operating systems, without any evidence of competitive blocking. Indeed, there are so many ways that it is possible to harden or speed-up an operating system that it is highly unlikely such blocking could occur.

¶ 72 A notable feature of Table 3 is that a majority of firms license technology from third party suppliers rather than develop their own anti-spam technology. This demonstrates the existence of a functioning and thriving market for anti-spam technologies, in which patents help innovators recoup their investments by reducing the risk of imitation through reverse engineering. Patent disclosure may also be facilitating the rapid evolution of anti-spam technologies because competitors can examine and improve, or invent around, an invention. However, proprietary technologies, patented or not, have not excluded open source vendors—Barracuda Networks, for example, has managed to successfully coexist with proprietary competitors using entirely open source anti-spam solutions.[109]

---

[104] IronPort Systems, Inc. Home Page, http://www.ironport.com/company/ (last visited Aug. 23, 2005).

[105] *See* Borderware Technologies Inc. Brochure, *at* http://www.borderware.com/pdfs/corporate.pdf.

[106] Press Release, BorderWare, 3Com & BorderWare Announce Strategic OEM Relationship (Feb. 7, 2005), *available at* http://www.borderware.com/press/releases.php?action=v&id=143.

[107] *Mirapoint Breaks Industry Speed Record For Messaging; Outperforms Competition with Fastest Email Security & Server Appliances Available*, BUS. WIRE, Feb. 17, 2004, http://www.findarticles.com/p/articles/mi_m0EIN/is_2004_Feb_17/ai_113339895.

[108] Press Release, Nokia, Nokia Announces the Availability of IP Clustering Technology for its IP Security Platforms (Jul. 3, 2002), *at* http://press.nokia.com/PR/200207/865550_5.html.

[109] Barracuda's web site states "[s]ome Barracuda Networks products utilize 'open source' programs in their operation. If we make any changes to open source programs we offer those changes, bug fixes, or improvements back to the team working on the project. Typically, our only changes to open source programs are bug fixes, which are incorporated into the open source project at the discretion of the open source project team." Barracuda Networks Frequently Asked Questions, *at* http://www.barracudanetworks.com/ns/company/company_faq.php (last visited Aug. 23, 2005).

C. Hosted Services

¶ 73 In a hosted service, incoming e-mail is redirected to an anti-spam processing service, which then sends filtered e-mail on to the subscriber, while quarantining suspected spam on the provider's site. The redirection process is achieved by changing the mail exchange record ("MX record") that controls the routing of e-mail to a mail server.[110]

¶ 74 There are several subscriber benefits of a hosted service. First, the set-up is very fast because there is no hardware or software to install. For example, MailWise claims that "[o]ur setup takes 10 minutes" and "all it takes is a single phone call to us."[111] Second, because subscribers are usually charged according to the volume of e-mail processed, a hosted service can be more economical than an appliance for organizations with small or fluctuating volumes of e-mail. Lastly, because quarantined e-mail does not actually reach the organization, the load on mail servers can be significantly reduced.

¶ 75 Table 4 lists the leading hosted-service providers in order of 2003 anti-spam revenues. The hosted-service sector is the most concentrated in the anti-spam industry, with the top three firms accounting for 80% of sales. The reason for this concentration is the high barrier to entry created by the need for major front-end investments in Internet infrastructure. The top firms are among the most heavily capitalized in the industry. For example, since its founding in 1999, Postini has received four rounds of venture funding totaling $26 million,[112] while MessageLabs has had three funding rounds totaling $58 million.[113] The more recently established FrontBridge and MX Logic have received $28 million and $26 million respectively.[114]

¶ 76 For hosted services, the reliability and reach of their infrastructures is paramount. For example, MessageLabs advertises an infrastructure that "spans 13 data centers on four continents in six countries," with Network Operation Centers in New York, London, Hong Kong and Sydney.[115] Postini has recently consolidated its ten data centers and re-

---

[110] The MX record maintains the translation between an e-mail domain name and the physical IP address of a mail server. The e-mail domain name is that part of the e-mail address that follows the @-sign (for example, the e-mail domain name for jsmith@travel.com is travel.com). The IP address is a four-part number, such as 212.84.110.143, that uniquely defines the location of the server. The relationship between the domain name and the IP address is somewhat like the relationship between the name of an organization and its telephone number, or the name of an individual and a bank account number. The MX record is replicated in the thousands of "domain name servers" that route information around the Internet.

[111] MailWise, LLC, Why MailWise Filter is Better, *at* http://www.mailwise.com/sv_why_we_are_better.htm (last visited Aug. 24, 2005).

[112] Primack, *supra* note 97, at 15.

[113] MessageLabs Group, Financial Profile, http://www.messagelabsgroup.com/Financials/financialprofile/index.htm (last visited Aug. 23, 2005).

[114] *FrontBridge Fuels Business Expansion with $10 Million in New Funding*, FRESHNEWS.COM, August 2, 2004, http://www.freshnews.com/cgi-bin/jsj_news/print.cgi?article_ID=18796; MX Logic, Inc., Investors, http://www.mxlogic.com/about/investors.html.

[115] MessageLabs Groups Quick Facts, *at* http://www.messagelabs.com/About_Us/Company_Profile/Quick_Facts (last visited Aug. 23, 2005).

architected its network by partnering with Equinix, a leading Internet access provider.[116] FrontBridge emphasizes the reliability of its system, which it claims has been running non-stop for 5 years.[117]

¶ 77 Hosted services compete far less on the efficacy of their anti-spam technology than on their infrastructures and quality of service. Indeed, two of the four leading providers in Table 4 rely largely on technology licensed from Symantec Brightmail. The most important patent in the hosted services sector has not been for spam elimination *per se*, but for the MX-record changing technique; Postini was the first mover in hosted services, and applied for a broad patent in September 2000.[118] This patent has been contentious because one of its primary claims is the MX record-changing technique which is the basis of all hosted services. The hosted-service market was barely established at this time, and the patent therefore had the potential to exclude all competitors from the sector. When the patent came to public attention on its issuance in November 2003, there was hostile press and Internet comment.[119] At first, Scott Petry, a co-inventor of the patent and co-founder of Postini, engaged in a little triumphalism, and it was reported that "Postini executives are studying the patent and considering ways to 'maximize' its value to the company."[120] Given that Postini already dominated the market and no doubt recognized that too much saber-rattling could be damaging to its image, Petry quickly adopted a more conciliatory tone. Petry was subsequently reported as saying "the patent was applied for so that he and the other patent applicants . . . could protect their work, show industry leadership and develop intellectual property to show public and private investors."[121] It is worth noting that, despite a patent with considerable blocking potential, the social disapproval of the Internet community had a powerful moderating effect.

VI. THE COEXISTENCE OF OPEN-SOURCE AND PROPRIETARY SOLUTIONS

¶ 78 There are several prominent open-source anti-spam products, and many more less prominent ones. Here we discuss the single most important open-source content filter, SpamAssassin, and the two most important product categories with competing open-source solutions, Distributed Checksum Filtering and DNS-based blacklists.

¶ 79 The non-proprietary open-source and the proprietary closed-source models of

---

116 Press Release, Equinix, Inc., Anti-spam Provider Postini Selects Equinix to Enhance Redundancy of its Operations (Apr. 5, 2004), *available at* http://www.equinix.com/press/press/2004/04_05_04.htm.

117 Frontbridge Technologies, Inc., Network Statistics (pdf on file with the author).

118 *See* U.S. Patent No. 6,650,890 (issued Nov. 18, 2003).

119 *See, e.g.*, Paul Roberts, *Postini Anti-spam Patent Could Cause Headaches*, COMPUTERWORLD, Mar. 26, 2004, *available at* http://www.computerworld.com/printthis/2004/0,4814,91685,00.html (discussing objections to the patent).

120 *Id.*

121 Cameron Sturdevant, *Anti-spam Patents: Precursor to Consolidation?*, EWEEK, Apr. 19, 2004, http://www.eweek.com/print_article2/0,2533,a=124683,00.asp.

software development are often portrayed as irreconcilable and mutually exclusive.[122] Our empirical study of the anti-spam technology prospect shows that this is far from the case. For example, some sixty proprietary anti-spam vendors make use of open-source SpamAssassin.[123] Rather than the two worlds being mutually exclusive, there is a spectrum of usage. At one end of the spectrum are firms such as Symantec, which use exclusively proprietary technology; at the other end there are firms such as Barracuda Networks, which use exclusively open source. However, the majority of firms use open- and closed-source in a wholly pragmatic way. For example, Sendmail, the supplier of the productized version of the most popular open-source mail server software, partners with open-source Cloudmark for anti-spam technology and Network Associates for proprietary anti-virus solutions.

¶ 80 It is important to understand that open-source companies are profit-seeking entities and not charitable institutions. Like their proprietary counterparts, they sell their products and services for money, have paid staffs, own IP (including patents), consume venture capital, and have products that are delivered to customers under mutually enforceable contracts. The most important difference between open- and closed-source solutions is in the software development practice.

¶ 81 A typical open-source project, such as SpamAssassin, consists of a body of code, freely available on the Internet, developed by a community of highly competent individuals, who might range from seasoned employees of profit-making firms to student volunteers at universities. Many organizations—often cash poor, but labor rich, such as universities—will use the raw, open-source code to implement their local solutions. A few will report and sometimes fix bugs. Other organizations, however, prefer to buy a productized solution from an open-source vendor who customizes the code (so that it runs out of the box) and provides after-sales support.

A. SpamAssassin

¶ 82 SpamAssassin is the best known and most widely deployed open-source anti-spam solution. It has evolved by integrating a number of anti-spam techniques, including textual analysis, Bayesian filtering, distributed checksums, blacklisting, and whitelisting. The SpamAssassin project was begun in 2001 by Justin Mason (later a co-founder of Deersoft), with the help of a small number of lead programmers, including Matt Sergeant (now senior anti-spam technologist with MessageLabs) and a hundred contributing developers.[124]

¶ 83 SpamAssassin is popular with hands-on systems administrators in Unix/Linux environments and is integrated into proprietary packages. SpamAssassin features

---

[122] *See, e.g., Open source vs. closed source, available at* http://en.wikipedia.org/wiki/Open_source_vs._closed_source (last visited Sept. 23, 2007).

[123] The Apache SpamAssassin Project Home Page, http://wiki.apache.org/spamassassin/FrontPage (last edited Oct. 18, 2004).

[124] Mike Cassidy, *Spam Fight Could Use Some Ruthless Soldiers*, SAN JOSE MERCURY NEWS, Jan. 14, 2003, at 1C, (on file with author).

prominently in the practitioner literature, including two dedicated monographs.[125] To deploy SpamAssassin, a systems administrator will download the source code and compile and install it on a mail server. Often incoming mail will also be filtered by other open-source programs such as the DSPAM or CRM114 Bayesian filters.[126] There is a considerable administrative burden for customizing the packages to local requirements and downloading the latest versions and bug fixes. However, such fine tuning of the anti-spam solution allows for superior filtering.

¶ 84 The direct deployment of open-source solutions is time consuming, technically challenging, and mainly limited to Unix/Linux systems.[127] Consequently, there are market opportunities for the development of packaged solutions for end users and OEM development kits for solution vendors. The originator of an open-source project is particularly well placed to exploit these commercial opportunities.

¶ 85 In the case of SpamAssassin, its originator Justin Mason, together with another developer and a venture-fund partner, established Deersoft in June 2002.[128] The company "took the open-source product . . . and turned it into a more slick, packaged tool that could appeal to a wider audience—and be sold at a profit. The results are SpamAssassin Pro and SpamAssassin Enterprise."[129] The company, however, remained committed to supplying the open-source community by maintaining the source code and keeping it online.

¶ 86 In January 2003, Deersoft was acquired by Network Associates for an undisclosed amount so that SpamAssassin could be integrated into McAfee SpamKiller.[130] However, SpamAssassin remains a freely available open-source development, and Mason is simultaneously an employee of Network Associates and a member of the SpamAssassin management committee. In effect, he is a conduit between the open source world and Network Associates, an important strategic asset in the ever-evolving software landscape.

¶ 87 Another re-packager of SpamAssassin was ActiveState, a Canadian company founded in 1997 by members of the open-source community in an effort to develop tools for the new wave of Internet programming systems, such as Perl, Python, and PHP.[131] Its first anti-spam product PerlMX was released in October 2000. SpamAssassin was

---

[125] ALISTAIR MCDONALD, SPAMASSASSIN: A PRACTICAL GUIDE TO INTEGRATION AND CONFIGURATION (Packt Publ'g 2004); ALAN SCHWARTZ, SPAMASSASSIN (O'Reilly 2004).

[126] For descriptions of DSPAM and CRM114 filters, see NuclearElephant.com, The DSPAM Project, http://www.nuclearelephant.com/projects/dspam (last visited Aug. 23, 2005) and CRM114 - The Controllable Regex Mutilator, http://crm114.sourceforge.net/ (last visited Aug. 23, 2005).

[127] Two practitioners suggest "between 4 and 20 hours of administrator time each week." Lorraine Faith Cranor & Brain A. LaMacchia, *Spam!*, 41 COMM. ASS'N COMPUTING MACHINERY 74, 76 (1998).

[128] Christopher Lindquist, *Dawn of a Spam Killer*, CIO MAGAZINE, Feb. 1, 2003, *available at* http://www.cio.com/archive/020103/et_company.html.

[129] *Id.*

[130] *McAfee Unleashes SpamKiller for Small Businesses*, CLICKZ NETWORK, Apr. 21, 2003, *available at* http://www.clickz.com/showPage.html?page=2194171.

[131] Rob Reilly, *PureMessage Raises E-mail Admin Standard ActiveState's Perl of a Product*, LINUXPLANET, Sept. 2, 2003, *available at* http://linuxplanet.com/linuxplanet/reports/4983/1/

incorporated into the product in 2002 and it was subsequently renamed PureMessage. In April 2003, ActiveState was acquired by the U.K. Internet-security vendor Sophos for $23 million—ActiveState was by that time a substantial company with over a hundred employees. PureMessage is now a Sophos Product. Interestingly, ActiveState operates as a subsidiary company, and Sophos has neither integrated it into its regular operations nor has it closed down ActiveState's non-security-related activities. Rather Sophos appears to be using ActiveState as a listening post and a strategic investment. Although Sophos was originally motivated by acquiring ActiveState's PureMessage software, according to an insider, Sophos is now trying "to figure out . . . how to marry the two cultures and help us do things better than we did before, and vice versa."[132]

B. Distributed Checksum Filters and DNS-based Blacklisting

¶ 88 The open-source community has been particularly effective in establishing collaborative schemes for identifying and reporting spam. In Distributed Checksum Filter ("DCF") schemes, individual users report spam, typically by clicking a button integrated into their e-mail reader software. Once spam e-mail has been reported by a sufficient number of users, it can be eliminated from the inboxes of all subscribers of the service.

¶ 89 The two best known DCF schemes are Vipul's Razor and the Distributed Checksum Clearinghouse. They flourished in the open-source community because they needed a network of servers to capture spam reports and enough volunteers came forward to offer their corporate servers without cost. It would have been much more difficult for a proprietary vendor to establish the relationships needed to free-ride on corporate servers.

¶ 90 The Vipul's Razor ("VR") project was begun in 1998, and is named for its inventor Vipul Prakash.[133] In the VR scheme, when a spam e-mail is identified by a user the software computes a checksum—rather like a fingerprint or a signature—that uniquely identifies the e-mail.

¶ 91 In 2001, Prakash co-founded Cloudmark to productize the technology for end users and for OEM licensing to proprietary vendors. Cloudmark has received $15 million in venture funding, and has been very successful—in 2003, it ranked number 11 among global anti-spam software product vendors by annual sales (Table 1). It sells a desktop product, SafetyBar, and an OEM solution. Licensees include proprietary firms AhnLab, Bizanga, Mayflower, and Secure Computing, and open-source vendors Sendmail and OpenWave.[134] Prakash has been personally successful too, being nominated a top-100 innovator by *Technology Review*.[135]

¶ 92 In May 2005, Prakash applied for two patents covering VR technology, both of

---

[132] Vance McCarthy, *ActiveState To Stay Open Source After Buyout*, OPEN ENTER. TRENDS, Oct, 1, 2003, *available at* http://www.oetrends.com/archive.php?action=view_record&idnum=265.

[133] Cloudmark, Cloudmark Overview, http://www.cloudmark.com/company/ (last visited Apr. 27, 2005).

[134] Cloudmark, OEM Partners, http://www.cloudmark.com/partners/oem/ (last visited Apr. 27, 2005).

[135] John Verity, *Computing*, TECH. REV., October 2003, at 58-68.

which are still pending.[136] Neither Cloudmark nor Prakash has commented publicly on these patent applications and whether they conflict with the usual open-source antipathy to software patents. However, regardless of the IP value of software patents, such patents would be useful in terms of signaling innovative quality, in attracting further venture funding, and in indemnifying users against potential infringements.[137] VR remains an open-source project.[138]

¶ 93 The Distributed Checksum Clearinghouse (DCC) scheme is somewhat similar to VR. It was developed as an open source project by Vernon Schryver in 2000, and is supported by his small company Rhyolite Software.[139] Unlike Prakash's Cloudmark, Rhyolite is not a significant company and operates in the classic manner of a one- or two-person custom programming and consulting operation, rather than offering a productized version of DCC. In March 2005, Schryver sold the rights for DCC to the Israeli anti-spam vendor Commtouch (Table 1), which now has exclusive worldwide licensing rights.[140] Schryver stated in an open-source forum that his decision to sell DCC for a "pittance" was because he did not have the financial resources for "things that cost money like a feed of the (formerly free) SBL from Spamhaus."[141] Schryver was evidently also piqued by the use of DCC by free-loading commercial anti-spam vendors, stating "I think it's perfectly fine to make a buck with other people's free source, but there is a difference between parasitism and commensalism, not to mention symbiosis."[142]

¶ 94 Prior to the DCC acquisition, Commtouch had developed a proprietary DCC-type technology, and had purchased U.S. Patent No. 6,330,590 from a private inventor in September 2005, which "stood out as the earliest and most important patent in the field."[143] One view might be that by acquiring the DCC rights, and giving it some IP protection, Commtouch saved DCC technology from extinction. A contrary view—that such an action imperiled the open-source spirit—was expressed in the DCC forum: "I lived through 9-11 so far worse possibilities exist, but not in relation to software."[144]

---

[136] US Patent Application No. 20050097435 (filed June 24, 2004); U.S. Patent Application No. 20050114452 (filed Nov. 3, 2003).

[137] There are, however, two prior DCF-type patents: U.S. Patent No. 6,330,590 (issued Dec. 11, 2001) and U.S. Patent No. 6,453,327 (issued Sept. 17, 2002).

[138] Vipul's Razor, http://razor.sourceforge.net/ (last visited Aug. 23, 2005).

[139] Rhyolite Software LLC Home Page, http://www.rhyolite.com/ (last visited Sept. 23, 2007).

[140] Commtouch DCC Licensing Program, http://www.commtouch.com/Site/Company/DCC.asp (last visited Sept. 23, 2007).

[141] Posting of Vernon Schryver to http://www.rhyolite.com/pipermail/dcc/2005/002570.html (Mar. 16, 2005, 16:36:52 MST). Spamhaus is another strapped-for-cash open-source project. *See infra* Section VI(C).

[142] Posting of Vernon Schryver to http://www.rhyolite.com/pipermail/dcc/2005/002579.html (Mar. 17, 2005, 10:24:49 MST).

[143] Press Release, Commtouch Software, Ltd., Commtouch Acquires Patent for Preventing Delivery of Unwanted Bulk Email, (Sept, 1, 2004), *at* http://www.commtouch.com//Site/News_Events/pr_content.asp?news_id=38&cat_id=1.

[144] Posting of Ruben Safir to http://www.rhyolite.com/pipermail/dcc/2005/002571.html (Mar. 17, 2005 01:17:16 EST).

¶ 95 In fact, Schryver had tried to negotiate a deal that would enable DCC to remain in the public domain for non-commercial users. Commtouch stated that the "[u]se of the DCC code and participation in the global network of DCC clients and servers is free to all public network users and organizations that are not reselling anti-spam services."[145] This allowable use included certain elements of U.S. Patent No. 6,330,590. However, to judge from the open-source debates, the position of non-public, non-anti-spam-vending organizations remains unclear.[146]

¶ 96 DNS-based blacklisting causes all e-mail from the domain of a presumed spammer to be rejected by a subscriber to the service. This is a very Draconian measure that comes close to censorship. It owes its initial acceptance to the reputation of its inventor Paul Vixie, an Internet pioneer.

¶ 97 Vixie, and his colleague David Rand, began development of their Mail Abuse Prevention System ("MAPS") in 1996.[147] Although technically an open-source project, MAPS might be better described as an open *service*. The code was relatively simple, but MAPS also required a significant infrastructure and a team of administrators to validate spam reports. By fall 2000, MAPS had 16 paid staff and 24 volunteers, funded partly by Vixie and Rand personally and partly by consulting operations.[148]

¶ 98 MAPS' activities were controversial from the beginning. Early on, it blocked some major ISPs who, unbeknownst to them, were harboring spammers.[149] As a result, ordinary subscribers of these ISPs were greatly inconvenienced when their outgoing mail was blocked to MAPS subscribers. To many, MAPS' high-handed policing of the Internet smacked of censorship and vigilantism.[150] When invited to respond to criticism by *Slashdot*, the company declined to comment.[151]

---

[145] Commtouch Software Ltd,, DCC Licensing Program, *at* http://www.commtouch.com/Site/Company/DCC.asp.

[146] SpamAssassin's Justin Mason has stated the he would no longer be able to assume default use of DCC (and Vipul's Razor) in open-source versions of SpamAssassin. *See* Justin Mason, Happy Software Prole, http://taint.org/2005/03/19/013823a.html (Mar. 19, 2005, 01:38 PST).

[147] MAPS, About MAPS, http://www.mail-abuse.com/company/ (last visited Sept. 23, 2007); MAPS, Management Team, http://www.mail-abuse.com/company/mgtteam.html (last visited Sept. 23, 2007).

[148] Patricia Odell, *Cease-fire*, DIRECT, Sep. 1, 2000, *available at* http://www.directmag.com/mag/marketing_ceasefire/.

[149] For example, in 1998, nearly 7000 subscribers to the ISP Internet Communications experienced e-mail blockage for a week. Paul Eng, *An Innocent Company Gets Snared in an Anti-spam Sweep*, BUSINESSWEEK ONLINE, Dec. 17, 1998, http://www.businessweek.com/smallbiz/news/date/9812/e981217.htm.

[150] Posting of Jamie to http://slashdot.org/article.pl?sid=00/12/13/1853237 (Dec. 13, 2000, 11:30 p.m.); Kiri Blakeley, *Spam Warfare*, FORBES.COM, Sept. 18, 2000, http://www.forbes.com/forbes/2000/0918/6608230a.html.

[151] *See* Posting of Jamie, *supra* note 150 (stating that "I contacted Paul Vixie to ask about AboveNet and how it uses the RBL, but he refused comment, sending me to AboveNet PR, who didn't get back to me by deadline time.").

¶ 99 In July 2001, MAPS announced that its services would no longer be free.[152] At the time, it was estimated that up to 40% of the world's Internet hosts subscribed to at least one MAPS list and the decision to charge for access "sent shockwaves throughout the community."[153] In July 2004, MAPS was fully privatized as Kelkea. Clients of Kelkea include anti-spam vendors Symantec, MessageLabs, and MXlogic, as well as global ISPs such as AOL, USA.net and BT.[154]

¶ 100 A similar trajectory to MAPS-Kelkea was followed by another major blacklisting service, SpamCop, created by Julian Haight in 1998. SpamCop was run as a "one man show", with the support of volunteers.[155] By 2003, however, the system had become unsustainable because of denial-of-service attacks, presumed to come from spammers.[156] In November 2003, Haight sold SpamCop to the anti-spam appliance vendor IronPort for an undisclosed sum, with the intention that IronPort would invest $1 million to build an infrastructure capable of withstanding denial-of-service attacks.[157] Haight is now supported by IronPort to maintain the SpamCop.net website, access to which, while no longer free, is modestly priced.[158] The open source code to access SpamCop has, however, become orphaned.[159]

¶ 101 Although the MAPS and SpamCop DNS-based blacklists have been lost to the public domain, there remain several hundred other freely available lists.[160] Among these are major services such as Spamhaus, SORBS and DSBL.org, although they are all vulnerable to denial-of-service attacks.

¶ 102 The privatization of MAPS and SpamCop (and Cloudmark and DCC) suggests that there are limits to the open-source model for anti-spam solutions. Open source seems to work well when developers are remunerated by regular employers or grants, and they are volunteering the skills that they love to exercise. However, the model does not extend to complementary activities such as database administration, tedious work for which few would volunteer, or investment in expensive infrastructure.

---

[152] Tom Geller, *The Future of MAPS*, SPAMCON FOUND. NEWSLETTER (SpamCon Foundation, S.F., Ca.) July 16, 2001, *available at* http://spamcon.org/about/news/newsletters/005/opinion.shtml.

[153] *Id.*

[154] Kelkea, *Customers*, www.kelkea.com/company/customers.html (last visited June 5, 2005).

[155] Arik Hesseldahl, *The Cop on the Spam Beat*, FORBES, Nov. 24, 2003, *available at* http://www.forbes.com/2003/11/24/cx_ah_1124tentech_print.html.

[156] Other services reportedly experiencing denial-of-service attacks included Spamhaus, SORBS, OpenRBL, Monkey.com, and Osirusoft. *See* Hiawatha Bray, *Saboteurs Hit Spam's Blockers*, BOSTON GLOBE, August 28, 2003, at A1, *available at* http://www.boston.com/news/nation/articles/2003/08/28/saboteurs_hit_spams_blockers/.

[157] Tony Kontzer, *IronPort Acquires Anti-spam Blacklist*, INFORMATION WEEK, Nov. 25, 2003, http://www.informationweek.com/showArticle.jhtml;jsessionid=2DEAACEPRPJPEQSNDBECKICCJUMEKJVN?articleID=16400810. .

[158] *Id.*; *see also* spamcop.net email, http://www.spamcop.net/ces/pricing.shtml (last visited Sept. 23, 2007).

[159] Posting of Julian Haight to http://www.spamcop.net/source.shtml (last visited Aug. 23, 2005).

[160] Declude, Inc., List of All Known DNS-based Spam Databases, http://www.declude.com/Articles.asp?ID=97 (last visited Aug. 23, 2005).

VII. CONCLUSION

¶ 103 The patent system has to work for the benefit of society as well as firms, and for small firms as well as large ones. In this article we have accepted software patents as a fact of life and have tried to evaluate whether they are in practice helping or hindering the industry. In order to frame the discussion, we found it useful to represent anti-spam as a technology prospect. We conclude by directly addressing the following issues as they relate to the mining of the anti-spam prospect:

1. Is the patent system achieving an orderly development of the anti-spam prospect?
2. Are big firms with strong patents excluding small firms?
3. Are broad patents blocking competitors?
4. Are there too many narrow patents, resulting in over-fishing of the prospect?
5. Is the patent system facilitating a market for anti-spam technology with reasonable transaction costs?
6. Is open-source activity being threatened by the patent system?

¶ 104 By an "orderly development" of the prospect, we mean the absence of a "gold rush" to particular hotspots, resulting in localized over-fishing. Rather we wish to see exploitation of the full range of possibilities in the prospect, including the less obvious areas. In section III and Appendix A, we examined in detail one hundred-plus anti-spam patents. These showed a wide range of distinct techniques, and there is probably no anti-spam technique currently in use that is not represented. Likewise in sections VI and V, we examined the forty-two most significant firms in the anti-spam industry, which collectively deploy the full range of approaches. Within the three broad categories of software products, appliances, and hosted services, although a small number of firms relied on a single technique (such as Cloudmark's use of Vipul's Razor), the majority of firms used a variety of techniques from the prospect.

¶ 105 Large firms with strong patent positions can intimidate small firms, discouraging them from entry. This was shown empirically to be the case in biotechnology in an influential article by Josh Lerner where he observed that small firms avoided competing in areas that were heavily patented by major firms.[161] In the anti-spam prospect, however, although there are several major firms with strong patent positions, there is no evidence that small firms are intimidated by these patents (or perhaps even aware of them). For example, AT&T, Digital Equipment Corp. (now HP), IBM, and Microsoft all have important anti-spam patents. All these firms include anti-spam techniques in their products and services, but none of them is a direct player in the anti-spam industry. Patent licensing is somewhat underdeveloped in the software industry, and we cannot determine whether infringement of any of these patents is taking place, although we think

[161] Josh Lerner, *Patenting in the Shadow of Competitors*, 38 J.L. & ECON. 463, 463 (1995). Lerner argues that "firms with high litigation costs are less likely to patent in subclasses with many other awards, particularly those of firms with low litigation costs." *Id*. at 463.

it is quite possible. As the industry adapts to the patent system in the years to come, we expect that patent owners will become more assertive. However, major firms in the IT industries—and certainly those noted above—have a good reputation for fair and non-discriminatory licensing.

¶ 106 One might have expected that the incumbent, mid-sized Internet-security vendors (i.e., the firms in Table 2) would cast a threatening patent shadow over the start-up anti-spam firms (i.e., most of the firms in Table 1). However, it is actually the start-up firms that are licensing technology to the incumbents; this suggests that the patent system may well be serving to protect start-up firms from having their technology appropriated by cloning or reverse engineering. On the surface, relations between large, medium, and small firms in the anti-spam industry seem positively cordial. It is not our experience that software entrepreneurs are more passive or diplomatic than participants in other industries such as biotechnology, so we look to another explanation. We think that it is in the nature of the anti-spam problem (and of software in general), that solutions typically require the integration of multiple techniques. Patents owners, we suspect, tread carefully: the owner of a strong patent may one-day need to take out a license on another firm's patent. We think this is a noteworthy phenomenon. Many critics of software patents have argued that the need for large numbers of patents will make software writing infeasible.[162] We think this fear may be ill-founded. Because every vendor is likely to need to license some other firms' patent at some time, there is an incentive for fair and non-discriminatory dealing with competitors in this repeated game. The evolving anti-spam prospect may well eventually result in a patent pool. The result of such a pool is that R&D costs would be more fairly shared among firms, and vendors of free-loading clone products would be obliged to invest in R&D or pay an appropriate license fee to join the pool.

¶ 107 Although we identified a number of broad anti-spam patents detailed in Appendix A,[163] only two patents have caused public concern because of their blocking potential. In both cases, the firms that owned the patents engaged in aggressive litigation or posturing. In the case of the Mailblocks patent, discussed in section IV(C), two infringement suits were resolved. The suits were not unreasonable, but probably ill-advised. In the event the Mailblocks technology proved ephemeral and—had the firm not been acquired by AOL—Mailblocks might well have found itself in the invidious position of having to license technology from competitors. We have noted elsewhere that because software evolves so rapidly, the effective life of an application-software patent is on average only five years.[164] It is therefore unlikely, however history had unfolded, that the Mailblocks patent could have blocked its competitors for very long. The Postini patent, discussed in

---

[162] *See, e.g.*, Richard Stallman & Simson Garfinkle, *Viewpoint: Against Software Patents*, 35 COMM. ASS'N COMPUTING MACHINERY 17, 17-22, 121 (1992), who argue "[s]oftware patents threaten to devastate America's computer industry." *Id.* at 17.

[163] U.S. Patent No. 5,930,479 (issued July 27,1999); U.S. Patent No. 6,321,267 (issued Nov. 20, 2001); U.S. Patent No. 6,161,130 (issued Dec. 12, 2000); U.S. Patent No. 6,633,630 (issued Oct. 14, 2003), U.S. Patent No. 6,654,787 (issued Nov. 25, 2003); U.S. Patent No. 6,757,830 (issued June 29, 2004).

[164] Martin Campbell-Kelly and Patrick Valduriez, *A Technical Critique of Fifty Software Patents*, 9 MARQ. INTELL. PROP. L. REV. 249, 274 (2005).

section V(C), had much stronger blocking potential and, if asserted, could have enabled Postini to exclude competitors from the hosted services industry. Postini initially engaged in a little saber-rattling, but quickly moderated its behavior and repositioned the patent as defensive rather than offensive.[165] We believe that the explanation for this turnaround is the powerful effect of social disapproval from the Internet community—a moral sanction made doubly effective because of the repeated game nature of the industry and its incremental innovations. The mechanisms by which the patent system is adapting to software may be unique, but the process is not. The patent system has a repeated history of adapting to radical new technologies: a period of turbulence, then of accommodation, and finally of normalcy.

¶ 108 We argued above that broad patents have not inhibited the entry of start-up firms in the anti-spam industry. We next examine whether there are too many narrow patents. Is the anti-spam prospect being sliced and diced into such small areas that firms are having to take out licenses on so many patents that their activities are being impeded? We can find no evidence of such a phenomenon. In section V we examined in detail the patent licensing activities of the forty-two largest firms: the number of patent licenses was too few for any general conclusion to be drawn, but clearly the fear that firms will have to license large numbers of patents is exaggerated.

¶ 109 It is possible that as the anti-spam problem evolves, there will be many more patents—we estimate that the number of anti-spam patents currently in the application stage is in the range of 120 to 200.[166] If the anti-spam prospect were static, as in the mineral claim analogy, we think that the constantly rising number of anti-spam patents would be a cause for concern. However, because spam techniques are constantly evolving, the anti-spam prospect is also constantly growing and one would expect the number of patents to increase pari-passu. As spam techniques become obsolete, however, so do their patents—long before their statutory twenty year term—and there are, even now, several patents "on the books" that are effectively obsolete.[167]

¶ 110 A problem related to narrow patents is over-fishing: broad patents discourage over-fishing, whereas narrow patents might encourage it. This issue is overshadowed by open-source developments. As we noted in section VI, there is a large number of open source projects in all of the major anti-spam categories. Perhaps some over-fishing is therefore occurring, but there is also a Darwinian process by which the better solutions (such as SpamAssassin and Vipul's Razor) are being selected. At the present time, we do not have sufficient data to make a judgment as to whether or not open-source anti-spam solutions are socially optimal. Given these intractable problems of analysis, we are unable to comment on the impact of the patent system on duplicate R&D investments,

---

165 Roberts, *supra* note 119; Sturdevant, *supra* note 121.

166 This is an estimate based on a search of United States Patent and Trademark Office's patent database conducted by the authors.

167 *See, e.g.*, U.S. Patent No. 5,377,354 (issued Dec. 27, 1994), the first patent for content-filtering using lexical analysis and user-defined rules based on keywords entered by the user, and can be considered obsolete as it has been superseded by Bayesian filtering. See also U.S. Patent No. 5,619,648 (issued Apr. 8, 1997), assigned to Lucent Technologies, which proposed a simple sender verification method for anti-spam which may also be obsolete as it creates an unacceptable burden on the sender.

other than that we see no cause for anxiety.

¶ 111 One objective of the patent system is to convert intangible inventions into tradable intellectual property. Commercial anti-spam products typically need to make use of several collateral technologies, and there is a concern that the transaction costs of patent licenses could be excessive. As we saw in section V, there is a thriving and apparently efficient market in anti-spam technologies. Occasionally, patent owners choose to license a patent by itself, but far more often vendors choose to license an OEM development kit. Such development kits not only make the process of software integration easier, but they bundle any patent transaction costs—whether there are one, two, or ten patents involved. As we noted earlier, the reduction of patent transaction costs could well be a stimulus to the market for software components.

¶ 112 Finally, one of the most surprising observations of this empirical study is the peaceful and productive co-existence—occasionally symbiosis—of the open- and closed-source worlds. The reason we say surprising, is that the open-source community is often portrayed, and often is, hostile to proprietary software and software patents. The community is hostile because it sees patents as a threat that could foreclose their participation in the software industry. For example, Richard Stallman, a stern software patent critic wrote in 1992 that "[s]oon new companies will often be barred from the software arena—most major programs will require licenses for dozens of patents, making them infeasible."[168] More than a decade since these fears were expressed, they seem to be unfounded. Our study demonstrates that there are two complementary anti-spam markets. First, there is an open-source market where, even though technologies can be freely appropriated, there is a virtual market for technological superiority, and a real market for productized solutions. Second, there is a conventional market for proprietary, sometimes patented, technologies. Firms have a choice: they can use an open source solution for "free," they can buy a productized open-source solution, or they can buy a license for a proprietary solution. We believe that firms make rational choices. For some vendors, SpamAssassin is free and good enough to serve their purposes; others will chose to license a patented product such as Symantec Brightmail, presumably because they think it is a better solution and worth the money.

[168] Stallman & Garfinkel, *supra* note 162, at 17.

VIII. APPENDIX A: A TECHNICAL ANALYSIS OF ANTI-SPAM PATENTS

¶ 113 We studied a set of 113 issued anti-spam patents which we obtained from the USPTO site,[169] starting from a partial list of anti-spam patents,[170] which we extended based on various searches using the names of major anti-spam companies or key-words like "spam," "junk," and "unsolicited email." All these patents are either "pure" anti-spam patents in the sense that they deal specifically with anti-spam, or are anti-spam-related patents in which dealing with spam is a side-effect of a more general security solution.

A. Introduction: How Email Works

¶ 114 Email systems use a "client-server" architecture. An email client (or mail client) is an application that runs on a desktop PC that lets the user send and receive email, and organize it within various folders. Email that a client receives is called inbound while email that a client sends is called outbound. Examples of popular graphical mail clients are Microsoft Outlook, Mozilla Thunderbird, and Netscape Messenger.

¶ 115 An email server (or mail server) is a computer that handles the storage and the transfer of messages from/to the local mail clients (within the organization) and from/to other mail servers (across organizations). It includes a database of local user accounts with authentication (password) information and other information to communicate with other servers. Examples of popular mail server programs are MS Exchange, Sendmail, and qmail. Mail servers are typically managed by a "postmaster" within an organization or an Internet Service Provider which provides email (and other) services to remote users through a telephone line, cable, or broadband connection.

¶ 116 While traveling to a mail server or mail client, a message may go through a firewall—a security program which isolates the resources of a private network or of a PC from users in other networks. Typically, organizations use firewalls on dedicated servers or "appliances." Firewalls prevent external users from accessing private data resources and also control the Internet resources an organization's users have access to (for example, by blocking inappropriate websites). Light firewalls are now available to protect PCs. A firewall is able to log and screen all inbound and outbound network traffic and determine, using packet filtering rules, whether to forward it toward its destination according to the organization's security policy.

¶ 117 The Internet Engineering Task Force (IETF) has standardized several protocols which are widely used for email client-server and server-server communication on the Internet. To retrieve messages from a server, a mail client can use either the Post Office Protocol (POP) or the Internet Message Access Protocol (IMAP). For a client to send a

[169] USPTO home page, http://www.uspto.org.

[170] [ASRG] US Spam Patents: Partial List, http://www1.ietf.org/mailarchive/web/asrg/current/msg05356.html (last visited Aug. 24, 2005).

message to a server, or for a server to send a message to another server, there is the Simple Mail Transfer Protocol (SMTP). The Extended Simple Mail Transfer Protocol (ESMTP) specifies extensions to the original protocol for sending messages with graphics, audio and video files, and text in various national languages.

¶ 118 The original specification of SMTP was published in 1982 and suited the needs of the Internet users. At that time, the Web did not exist and the Internet was used essentially by universities and research organizations, but not yet for commercial purposes. Thus, SMTP remains inherently insecure[171] because it is easy for a user to create messages with fake sender email addresses. Also, as noted in section II, SMTP allows mail servers to act as open relays, allowing the sending or receiving email that is not for or from a local (authenticated) user. This is useful, for instance, for traveling users to access their corporate network by going first to a local ISP which forwards their messages to their corporate server. However, open relays can be used by spammers to send large volumes of email without being (easily) detected. This is because messages are not authenticated. There are proposals for message authentication and other security features in the SMTP and other email protocols[172] but they will not be standardized for some time, and will take longer still to be universally adopted because of the current installed base.

B. Taxonomy of Anti-spam Technologies

¶ 119 Before legal or standardized protocol solutions become effective to eradicate spam, we are left with technology. The main objective of anti-spam technologies is to block spam, i.e., detect and mark it at the mail servers or clients. But another important objective is to reduce the overhead for users and postmasters in dealing with spam, for instance, in removing marked messages that have ended up in a spam folder. But the main problem with technology is that spam is defined based on the user's perspective (one user's spam may be another user's legitimate mail). Blocking spam can also create the problem known as false negatives and false positives. A false negative is a spam message that is not blocked and makes it to the mail client. Thus, it is up to the user to detect and remove it. A false positive is a legitimate message that is incorrectly blocked. This is much more inconvenient for the user who will not even see the message until checking the spam folder. Thus, besides spam capture efficiency, an important measure for comparing anti-spam technologies is false positive efficiency.[173]

¶ 120 Designing an anti-spam solution for an organization is complex because it requires the combination of various anti-spam technologies and the implementation of an anti-spam policy. Ideally, anti-spam should be one aspect of a more general security

---

[171] Memorandum from The Internet Society to the Internet Community requesting discussion and suggestions, RFC 2821: Simple Mail Transfer Protocol (April 2001), http://www.ietf.org/rfc/rfc2821.txt.

[172] *See, e.g.*, Memorandum from the Internet Society to the Internet Community requesting discussion and suggestions, RFC 2487: SMTP Service Extension for Secure SMTP over TLS (Jan. 1999), http://www.faqs.org/rfcs/rfc2487.html.

[173] Osterman Research, White article on A Comparative Analysis of Leading Anti-spam Solutions (2004), http://www.ostermanresearch.com/whitearticles/download11.htm.

architecture that must deal with other Internet-related threats such as viruses in email or denial-of-service attacks of Web servers.[174] This explains why classifying and studying anti-spam patents is not easy. First, these patents may deal with one specific technology or a combination of technologies to be integrated in a more complete system including mail servers and firewalls. Second, they may deal with issues which are related to spam as well as other security threats. Finally, anti-spam technologies rely on more general technologies such as natural language processing (to analyze the text content of messages), information retrieval (to classify and rank messages), artificial intelligence (to train classifiers), etc.

¶ 121 We define four classes of anti-spam techniques in our taxonomy, discussed below.

1. Access control

¶ 122 Access to the email and network resources can be controlled in several ways at the mail client or mail servers, at different times in the transfer of messages. This can be done through header checking, blacklists and whitelists, and sender and message authentication.

¶ 123 Message header checking is usually done by mail servers. A typical check is to verify that the sender's address has not been forged by ensuring that the domain in the address is valid or that the sender's server has a correct Domain Name Service (DNS) setup. However, a server that is not configured properly may create false negatives.

¶ 124 Blacklists (also called blackhole lists or block lists) and whitelists of senders can be used at the mail server or mail client to deny or allow the delivery of messages. These lists can contain various attributes such as sender user name, sender address, receiver user name, receiver domain, server name, server address, etc. These lists can be stored locally in the mail server database or even in a client file. To be effective, this approach requires the lists to be up-to-date.

¶ 125 In addition to the previous techniques, senders can be required to provide authentication information, e.g. some password as part of their address, so their messages can be also authenticated. The authentication information is then compared with that in a database of registered users.

2. Content filtering

¶ 126 Content filtering involves looking inside the messages with more sophisticated techniques that often rank the message as potential spam. The main techniques are fingerprint analysis, lexical analysis, Bayesian filtering, heuristics, and checksum filtering.

¶ 127 Fingerprint analysis exploits the fact that spam messages are typically sent in very large numbers of copies during a period of time (e.g. a marketing campaign). Once a

[174] *See, e.g.*, EGAN, *supra* note 26.

message has been identified (by a user) as spam, a fingerprint of it (obtained by creating a unique signature from its content) can be stored in a database. Fingerprint analysis is thus useful to detect spam that is already known.

¶ 128 Lexical analysis deals with unknown spam by analyzing the content of the message. This is done by extracting elements (combinations of words or phrases) which are used as input to filtering rules which assign a weight to each element. Then the individual weights are combined to yield a score for the message as potential spam. The filtering rules are user-defined and their quality is critical for spam capture efficiency.

¶ 129 Instead of user-defined filtering rules which may be hard to specify, a better approach[175] is Bayesian filtering.[176] Because it can adapt automatically to a user's definition of what is and is not spam (which may vary much from one user to another),

¶ 130 Bayesian filtering has become the most effective content filtering method. It requires an automatic training period whereby a large number of spam and legitimate messages are analyzed to produce a database of all the words found with the probability that a particular word belongs to a spam message. Then, using this database, it is easy to compute the overall probability that a new message is spam. The efficiency of a Bayesian filter thus depends heavily on how well it is trained. Also filtering message headers in addition to message bodies improves spam capture efficiency.[177] Bayesian filtering can also be used for classifying messages in various categories such as urgent, personal, etc.

¶ 131 The heuristics approach does not use a single filtering method but applies successive tests to a message to infer whether it is spam. The tests can be based on access control methods or the content filtering methods discussed above. Each test gives a score which is added to the current sum of scores and the process ends when the score exceed a given threshold. Thus, the most selective or least costly tests should be done first to avoid wasting resources.

¶ 132 Checksum filtering is a collaborative approach involving several mail servers which count the number of times they have seen the same message, identified by a checksum. When the count for a particular message is high, there is a probability that the message is spam. However, it may also be a legitimate message to a mailing list. Thus, to avoid false positives, it is important that the sender be on a whitelist.

3. Sender verification

¶ 133 Sender verification is a method whereby the receiver of a message requires the sender to perform some action in order to prove that the message is legitimate. Thus, if a spammer has used a fake return address, the message will be easily blocked as spam.

---

[175] Paul Graham, *A Plan for Spam*, PAULGRAHAM.COM (Aug. 2002), http://paulgraham.com/spam.html.

[176] Named after Thomas Bayes' famous probability theory which was discovered after his death in England in 1761.

[177] Le Zhang, Jingbo Zhu & Tianshun Yao, *An Evaluation of Statistical Spam Filtering Techniques*, 3(4) ACM TRANSACTIONS ON ASIAN LANGUAGE INFORMATION PROCESSING 243, 243-269 (2004).

¶ 134 The main sender verification methods are sender compute and challenge/response. With sender compute, the receiver of a message requires the sender to perform some non-trivial computation and convey the result within the message, typically in the message header. Thus, by performing the same computation (or using the stored result), the receiver can detect whether the message is legitimate. This method is often used with whitelists to avoid the burden of repetitive computation from regular senders. Instead of performing a computation, the sender may also be required to use a password, or even an electronic stamp and thus be charged a fee.

¶ 135 With challenge/response, the receiver requires the sender to send a response to acknowledge he or she is not an automated (spam) email sender. While the response has not arrived, the message stays in quarantine in a staging area. It is delivered to the receiver only after receiving the response.

¶ 136 Other sender verification methods include electronic stamps or signing messages.

4. Email management

¶ 137 We put in this class the technologies that deal with the efficient and automated management of the resources involved in an anti-spam solution. Thus, they can integrate the anti-spam techniques described above in a more complete system in many different ways. The technologies in this class are typically larger grain in that they use several components. Examples of resources that need be managed are spam messages, email addresses, filtering rules and filtering events.

¶ 138 An example of a useful technology in this class is for the management of spam messages, e.g. which can be automated using self-removable messages or using a one-click interface that requests removal from a spam mailing list. Another example is the hiding of private email addresses and the automatic forwarding to real addresses, after anti-spam filtering. Managing filtering rules can also be made easier using a graphic user interface.

C. The Patent Set

¶ 139 Table A1 lists the number of anti-spam patents per assignee. It gives for each assignee, the numbers of patents granted for each class of anti-spam technology (access control, content filtering, sender verification, and email management), ordered alphabetically by assignee name. There are also several patents granted to individuals, in which case we prefix the first author name by the key-word "Ind:". A first observation about the table is that there are many different players, in addition to email or anti-spam management start-up companies. Besides the 9 individuals who still hold patents (i.e., they have not transferred it to a company), there are several large computer and/or software companies (Apple, HP, IBM, Intel, Fujitsu, Microsoft, Minolta, Siemens, Sun, Xerox), several telecom companies (AT&T, Lucent, Motorola, Nokia, Telecom Systems), and smaller email management or anti-spam companies (Brightmail, Cloudshield, Mailfrontier, Nixmail, Postini, Secure Computing, Sendmail, Surfcontrol, Trend Micro, Tumbleweed, etc.). The large companies with a long tradition of patenting have much

higher numbers of patents than smaller companies. For instance, IBM has 11 patents, AT&T has 7, Microsoft has 7, HP has 4 (coming from Digital), and Lucent and Xerox have 3 each. Smaller email management or anti-spam companies typically have one patent. The exceptions are Brightmail and Network Associates, which each have 3. The variety of the companies that hold anti-spam patents suggests that this is not so much a niche market but an important aspect of a more general security solution.

¶ 140 A second observation from Table A1 is that there is a good balance of the total numbers of patents per class. This gives us a good level of confidence in our taxonomy of anti-spam technologies from the viewpoint of patents. There are 30 patents in access control, 35 in content filtering, 18 in sender verification, and 30 in email management. The fact that the highest number of patents is in content filtering (38) confirms it is the most important and difficult way to fight spam.

¶ 141 Table A2 summarizes our set of anti-spam patents, ordered by technology (AC: access control, CF: content filtering, SV: sender verification, MM: email management) and patent number (patent#). For each patent, we list the company name, the year it was granted, the technology class, the number of citations by other patents, the number of claims, and the title. We also added a comment whenever applicable. This comment may simply indicate that a patent is a continuation or a division of a previous patent, of the same grantee. We also comment on some patents' scope. From a legal standpoint, the scope of a patent can only be determined by a Court in the event that the patent is infringed or challenged. As technical experts, we cannot evaluate the scope of all patents. However, we do indicate, for the patents with a large number of claims (respectively, small number of claims) if we feel they have a potentially broad scope (respectively, narrow scope).

¶ 142 An interesting observation is that the large majority of patents is fairly recent. Only 23 patents were issued before 2000, the oldest one being patent 5,208,748 issued to Action Technologies in 1993.

¶ 143 In the rest of this section, we discuss in more details some representative patents for each class. The criteria we used to select those patents are interesting features such as leading anti-spam firm, potentially broad or narrow scope, foundation patent (having a strong impact on subsequent innovation, in general, with a high number of citations).

1. Access Control Patents

¶ 144 Access to the email and network resources can be controlled in several ways at the mail client or mail servers, at different times in the transfer of messages. This can be done through header checking, blacklists and whitelists, and sender and message authentication.

¶ 145 Xerox has three patents in access control: patent 5,513,126 issued in 1996 followed by two continuation patents, 5,657,461 and 5,689,642, in 1997. We discuss patent 5,513,126 only which has a high number of cites (164). The patent deals with the general problem of access control over various network resources such as mail servers,

printers and fax machines. It describes a method for a sender to send information to a receiver on a network resource as defined in a receiver profile. The receiver profile establishes the properties and mode for receipt of information for receivers on the network. It is published in a network repository for all network users or is accessible by selected users in the network. Receivers have additional control over senders by defining some network resources as having priority of access such as direct or delayed access. Consequently, receiver profiles provide a variable receiver definable link to senders using multiple forms of media as well as multiple network configurations. Although the patent is not specifically designed for anti-spam, making the profiles accessible by selected senders in the network provides an early form of whitelist. In addition since it has a high number of cites, this patent can be considered foundation.

¶ 146 AT&T has two complementary patents issued in 1999 in access control. Patent 5,905,777 describes one of the first methods based on whitelists. It is used by a mail server to separate between legitimate messages and spam messages, but also to forward legitimate messages to a destination specified by the receiver. This illustrates how anti-spam processing can be included in a more general email solution. The email server has a database which stores, for each mail receiver record, a whitelist of selected email senders and forwarding destinations for the email messages such as the network address of another computer, fax machine or other appliance. The method has a number of specific steps for which some particular network is used. For instance, when a legitimate message is detected, the mail server sends the receiver an alert message over a radio communication system to ask for the choice of a forwarding destination address. This makes the patent's technical breadth narrow.

¶ 147 Patent 5,930,479 also from AT&T in 1999 describes a method to authenticate messages. To convey authentication information within a message, the receiver address is extended with an access address which is a data string that is hard to guess, e.g. using randomly generated numbers. Each legitimate sender is allowed to know one or more of these access addresses which the email server maintains in a file for determining whether a message is authorized. Although the patent has medium technical breadth, its scope with 68 claims on many variations of the method is potentially broad.

¶ 148 Patent 6,052,709, issued to BrightLight in 2000 (and later on reassigned to BrightMail), describes a method based on blacklists to be used at a mail server or client. The way blacklists are automatically generated and updated is quite original, based on spam probes. A spam probe is an email address specifically selected to make its way onto as many spammer mailing lists as possible. The spam probe is also selected to appear high up on the spammers' lists in order to receive mailings relatively early in the spammers' mailing process. For example, the spam probe address may be selected to appear at the top of an alphabetized mailing list (e.g., "Aardvark @aol.com"). The mailboxes corresponding to the spam probe email addresses are monitored for incoming email by a spam controller. Upon receipt of incoming email addressed to the spam probe addresses, the spam controller analyzes the received messages to identify the source of the message, extracts and processes the source data from the received message, and sends an alert message, with the source data and spam filtering instructions, to all mail servers in the network. When receiving an alert message, a mail server processes the spam

message using the filtering instructions and updates its blacklist of spam sources which it will use to detect future spam messages from that source. The method is effective to deal with spam mailing lists, but not with forged sender email addresses.The patent is rather short (13 pages) and has relatively narrow scope with 7 claims.

¶ 149 Patent 6,321,267, issued to Escom in 2001, describes a general method for access control at a firewall using an active filter proxy. The proxy actively probes remote hosts that attempt to send mail to the protected mail server in order to identify dialup PCs, open relays, and forged email. The method uses multiple defense tests to detect spam and viruses including: connect-time filtering based on sender addresses maintained in whitelists and blacklists, identification of dialup PCs attempting to send email, testing for permissive (open) relays, testing for the validity of the sender's address, and message header filtering. A sender's message must successfully pass through all these steps, or it is rejected and logged, and not sent to the email server. Because these tests are performed at the time of the initial data connection, they characterize the remote host as it is configured at that time, thus avoiding the latency problems of static blacklists Furthermore, the sender addresses of rejected messages are added to blacklists so subsequent mail from the same host can be rapidly blocked. The patent is 43-page long, has good disclosure and high technical depth. However, the high number of claims and the generality of the method with multiple tests makes its scope potentially broad.

¶ 150 Patent 6,865,671 issued to Sendmail in 2005 describes an authentication method for supporting relaying in mail servers. The method addresses the specific security problem of SMTP which allows mail servers to act as open relays which are used by spammers to send large volumes of email without being detected. This is because clients of messages are not authenticated. It is therefore of utmost importance for Sendmail which ships the most used open source mail server. Upon receiving a message, the method first checks whether the client has been authenticated. If not, the decision of whether relaying is allowed may be subject to other rules in the system, such as whether the user currently resides behind the organization’s firewall. If the client has been authenticated, the method can allow relaying for everyone who has a certificate signed by certain certificate authorities. The patent is short (13-page long), has medium disclosure and medium technical depth and is relatively narrow.

2. Content Filtering Patents

¶ 151 Patent 5,377,354 issued to Digital in 1994 (reassigned to HP) is the first one on content- filtering using lexical analysis. It describes a method for sorting and prioritizing email messages using user-defined filtering rules. Thus, although not mentioned in the patent description, it can apply to spam detection. The user-defined rules are based on keywords entered by the user. By applying the user-defined rules, the received messages are assigned a priority and sent to a main folder store, forwarded or put away as appropriate. Considering its originality and high number of cites (121), this patent may be considered a foundation patent. However, it is rather simple, short (9-page long) and has low technical depth and disclosure.

¶ 152 Patent 5,999,932 issued to Bright Light in 1999 describes what is probably the

first heuristics method which applies successive tests to a message to infer whether it is spam. Considering its relatively high number of cites (56 since 1999), it may also be viewed as foundation. The tests are based on header checking and can include content filtering. To perform header checking, a user inclusion list includes identification header data for identifying email desired by the user. If an email message data matches corresponding identification data from the user inclusion list, the message is considered legitimate. If no match is detected, the method performs one or more heuristic tests to determine whether the message may be of interest to the user or if it is spam. The heuristic tests can include content filtering. The patent is relatively simple (15-page long), has medium technical depth and disclosure, and is narrow.

¶ 153 Patent 6,161,130 issued to Microsoft in 2000 describes the first probabilistic method for content-filtering and thus can be considered foundation. It uses a probabilistic classifier trained on prior content classifications using sets of legitimate and spam messages. The patent mentions various ways to implement the classifier, including Bayesian networks and neural networks, but describes in details the use of a support vector machine. Through a resulting quantitative probability measure, i.e., an output confidence level, produced by the classifier for each message and subsequently compared against a predefined threshold, the message is classified as either spam or legitimate mail and stored in a corresponding folder for subsequent retrieval by the receiver. The patent is 28 page-long, has high technical depth, medium disclosure and narrow technical breadth. However, it has a large number of claims (65) and one claim (50) is rather broad, claiming various implementations of the probabilistic (Naive Bayesian classifier, limited dependence Bayesian classifier, Bayesian network classifier, decision tree, a support vector machine, or use of content matching.). Thus, it has a potentially wide scope.

¶ 154 Patent 6,330,590 issued to William D. Cotton in 2001 and acquired by Commtouch describes a fingerprint analysis method to be used at a mail server or client. Each received e-mail message, after elimination of source and sender identification, is scanned and coded to provide a unique signature from its content. The signature code is typically calculated numerically, using a checksum in a 16-bit cyclic redundancy check routine. This kind of numerical signature can be produced quickly and eases comparison of all inbound email messages. Detecting a set of typically three identical messages (having the same signature), going to different email addresses means spam. Then, the spam message signature is stored in the signature database for use in eliminating future such spam. The patent is short (7 page-long), has low technical depth, low disclosure and narrow technical breadth.

¶ 155 Patent 6,732,157 issued to Network Associates in 2004 describes a heuristics method which filters message content. The email messages may be filtered as being unwanted based on a comparison involving the probability and a threshold which can be user-defined. The method uses a combination of techniques including compound filters, paragraph hashing, and Bayes rules. The compound filters may use Boolean logic or conditional logic. The content of the messages may be normalized prior to utilizing the paragraph hashing. Such normalizing may include removing punctuation of the content, normalizing a font of the content, and/or normalizing a case of the content. The compound filters and paragraph hashings may have an associated level so the higher-

level checks are applied first. The use of Bayes rules occurs after the use of the compound filters and the paragraph hashings and may include identifying words of the email messages. This may further include identifying a probability associated with each of the words. Optionally, the probability associated with each of the words may be identified using a Bayes rules database. The patent is 19 page-long, has medium technical depth, medium disclosure and medium technical breadth.

¶ 156 Patent 6,778,834 issued to Nokia in 2004 describes a key-word based method to filter messages sent to a mobile terminal and alert the user when a message of interest arrives. The user of a mobile terminal can enter the keywords indicative of the desired categories of messages. When the mobile terminal receives a message, the keywords associated with the message are compared with the user-entered keywords. If there is a match, a banner portion of the message is displayed on the mobile terminal which issues a sensible alert. The user may request to view the message body associated with the banner portion or to take actions regarding the message body, such as storing the message body in the mobile terminal or calling a telephone number contained in the message body

¶ 157 If there is no match between the user-entered keywords and an incoming message, the message may be displayed in a de-emphasized manner ("grayed out"), or portions of the message may be omitted from display. This constitutes a light way to filter spam while avoiding false positives. The patent is 20 page-long, has medium technical depth, medium disclosure and narrow technical breadth.

¶ 158 Patent 6,845,374 issued to Mailfrontier in 2005 describes a general method for selecting relevant electronic documents to recommend to a requester. The method applies to all kinds of electronic documents and thus can be used for automatic classification of personal email and automatic routing of customer email. Thus, although not explicitly mentioned in the patent description, it can be used to filter spam. The method uses automatic classification of documents based on key-word extraction and clustering into categories with a high measure of statistical relevancy. The method is either automatically or manually invoked and it presents the recommendation set in real-time in different ways, e.g. notification, alert, fax, voicemail, email, etc. The recommended set may also consist of Internet bookmarks or subscriptions to publications for a community of interest. The patent is 20 page-long, has low technical depth, low disclosure and medium technical breadth.

3. Sender Verification Patents

¶ 159 Patent 5,208,748 issued to Action Technologies in 1993, is the oldest patent in our set. It describes a general method for email conversation management by defining the explicit types of communications between participants. It does not specifically address the problem of spam although some aspects may be useful for sender verification. The method is based on recording the "moves" in conversations with a pre-defined structure in a database. Each record of a "move" is identified with a particular participant that produced it, and others who are involved in the conversation and will receive it. This provides for some form of sender verification during conversations. The patent is 171 page-long, has high disclosure, high technical depth and wide technical breadth.

¶ 160 Patent 5,619,648 issued to Lucent in 1997 is the first one to propose a simple sender verification method for anti-spam. It has also a high number of cites (135) and thus may be considered foundation. The method requires the sender to add some non-address information to specify the receivers of an email message. Then a mail filter for a given receiver uses the non-address information in the email message to determine whether the message should be provided to the given receiver or blocked. The patent is 10 page-long, has low disclosure, low technical depth and narrow technical breadth.

¶ 161 Patent 6,195,698 issued to Digital in 2001 (reassigned to HP) describes a method for challenge/response in a server to detect automated agents that may send bulk email messages or request a search engine to index useless Web pages. Upon receipt of a client request, the server generates a predetermined number of random characters to form a string. The string is randomly modified either visually or audibly to form a riddle. The original string becomes the correct answer to the riddle. The server then challenges the client by sending it the riddle. In response, the client must send a user's guess for the correct answer. The server determines if the guess is the correct answer, and if so, the access request is accepted. If the correct answer is not received within a predetermined amount of time, the connection between the client and server computer is terminated by the server on the assumption that an automated agent is operating in the client on behalf of the user. The patent is 18 page-long, has medium disclosure and technical depth, and medium technical breadth.

¶ 162 Patent 6,393,465 issued to Nixmail in 2002 describes a challenge/response method to detect spam. Upon receipt of a message, the email server attempts to contact the purported sender in order to verify that the identified host computer actually exists and accepts outgoing mail services for the specified user. The routing history is also examined to ensure that identified intermediate sites are also valid. Likewise, seed addresses can alert an e-mail provider to potential mass mailings by reporting when mail is received for ghost or non-existent accounts. The patent is 16 page-long, has low disclosure and technical depth, and narrow technical breadth.

4. Email management

¶ 163 Patent 5,283,856 issued to Beyond in 1994 is the first one to propose a rule-based method for managing incoming email, and in particular, block spam. It has also a high number of cites (140) and thus can be considered foundation. The rule mechanism supports the general "When-If-Then" event-driven, conditional, action-invoking paradigm which permits the definition of a number of events considered to be significant events upon which to trigger actions in the email system. Each particular event may be associated with a specific email message or rules to promote efficient mapping of messages, events and rules so that only rules associated with a specific event are invoked upon occurrence of the event. Only the relevant rules, i.e. those associated with a satisfied event, need be further processed. In particular, rules can be defined for anti-spam to perform header checking or content filtering. Typical actions include forwarding, filing and deleting the mail message(s). A graphical user interface to a structured rule editor facilitates synthesis of rules by a user via a transparent rule engine. The patent is 30 page, high technical depth, high disclosure (a detailed rules syntax is given in an appendix) and

wide technical breadth.

¶ 164 Patent 6,094,681, issued to Siemens in 2000 describes a method for automated event notification. It provides for remote user notification of an event when the user is determined to be unavailable to locally receive the notification. The method includes receiving data and analyzing the content of the data using a data filter. The data filter is configured to detect an indication of a predetermined event within the data. If an event is detected, the data filter activates a local event monitor which performs the automatic notification to the user. The data filter is capable of analyzing data included in email messages, web page updates transmitted to a web browser of the computer, scheduling updates and requests transmitted to an electronic calendar, and scheduling reminders transmitted by the electronic calendar. Thus, although not specifically designed for anti-spam, the method can be used to notify users that spam events had occurred. The patent is only 9-page long, has low technical depth and low disclosure, and medium technical breadth.

¶ 165 Patent 6,650,890 issued to Postini in 2003 describes a method of providing an email preprocessing service in an email client or server. The service can detect and detain damaging or unwanted messages, such as spam or viruses, and route email messages from various sources to wired and wireless destinations, in addition to the intended recipient email address, in various formats. The service uses stored user profiles and can perform header checking and content filtering of messages. The patent is 15-page long, has low technical depth and low disclosure, and medium technical breadth.

¶ 166 Network Associates has two complementary email management patents issued in 2004. Patent 6,757,830 describes a method that computes a minimum delay period before releasing an email message to a receiver so as to make sure that all most recent tests have been applied. This avoids that spam messages or viruses, being discovered for some other users, make their way to the receiver. Prior to release of the email message upon expiry of the minimum delay period, a check is made that the most up-to-date antivirus and anti-spamming tests have been applied to the email message. Characteristics that may be used to determine the minimum delay period applied include sender characteristics, recipient characteristics, attachment type characteristics and message content type characteristics. The patent is 16-page long, has low technical depth and low disclosure, and medium technical breadth. It has 45 claims and has potentially broad scope.

¶ 167 Patent 6,802,012 also from Network Associates describes a method for the efficient scanning of files for unwanted properties, such as containing viruses or being spam email. It allocates a priority to pending scan requests based upon the identity of a computer user associated with the scan request. In the case of scan requests associated with emails, the sender or recipient computer user may be used in the allocation of a priority level for improving the time of the scan request. The patent is 15-page long, has low technical depth and low disclosure, and medium technical breadth.

¶ 168 Patent 6,829,654 issued to Cloudshield in 2004 describes a method for enhancing the infrastructure of a network, in particular for web access and email. The method uses multiple edge servers and edge caches in the network so as to cover and monitor all

points of presence. The edge servers selectively intercept domain name translation requests generated by downstream clients, coupled to the monitored points of presence, to subscribing Web servers and provide translations which either enhance content delivery services or redirect the requesting client to the edge cache to make its content requests. Furthermore, the method provides network traffic monitoring in order to detect malicious or otherwise unauthorized data transmissions. Through the provision of additional processing capabilities within the edge servers, service applications such anti-spam filtering and spam source detection, can be implemented. The patent is 26-page long, has low technical depth and low disclosure, and medium technical breadth.

¶ 169 Patent 6,836,792 issued to Trend Micro in 2004 describes a general technique to extend an existing email server with add-on services. Examples of add-on services include anti-spamming, virus scanning, anti-spamming, paging, auto-redirection of received messages to another domain, auto-reply back to the sender with a pre-selected message, conversion of the information contained in the message to voice or fax or another medium, mailing list, security encryption prior to forwarding to the destination email server, etc. The email server maintains a user profile database which records the add-on services which are subscribed by the users. Upon receiving a message, the email server checks using the user profile database if the receiver has subscribed to at least one add-on service, performs the subscribed services and then sends a post add-on service message to the receiver. The proposed technique is rather simple as it mainly involves selecting and performing a subscribed service. It is general as it can include any kind of add-on service, not limited to anti-spam. The patent is 14 page-long, has low disclosure and technical depth and has narrow technical breadth.

D. Conclusions

¶ 170 This discussion of representative patents in our set of 113 issued anti-spam patents illustrates why there are so many in a relatively short time (most of these patents (90) have been issued since 2000). A first reason is that an effective anti-spam solution requires the combination of various anti-spam technologies. Another important reason is that spam cannot be treated in isolation from other security issues such as viruses and is one aspect of a more general communication architecture including email and web sites. Thus, there are many different patents simply because they either consider one particular anti-spam technology (access control, content filtering, sender verification) or combine anti-spam processing with other email management issues, e.g. user notification, add-on pre-processing services, etc. Even within a given class, e.g. access control, there are many different ways and variants to solve the problem. We found that only 6 patents have potentially wide scope.

¶ 171 We found that 7 patents may be considered foundation because they have both a high number of cites and typically invent a new way to deal with spam. However, even the most inventive patents do not invent new basic techniques. Rather they borrow basic techniques coming from academic research in various domains: natural language processing (to analyze the text content of messages), information retrieval (to classify and rank messages), artificial intelligence (to train classifiers), etc. And they apply it in an inventive way to come up with a real solution. Thus, as for many research areas for which

there is a killer application, there is now much synergy between academic research and industrial research in order to fight spam and related email threats. This is exemplified by the many recent conferences in Internet and Web security, ISP, etc. which regularly feature anti-spam sessions. In 2005, Stanford University is organizing a conference specifically on anti-spam.[178]

[178] The 2005 Conference on Email and Anti-spam, July 21-22, 2005 at Stanford University, covering all aspects of improving email systems and methods for stopping spam.

IX. APPENDIX B: BAYESIAN-RELATED PATENT APPLICATIONS

¶ 172 Most Bayesian content filtering solutions build on a "raw" Bayesian approach which yields a score for spam and a score for non-spam. Many refinements and extensions of this basic method have been proposed by spam researchers to improve the probability of success while reducing false positives. These extensions include using more complex methods for computing the probability that a message is spam (for example, by introducing a "maybe spam" intermediate state between spam and non-spam).[179] Such extensions are public domain and used in most open-source Bayesian filters such as Bogofilter, Mozilla, SpamBayes, and SpamAssassin. It is very likely that the Bayesian filtering methods will continue to improve as a result of using better statistical methods.

¶ 173 In our set of 40 Bayesian-related anti-spam patent applications, none describes a radically new Bayesian filtering method (for example, by using a different probability measure). Rather, they describe diverse approaches of using Bayesian analysis for anti-spam solutions or for more general purposes (such as security or document retrieval applications) which include anti-spam.

¶ 174 We can further classify these Bayesian-related applications in three sub-classes: Bayesian analysis, Bayesian filtering, and combined analysis. Bayesian analysis is a general technique, not limited to anti-spam, to automatically learn from a training set of data and yield scores for incoming data. Applications in this class use Bayesian analysis for specific tasks such as rating a document according to a learned user's notion of "quality", automatic learning of users' profiles for targeted marketing, or message preprocessing. As a particular case or side effect, they can also be used to reduce spam. Bayesian filtering refers to the use of Bayesian analysis for spam content detection. Applications in this class make use of Bayesian filtering, possibly with other filtering techniques such as key-word extraction, in an anti-spam solution. Combined analysis refers to the recent use of multiple techniques for improving learning from training data.[180] This approach exploits the equivalence of Bayesian logic and fuzzy logic[181] to better deal with uncertainty and adaptation (to the incoming messages). Applications in this class use fuzzy logic for machine learning, and can apply it to anti-spam as a particular case.

¶ 175 Table B.1 describes our set of 40 Bayesian-related patent applications,[182] ordered by technology (BA: Bayesian analysis, BF: Bayesian content filtering, CA: combined

---

[179] Gary Robinson, *A Statistical Approach to the Spam Problem*, LINUX J. (Mar. 1, 2003), *available at* http://www.linuxjournal.com/article.php?sid=6467.

[180] A. B. BADIRUI & J. Y. CHEUNG, FUZZY ENGINEERING EXPERT SYSTEMS WITH NEURAL NETWORK APPLICATIONS (John Wiley & Sons 2002).

[181] A superset of conventional (Boolean) logic to handle the concept of partial values between "completely true" and "completely false". It was introduced by Lotfi Zadeh of UC Berkeley in the 1960's as a means to model the uncertainty of natural language.

[182] USPTO.ORG, *at* http://www.uspto.org (last visited June 26, 2005).

analysis) and patent number. When applicable, we give the assignee company name or the company name which is listed as correspondent (when it is not a law firm).

¶ 176 There are 16 applications in Bayesian analysis which we illustrate with two sample applications. Application 20020062245 describes a method for generating real-time promotions on an e-commerce website. It computes the Bayesian probability that a visitor will leave the website or make a purchase based upon entered data and decides whether or not real-time promotions tailored to the visitor's display preferences should be generated on the website. Such precise targeted marketing can help reduce spam. Application 20040083129 describes a distributed multi-agent system for real time collection, monitoring and analysis of network resources. It uses Bayesian analysis to rapidly identify abnormal conditions such as scam or spam attacks. Interestingly, this patent has only one claim and is short (12 pages) with low depth and disclosure.

¶ 177 There are 19 applications in Bayesian filtering which we illustrate with some sample applications. Application 20040139160 assigned to Microsoft describes a framework to integrate several anti-spam filters, including Bayesian filters, by comparing the confidence levels in the filtering results against predefined thresholds to decide whether or not to continue further filtering. Application 20040167964 describes an adaptive use of filters for anti-spam based on adjustable false positive rates and false negative rates which are initially assigned and subsequently adjusted as a result of filtering. It uses Bayesian analysis to estimate uncertainty of the rates. There are also 5 related applications from W.T. Daniell, with Bellsouth as correspondent, which use Bayesian filtering in different ways (phonetic filtering, filtering of message attachments, etc.).[183]

¶ 178 There are 5 applications in combined analysis which we illustrate with one sample application. Application 20040267893 describes a fuzzy logic voting method for classifying e-mail messages using inputs from multiple spam classifiers.

¶ 179 To summarize, even in this narrow set of Bayesian-related patents, we found that they are all very different, except those explicitly related from the same inventors.

[183] U.S. Patent Application No. 20050080642; U.S. Patent Application No. 20050080860 ; U.S. Patent Application No. 20050080864; U.S. Patent Application No. 20050091321; U.S. Patent Application No. 20050097174.

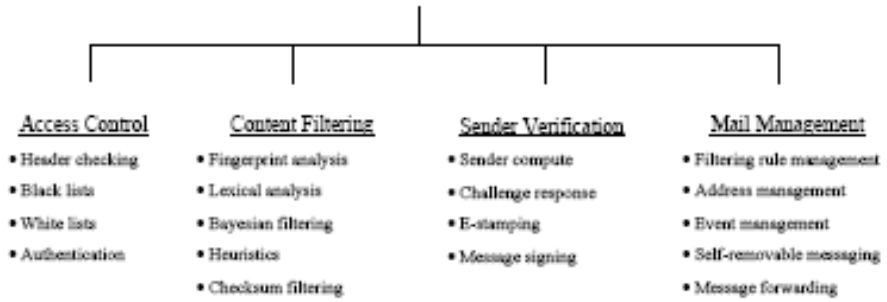

Figure 1 Taxonomy of Antispam Techniques

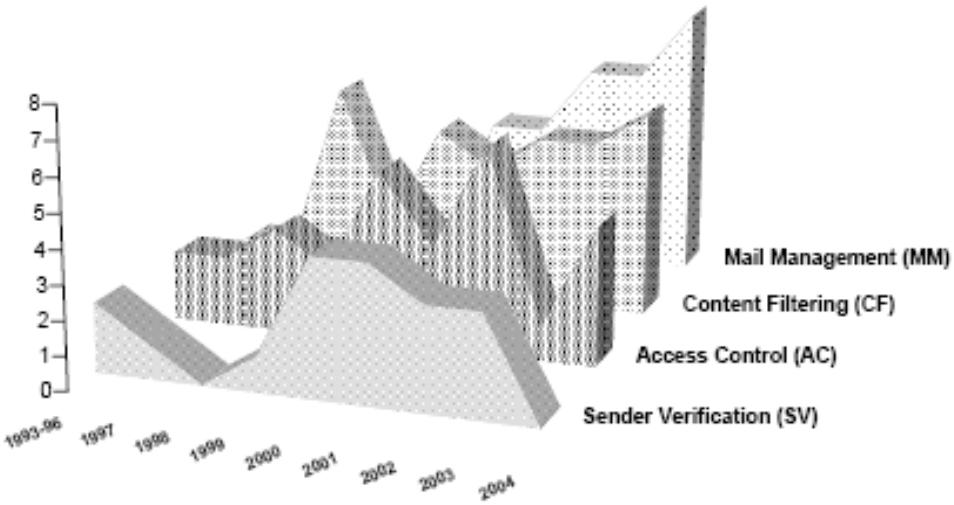

Figure 2 Number of Patents Issued by Antispam Technology Class

Table 1 Specialist Anti-Spam Vendors

Vendor	Date founded	Country	2003 antispam revenues ($m)	Antispam patents	Principal activities
Brightmail (acq'd by Symantec, 2005)	1998	US	20.8	Yes	Anti-spam software products
Tumbleweed	1993	US	15.7	Yes	Anti-spam software products; technology licensing
Cloudmark (open source)	2002	US	7.5	Pending	Anti-spam software products; technology licensing
Proofpoint	2002	US	4.5	-	Anti-spam software products
Coblon (acq'd by ISS, 2004)	1997	Germany	2.0	-	Anti-spam software products
Mail-Filters	2001	US	1.9	-	Technology licensing
MessageGate	2003	US	1.6	Note 1	Anti-spam software products
MailFrontier	2002	US	1.5	Yes	Anti-spam software products; technology licensing
Mailshell	1999	US	1.4	Pending	Technology licensing
WebWasher (acq'd by CyberGuard, 2004)	1999	Germany	1.1	Note 2	Anti-spam software products
Commtouch	1991	Israel	0.9	Yes	Technology licensing

Source (revenues only):
IDC, *Worldwide Antispam Solutions 2004-2008 Forecast and 2003 Vendor Shares*, December 2004

Notes
1. MessageGate is a spin-off from Boeing, which has an antispam patent, but it has not been assigned to MessageGate.
2. WebWasher is a spin-off from Siemens, which has an antispam patent, but it has not been assigned to WebWasher.

Table 2 Internet Security Incumbents Supplying Anti-Spam Software Products

Vendor	Date founded	Country	2003 antispam revenues ($m)	External sources of anti-spam technologies
SurfControl	1998	UK	16.5	Acq'd SecureM, 2004
Clearswift	1982	UK	16.3	Acq'd Content Technologies, 2002
Symantec	1982	US	14.2	Acq'd Brightmail, 2004; TurnTide, 2004
NetIQ	1995	US	11.0	Acq'd Marshal Software, 2002. Tumbleweed licensee
GROUP Technologies	1992	Germany	9.8	-
Network Associates	1989	US	9.2	Acq'd SpamKiller, 2002; Deersoft, 2003
Sophos	1985	UK	8.4	Acq'd ActiveState, 2004
Trend Micro	1988	Japan	8.1	Postini licensee
Sybari	1995	US	7.1	Commtouch licensee
Sendmail (open source)	1991	US	4.5	Cloudmark licensee
ZixCorp	1988	US	4.0	Acq'd Elron Software, 2003
Panda Software	1990	Spain	2.8	MailShell licensee
Aladdin Knowledge Systems	1985	Israel	2.2	Corbion licensee
F-Secure	1988	Finland	1.9	-
Computer Associates	1975	US	1.8	Qurb licensee
Norman ASA	1984	Norway	1.2	-
AhnLab	1995	Korea	1.1	Cloudmark licensee
Finjan	1996	Israel	1.1	Mailshell license

Source (revenues only):
IDC, *Worldwide Antispam Solutions 2004-2008 Forecast and 2003 Vendor Shares*, December 2004

Table 3 Antispam Appliance Vendors

Vendor	Date founded	Country	2003 antispam revenues ($m)	Sources of anti-spam technologies	Technology licenses for Internet security and encryption
CipherTrust	2000	US	20.8	Own technologies	Sophos, Network Associates, PostX
IronPort Systems	2000	US	8.0	Tumbleweed licensee; Brightmail licensee; acq'd SpamCop, 2003	Sophos, PGP
BorderWare	1994	Canada	4.0	Brightmail licensee	Kapersky, RSA
MiraPoint	1997	US	3.5	Own technologies; Commtouch licensee	Sophos
Barracuda Networks	2002	US	2.1	Open source	Open Source
Corvigo (acq'd by Tumbleweed, 2004)	2003	US	2.0	Note 1	Note 1
Nokia Internet Communications	2000	Finland	1.0	Own technologies; Postini licensee	SurfControl, CheckPoint
BlueCat Networks	2000	Canada	1.0	Open source; Commtouch licensee	F-Secure, Borderware

Source (revenues only):
IDC, *Worldwide Antispam Solutions 2004-2008 Forecast and 2003 Vendor Shares*, December 2004

Note:
1 No public domain information has been available on Corvigo since its acquisition by Tumbleweed.

Table 4 Hosted Services

Vendor	Date founded	Country	2003 antispam revenues ($m)	Sources of anti-spam technologies	Technology licenses for Internet security
Postini	1999	US	16.7	Own technologies	Authentium, Network Associates
Messagelabs	1999	UK	12.8	Own technologies, Brightmail licensee, Kelkea licensee	Symantec, Network Associates, F-Secure
FrontBridge	1999	US	7.5	Own technologies	Symantec, Kapersky, Sophos
MX Logic	2002	US	1.7	Brightmail licensee, Kelkea licensee	Authentium, Network Associates, Sophos
MailWise	2002	US	0.5	Not stated	Not stated

Source (revenues only):
IDC, *Worldwide Antispam Solutions 2004-2008 Forecast and 2003 Vendor Shares*, December 2004

Table A1: Anti-spam patents per assignee

This table gives for each assignee (in general a company name, or an individual - prefixed by the keyword "Ind:" – if there is no company) the numbers of patents granted for each class of technology (access control, content filtering, email management, sender verification). It is ordered by increasing order of assignee.

Assignee	Access control	Content filtering	Sender verification	Email management	Total
Abuzz Technologies		1			1
Action Technologies			2		2
AltaVista		2			2
Ameritech	1			1	2
Anti-spam Technology	2				2
AOL			2		2
Apple		1			1
AT&T	3		2	2	7
Beyond				1	1
Boeing		1			1
BrightMail	1			2	3
CBT Flint Partners				1	1
Cloudshield				1	1
CommTouch		1			1
Compuserve		1			1
Cranberry Properties		1			1
Duquesne Univ.		1			1
Escom	1				1
Fujitsu				1	1
HP		3	1		4
IBM	3	4	2	2	11
Ind.: Cohen		1			1
Ind.: Council et al.				1	1
Ind.: Ogilvie et al.				4	4
Ind.: Olivier	1				1
Ind.: Pace et al.		1			1
Ind.: Pang				2	2
Ind.: Reuning				1	1
Ind.: Stmdsted			1		1
Ind.: Wang	1				1

Assignee	Access control	Content filtering	Sender verification	Email management	Total
Infoseek			1		1
Infospace	2				2
InfoSpace				1	1
Intel	2		1		3
Jfax Communications				1	1
Lucent			2	1	3
Mailfrontier		1			1
Micron	1				1
Microsoft		6		1	7
Minolta		1			1
Mitel Knowledge			1		1
Motorola				1	1
Netcreations	1				1
Network Associates		1		2	3
Nixmail			1		1
Nokia		1			1
Omron		1			1
Paratran	1				1
PinPoint		3			3
Postini				1	1
Propel Software		1			1
SBC Properties	1				1
Secure Computing		1			1
Sendmail	1				1
Siemens				1	1
Sun	2				2
Surfcontrol	1				1
TeleCommunication Systems			1		1
The Robert Uomini and Louise Bidwell Trust			1		1
Titanker Software	1				1
Trend Micro				1	1
Tumbleweed		1			1
Univ. Central Florida		1			1
Vanguish	1				1
Xerox	3				3
Total numbers per class	**30**	**36**	**18**	**29**	**113**

Table A2: Patent descriptions

This table describes our set of anti-spam patents[1], ordered by technology (AC: access control, CF: content filtering, SV: sender verification, MM: email management) and patent#. We give the last assignee of the patent, and when applicable, the previous main assignee(s) in parenthesis.

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
5315504	IBM	1994	AC	28	12		Electronic document approval system
5513126	Xerox	1996	AC	164	45	foundation	Network having selectively accessible recipient prioritized communication channel profiles
5657461	Xerox	1997	AC	37	18	cont. of 5513126	User interface for defining and automatically transmitting data according to preferred communication channels
5689642	Xerox	1997	AC	68	22	cont. of 5513126	Recipient prioritized communication channel profiles
5742769	Infospace (Banyan)	1998	AC	31	10		Directory with options for access to and display of email addresses
5757891	Ind.: Wang	1998	AC	21	35		Ever ready telephonic answering-machine for receiving and delivering electronic messages
5826022	Sun	1998	AC	35	20		Method and apparatus for receiving electronic mail
5905777	AT&T	1999	AC	44	33		E-mail paging system
5930479	AT&T	1999	AC	38	68	broad scope	Communications addressing system
6023723	Anti-spam Technology	2000	AC	30	23		Method and system for filtering unwanted junk e-mail utilizing a plurality of filtering mechanisms
6052709	Brightmail (BrightLight)	2000	AC	45	7	narrow scope	Apparatus and method for controlling delivery of unsolicited electronic mail
6073167	Paratran	2000	AC	15	20		Distribution limiter for network messaging
6108691	Infospace (Switchboard)	2000	AC	9	20	cont. of 5742769 - Banyan	Directory with options for access to and display of email addresses
6167435	Netcreations	2000	AC	4	22		Double opt-in TM. method and system for verifying subscriptions to information distribution services
6219786	Surfcontrol	2001	AC	16	18		Method and system for monitoring and controlling network access
6249805	Micron	2001	AC	17	9		Method and system for filtering unauthorized electronic mail messages
6321267	Escom	2001	AC	8	68	broad scope	Method and apparatus for filtering junk email
6356935	Intel (Omnipoint)	2002	AC	3	31		Apparatus and method for an authenticated electronic userid

[1] Downloaded from www.uspto.org on May 5, 2005.

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
6400810	Ameritech	2002	AC	4	2	narrow scope	Method and system for selective notification of E-mail messages
6421709	Anti-spam Technology	2002	AC	4	2	narrow scope	E-mail filter and method thereof
6453327	Sun	2002	AC	7	36		Method and apparatus for identifying and discarding junk electronic mail
6473758	Intel	2002	AC	1	33		Method and apparatus for private and restricted-use electronic addresses
6480885	Ind.: Olivier	2002	AC	10	22		Dynamically matching users for group communications based on a threshold degree of matching of sender and recipient predetermined acceptance criteria
6535586	AT&T	2003	AC	3	7	narrow scope	System for the remote notification and retrieval of electronically stored messages
6691156	IBM	2004	AC	0	3		Method for restricting delivery of unsolicited E-mail
6697462	Vanguish	2004	AC	0	9		System and method for discouraging communications considered undesirable by recipients
6782079	SBC Properties	2004	AC	0	14	cont. of 6438215 - Ameritech	Method and system for filter based message processing in a unified messaging system
6829631	IBM	2004	AC	0	28		Method and system for screening electronic messages
6865671	Sendmail	2005	AC	0	15		Electronic mail system with authentication methodology for supporting relaying in a message transfer agent
6868498	Titanikey Software (Katsikas)	2005	AC	0	20		System for eliminating unauthorized electronic mail
5377354	HP (Digital)	1994	CF	121	16	foundation	Method and system for sorting and prioritizing electronic mail messages
5613108	Minolta	1997	CF	41	20		Electronic mail processing system and electronic mail processing method
5717913	Univ. Central Florida	1998	CF	29	11		Method for detecting and extracting text data using database schemas
5724567	Apple	1998	CF	113	39		System for directing relevance-ranked data objects to computer users
5754938	PinPoint (Herz et al.)	1998	CF	144	36	foundation	Pseudonymous server for system for customized electronic identification of desirable objects
5754939	PinPoint (Herz et al.)	1998	CF	219	22	cont. of 5754938	System for generation of user profiles for a system for customized electronic identification of desirable objects
5796948	Ind.: Cohen	1998	CF	19	12		Offensive message interceptor for computers
5835087	PinPoint (Herz et al.)	1998	CF	180	24	cont. of 5754938 and 939	System for generation of object profiles for a system for customized electronic identification of desirable objects
5999932	BrightMail (BrightLight)	1999	CF	56	31	foundation	System and method for filtering unsolicited electronic mail messages using data matching and heuristic processing
6003027	IBM	1999	CF	16	20		System and method for determining confidence levels for the results of a

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
							categorization system
6023700	Compuserve	2000	CF	41	12		Electronic mail distribution system for integrated electronic communication
6072942	Secure Computing	2000	CF	22	27		System and method of electronic mail filtering using interconnected nodes
6092101	HP (Digital)	2000	CF	20	3	narrow scope	Method for filtering mail messages for a plurality of client computers connected to a mail service system
6119124	Altavista (Digital)	2000	CF	11	24		Method for clustering closely resembling data objects
6161130	Microsoft	2000	CF	38	65	foundation, broad scope	Technique which utilizes a probabilistic classifier to detect ""junk"" e-mail by automatically updating a training and re-training the classifier based on the updated training set
6185551	HP (Digital)	2001	CF	8	12		Web-based electronic mail service apparatus and method using full text and label indexing
6192360	Microsoft	2001	CF	30	44		Methods and apparatus for classifying text and for building a text classifier
6199103	Omron	2001	CF	9	20		Electronic mail determination method and system and storage medium
6330590	CommTouch (Cotten)	2001	CF	9	31		Preventing delivery of unwanted bulk e-mail
6349296	AltaVista Company	2002	CF	1	20	cont. of 6119124 - Digital	Method for clustering closely resembling data objects
6370526	IBM	2002	CF	4	30		Self-adaptive method and system for providing a user-preferred ranking order of object sets
6397205	Duquesne Univ.	2002	CF	3	14		Document categorization and evaluation via cross-entropy
6415304	Microsoft	2002	CF	0	17		Waiting prior to engaging in action for enhancement of automated service
6460050	Ind.: Pace et al.	2002	CF	6	25		Distributed content identification system
6505167	Microsoft	2003	CF	0	32		Systems and methods for directing automated services for messaging and scheduling
6546390	Abuzz Technologies	2003	CF	1	45		Method and apparatus for evaluating relevancy of messages to users
6553358	Microsoft	2003	CF	0	46		Decision-theoretic approach to harnessing text classification for guiding automated action
6609196	Tumbleweed	2003	CF	2	19		E-mail firewall with stored key encryption/decryption
6633630	Cranberry Properties	2003	CF	3	89	broad scope	System for integrated electronic communications
6728690	Microsoft	2004	CF	0	27		Classification system trainer employing maximum margin back-propagation with probabilistic outputs
6732149	IBM	2004	CF	0	42		System and method for hindering undesired transmission or receipt of electronic messages

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
6732157	Network Associates	2004	CF	0	17		Comprehensive anti-spam system, method, and computer program product for filtering unwanted e-mail messages
6772196	Propel Software	2004	CF	0	3	narrow scope	Electronic mail filtering system and methods
6778834	Nokia	2004	CF	0	33		Push content filtering
6779021	IBM	2004	CF	0	33		Method and system for predicting and managing undesirable electronic mail
6845374	Mailfrontier	2005	CF	0	9		System and method for adaptive text recommendation
5208748	Action Technologies	1993	SV	50	19	Longest (171 pages)	Method and apparatus for structuring and managing human communications by explicitly defining the types of communications permitted between participants
5216603	Action Technologies	1993	SV	47	9	cont. of 5208748	Method and apparatus for structuring and managing human communications by explicitly defining the types of communications permitted between participants
5619648	Lucent (AT&T)	1997	SV	135	16	foundation	Message filtering techniques
5999967	Ind.: Sundsted	1999	SV	18	12		Electronic mail filtering by electronic stamp
6018761	The Robert Uomini and Louise Bidwell Trust	2000	SV	11	3	narrow scope	System for adding to electronic mail messages information obtained from sources external to the electronic mail transport process
6064878	AT&T	2000	SV	3	22		Method for separately permissioned communication
6085321	Intel (Omnipoint)	2000	SV	10	26		Unique digital signature
6112227	AOL (Heiner, Mailblocks)	2000	SV	7	18		Filter-in method for reducing junk e-mail
6195698	HP (Digital)	2001	SV	3	75		Method for selectively restricting access to computer systems
6199102	AOL (Cobb, Mailblocks)	2001	SV	12	37		Method and system for filtering electronic messages
6266692	IBM	2001	SV	14	16		Method for blocking all unwanted e-mail (SPAM) using a header-based password
6332164	AT&T	2001	SV	6	22		System for recipient control of E-mail message by sending complete version of message only with confirmation from recipient to receive message
6363140	Mitel	2002	SV	1	34		Dialable screening profile
6393465	Nixmail	2002	SV	9	20		Junk electronic mail detector and eliminator
6484197	IBM	2002	SV	1	16		Filtering incoming e-mail
6546416	Infoseek	2003	SV	2	24		Method and system for selectively blocking delivery of bulk electronic mail
6574658	Lucent	2003	SV	1	28		System and method for secure classification of electronic mail

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
6658260	TeleCom-munication Systems	2003	SV	0	14		Inter-carrier short messaging service providing phone number only experience
5283856	Beyond	1994	MM	140	19	foundation	Event-driven rule-based messaging system
6073165	Jfax Communica-tions	2000	MM	28	28		Filtering computer network messages directed to a user's e-mail box based on user defined filters, and forwarding a filtered message to the user's receiver
6094681	Siemens	2000	MM	25	11		Apparatus and method for automated event notification
6101531	Motorola	2000	MM	33	11		System for communicating user-selected criteria filter prepared at wireless client to communication server for filtering data transferred from host to said wireless client
6167434	Ind.: Pang	2000	MM	12	20		Computer code for removing junk e-mail messages
6185603	AT&T	2001	MM	24	37		Method and system for delivery of e-mail and alerting messages
6192114	CBT Flint Partners	2001	MM	2	14	Shortest (6 pages)	Method and apparatus for billing a fee to a party initiating an electronic mail communication when the party is not on an authorization list associated with the party to whom the communication is directed
6230188	InfoSpace	2001	MM	6	40		System and method for providing a proxy identifier in an on-line directory
6324569	Ind.: Ogilvie et al.	2001	MM	13	18		Self-removing email verified or designated as such by a message distributor for the convenience of a recipient
6381592	Ind.: Reuning	2002	MM	5	15		Candidate chaser
6438215	Ameritech	2002	MM	5	11		Method and system for filter based message processing in a unified messaging system
6442589	Fujitsu	2002	MM	7	22		Method and system for sorting and forwarding electronic messages and other data
6487586	Ind.: Ogilvie et al.	2002	MM	1	21	Division of 6324569	Self-removing email verified or designated as such by a message distributor for the convenience of a recipient
6490574	IBM	2002	MM	1	26		Method and system for managing rules and events in a multi-user intelligent agent environment
6493007	Ind.: Pang	2002	MM	1	27		Method and device for removing junk e-mail messages
6560632	IBM	2003	MM	0	33		System and method for managing files in a distributed system using prioritization
6587550	Ind.: Council et al.	2003	MM	0	14		Method and apparatus for enabling a fee to be charged to a party initiating an electronic mail communication when the party is not on an authorization list associated with the party to whom the communication is directed

Patent#	Assignee	Yr	Tech.	#Cites	#Claims	Comment	Title
6591291	Lucent	2003	MM	4	32		System and method for providing anonymous remailing and filtering of electronic mail
6643686	AT&T	2003	MM	0	10		System and method for counteracting message filtering
6650890	Postini	2003	MM	2	36		Value-added electronic messaging services and transparent implementation thereof using intermediate server
6654787	BrightMail (BrightLight)	2003	MM	8	60	broad scope	Method and apparatus for filtering e-mail
6701347	Ind.: Ogilvie	2004	MM	0	22		Method for including a self-removing code in a self-removing email message that contains an advertisement
6745193	Microsoft	2004	MM	0	26		System and method for defining, refining, and personalizing communications policies in a notification platform
6754230	Boeing	2004	MM	0	21		User bandwidth monitor and control management system and method
6757713	Ind.: Ogilvie et al.	2004	MM	0	25		Method for including a self-removing indicator in a self-removing message
6757830	Network Associates	2004	MM	1	45	broad scope	Detecting unwanted properties in received email messages
6802012	Network Associates	2004	MM	0	36		Scanning computer files for unwanted properties
6829654	Cloudshield	2004	MM	0	30		Apparatus and method for virtual edge placement of web sites
6836792	Trend Micro	2004	MM	0	14		Techniques for providing add-on services for an email system

Table B.1: Bayesian-related patent applications

This table describes our set of 40 Bayesian-related patent applications ordered by technology (BA: Bayesian analysis, BF: Bayesian content filtering, CA: combined analysis) and application number. Where applicable, we give the assignee company name or the company name which is listed as correspondent (when it is not a law firm). [Source: Downloaded from www.uspto.org on June 26, 2005.]

App. #	Inventor	Assignee	Yr	Tech.	Title
20020055940	Elkan, Charles		2002	BA	Method and system for selecting documents by measuring document quality
20020059425	Belfiore, Joseph ; et al.	Microsoft	2002	BA	Distributed computing services platform
20020062245	Niu, David ; et al.		2002	BA	System and method for generating real-time promotions on an electronic commerce world wide website to increase the likelihood of purchase
20030182310	Charnock, Elizabeth ; et al.		2003	BA	Method and apparatus for sociological data mining
20040083129	Herz, Frederick S. M.		2004	BA	SDI-SCAM
20040177110	Rounthwaite, Robert L. ; et al.		2004	BA	Feedback loop for spam prevention
20040221062	Starbuck, Bryan T. ; et al.		2004	BA	Message rendering for identification of content features
20050010472	Quatse, Jesse T. ; et al.		2005	BA	High-precision customer-based targeting by individual usage statistics
20050030589	El-Gazzar, Amin ; et al.		2005	BA	Spam fax filter
20050050150	Dinkin, Sam		2005	BA	Filter, system and method for filtering an electronic mail message
20050055399	Savchuk, Gene		2005	BA	High-performance network content analysis platform
20050055416	Heikes, Brian Dean ; et al.		2005	BA	Managing instant messages
20050060295	Gould, Stephen ; et al.	Sensory Networks	2005	BA	Statistical classification of high-speed network data through content inspection
20050071432	Royston, Clifton W. III		2005	BA	Probabilistic email intrusion identification methods and systems
20050097435	Prakash, Vipul Ved ; et al.		2005	BA	Methods and apparatuses for classifying electronic documents
20050102366	Kirsch, Steven T.		2005	BA	E-mail filter employing adaptive ruleset
20030204569	Andrews, Michael R. ; et al.	Lucent (corresp.)	2003	BF	Method and apparatus for filtering e-mail infected with a previously unidentified computer virus
20040139160	Wallace, Andrew J. ; et al.	Microsoft	2004	BF	Framework to enable integration of anti-spam technologies
20040139165	McMillan, Bruce A. ; et al.	Microsoft	2004	BF	Framework to enable integration of anti-spam technologies
20040148330	Alspector, Joshua ; et al.		2004	BF	Group based spam classification
20040167964	Rounthwaite, Robert L. ; et al.		2004	BF	Adaptive junk message filtering system
20040215977	Goodman, Joshua T. ; et al.		2004	BF	Intelligent quarantining for spam prevention
20050015452	Corson, Gregory	Sony	2005	BF	Methods and systems for training content filters and resolving uncertainty in content filtering operations

App. #	Inventor	Assignee	Yr	Tech.	Title
20050015626	Chasin, C. Scott		2005	BF	System and method for identifying and filtering junk e-mail messages or spam based on URL content
20050076240	Appleman, Barry		2005	BF	Degrees of separation for handling communications
20050076241	Appelman, Barry		2005	BF	Degrees of separation for handling communications
20050080642	Daniell, W. Todd	Bellsouth (corresp.)	2005	BF	Consolidated email filtering user interface
20050080860	Daniell, W. Todd ; et al.	Bellsouth (corresp.)	2005	BF	Phonetic filtering of undesired email messages
20050080864	Daniell, W. Todd	Bellsouth (corresp.)	2005	BF	Processing rules for digital messages
20050081059	Bandini, Jean-Christophe Denis ; et al.		2005	BF	Method and system for e-mail filtering
20050091321	Daniell, W. Todd ; et al.	Bellsouth (corresp.)	2005	BF	Identifying undesired email messages having attachments
20050097174	Daniell, W. Todd	Bellsouth (corresp.)	2005	BF	Filtered email differentiation
20050097179	Orme, Gregory Michael		2005	BF	Spam prevention
20050120019	Rigoutsos, Isidore ; et al.	IBM	2005	BF	Method and apparatus for the automatic identification of unsolicited e-mail messages (SPAM)
20050131811	Ranzini, Stephen Lange ; et al.		2005	BF	System and method for message handling
20030158830	Kowalczyk, Adam ; et al.		2003	CA	Gradient based training method for a support vector machine
20040177081	Dresden, Scott		2004	CA	Neural-based internet search engine with fuzzy and learning processes implemented at multiple levels
20040267893	Lin, Wei		2004	CA	Fuzzy logic voting method and system for classifying E-mail using inputs from multiple spam classifiers
20050004905	Dresden, Scott		2005	CA	Search engine with neural network weighting based on parametric user data
20050055340	Dresden, Scott	Brainbow	2005	CA	Neural-based internet search engine with fuzzy and learning processes implemented by backward propagation