

COMMERCIAL AND NON-COMMERCIAL APPROACHES TO FIGHTING SPAM

Oren Drori

CommTouch Software, Ltd., 1A Hatzoran St,
P.O. Box 8511, Netanya 42504, Israel

Tel +972 9 8636837 • Email
orend@commtouch.com

ABSTRACT

With the exponential rise in spam over the past two years, a number of commercial and non-commercial approaches to fighting spam have been proposed. This paper provides an overview of trends in the spam industry, and evaluates the viability of commercial and non-commercial anti-spam techniques.

Perhaps the most significant trend in spam is the profile of the spammers that stand behind most of the junk mail that floods our inboxes. Whereas two or three years ago spam was mostly used by quasi-legal operations, today criminals and organized crime are responsible for most spam. Spam-related e-crime activities include: the use of viruses as a distribution platform, penetration and hijacking of computer stations, directory attacks and email fraud.

In this context, we will discuss the major non-commercial efforts at containing the spam problem, including: legislation, Internet standardisation and non-profit organizations. We will explain why these efforts have been largely ineffective, and sometimes even counterproductive.

Finally, we will review several commercial approaches, which effectively handle spam and block it almost completely. The expected long-term effectiveness of these solutions will be discussed. In conclusion, we will show that the overall cost of commercial solutions is far below that of either accepting spam or implementing non-commercial alternatives.

INTRODUCTION

Techniques to block spam are almost as old as spam itself. Since the first unsolicited mass emailing in 1994, multiple approaches to blocking spam have been used, with varied degrees of success. Today, organizations and individuals have a wide range of choices, both commercial and non-commercial.

Non-commercial attempts include legislation and Internet standardization. Legislating against spam has proven to be complex and problematic, while Internet standards offer only a very partial remedy to the problem.

On the other hand, commercial approaches are evolving constantly, responding to spammers' innovations. During the last three years there has been a marked improvement in the performance of the top commercial approaches. Industry analysts believe that future methods will be based on network-wide spam detection rather than on analysis of individual messages.

This paper describes commercial and non-commercial approaches to fighting spam. It evaluates the pros and cons of each method, explaining why, so far, commercial approaches have far outperformed non-commercial ones.

HOW SPAMMERS WORK

The criminalization of spam

'Anti-virus experts have detected signs of a massive, well-coordinated Trojan attack capable of creating botnets-for-hire. Is it the work of organized crime?'

Ryan Naraine, *Tech News*, June 2005.

In the 'good old days', say the years 1996–2001, spam was not altogether different from direct mail advertising. The key difference was the 'e' in email, but the spammers themselves and the contents of their messages were relatively similar. This is changing rapidly. Spam is increasingly becoming the province of criminals.

A few years ago, a typical spammer might have been a flower delivery business or a dating service. Today, such mainstream businesses are the minority. For a few years, spam has been the province of extremely aggressive quasi-legal marketers, specializing in pharmaceuticals, pornography and gambling. Worse yet, these lightweights are making way for or becoming hard-core criminals who traffic in illegal substances or fraud. Financial fraud accounted for over 10% of spam in the first half of 2005. The most popular types are phishing, stock pump-and-dump and Nigerian-sting schemes.

Who Are The Spammers?

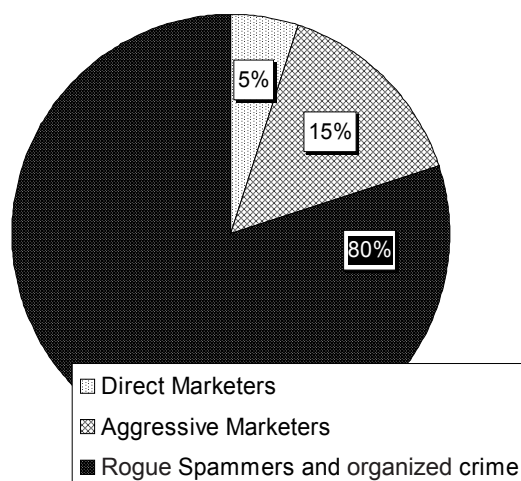


Figure 1: Spammer types.

Spamming technologies

- The same spammers who send criminal content also use criminal distribution methods.
- 60–80% of global spam is sent using viruses.

'The notion of purchasing the use of botnets, or zombie grids, is well-known. There's a sharp uptake in the amount of spam being generated by these zombies. It's pretty well-organized.'
Andrew Jaquith, security analyst, *Yankee Group*.

Spammers are constantly devising new tactics to make their email messages appear innocent, and allow them to pass

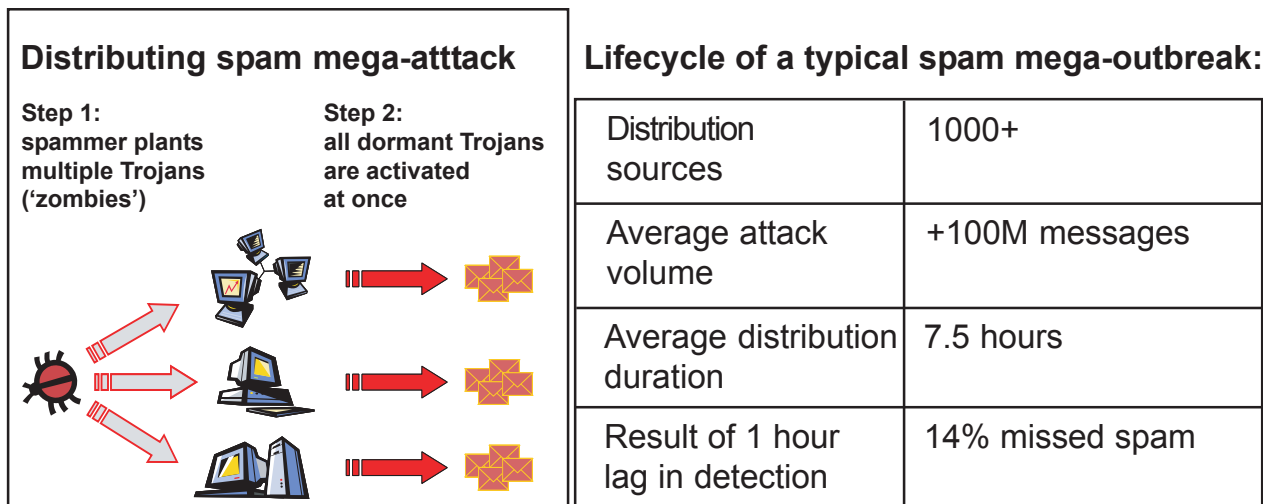


Figure 2: Mega-attack technique and results.

through spam detection applications. Their increasingly sophisticated methods for delivering spam include: sending the entire spam message as an image; slightly altering the content, subject line, or 'From' field; avoiding the use of known spam 'keywords'; disguising URLs; using quotations from innocent sources (e.g. poetry or news articles) and more.

Given that the majority of today's spam is distributed by professional criminals, it is hardly surprising that they are eager to use the most aggressive distribution techniques available – be they legal or illegal. Trojans are a particularly widespread spam distribution technique: dormant Trojans are planted in advance in hundreds of susceptible computers over the web ('Zombies', aka 'Botnets'). On 'D-day' they are all activated together to create fast and massive mega-attacks, distributing enormous volumes of spam over short periods. Of the 25 billion spam message sent each day, 60–80% are sent in such mega-attacks using Trojans and viruses.

Spam-virus – a dangerous symbiosis

The criminalization of spam has created a deeper affinity between spam and viruses. Just as Trojans are used to distribute spam, the reverse is also true. Many of the successful virus attacks in 2004–2005 used spam distribution technology.

In this universe of hybrid threats, speed of distribution has become the key success parameter for both spammers and virus and malware creators. This weapon is aimed directly at the weak spot of many security solutions – the time lag involved in adapting to the new spam or spam-virus attack.

NON-COMMERCIAL APPROACHES TO FIGHTING SPAM

The challenge

Non-commercial approaches to fighting spam range from new legislation to various Internet standardization methods. By harnessing the public, governments and enterprises – it sometimes seems that spam can be blocked. After all, everyone wants to get rid of it.

The global applicability and cost benefits of using non-commercial methods clearly explain the widespread

interest in them. Nevertheless, we should be cautious and ask: can these approaches deliver? Can they protect us effectively against spam, for free and over time? Unfortunately, in this case as in many others, there are no free rides.

Public blacklists

Among the earliest techniques for fighting spam and spammers were public blacklists (RBLs or Real-time Blackhole Lists), which identify the originating sources of spam (IPs) by tracing users' complaints.

This method served its original purpose very well: it was a powerful antidote to the bad habit of using open proxy servers. Indeed, most ISPs will automatically block a blacklisted IP, and open proxies can hardly be found any more.

Obviously, one thing that public blacklists did not achieve was to solve the spam problem. Spammers have simply stopped spamming from static sources. Almost all spam is sent from temporary sources (IPs), which cannot be blacklisted in time.

Internet standardization

Sender identification

A number of new standards have been proposed to make the SMTP email protocol more secure and reliable. Specifically, there has been a focus on making sender-forging more difficult. These efforts include standards such as SPF/Caller-ID initiatives and their recent incarnation, Sender-ID.

The effectiveness of any standard is inherently dependent upon its adoption – a critical mass of users is required to make a new standard effective. Even if adoption is not immediate, it can still be effective if it is backed by key players such as *Microsoft* and *AOL*.

Will sender-forging prevention (via improved Internet standards) stop spam? Most experts agree that such an assumption is far too optimistic. Even the authors of these standardization initiatives aim for more modest goals. This family of standards will ensure that an email sent from john@doe.com was actually sent from that domain. They will make the Internet a more trustworthy environment, and that is a great deal.

A successful standard might make it more difficult for spammers to cover their tracks – meaning increased costs, decreased incentives and a step in the right direction. Yet expecting this type of intervention to wipe out the spam problem is probably somewhat naïve.

The ‘Penny Black’ initiative

At the beginning of 2004 Bill Gates suggested an initiative in a different direction. We can make spamming unprofitable, suggested Gates, by levying a 1-cent tax for each message sent (through ISP billing or CPU overhead). The idea relies on the fact that, for the average user this would cost just nickels and dimes, but for spammers, operations would be impractically expensive.

Why did ‘Penny Black’ not take off? Sometimes, the cure is worse than the disease: considering that 25 billion messages are sent on an average day, implementing ‘Penny Black’ would have cost the industry around \$100 billion, annually. Assuming that half of it is spam, we’re looking at a realistic cost of \$50 billion a year – far more expensive than providing top-level commercial solutions to all.

Legislation

Challenges for legislators

Almost everyone (except spammers) agrees that a world without spam would be a much better world. On this issue the interests of individuals, enterprises and governments are aligned. Given this congruence of influence, it can seem quite surprising that spam has not been eradicated effectively.

In fact, anti-spam legislation is far from simple. The first challenge is identification. It is not always easy to distinguish between spam and legitimate use of the Internet for direct marketing. Spreading the anti-spam legislation net too wide means threatening a legitimate industry.

Another challenge is prosecution. The largest and most significant spamming networks are international. Even if an outbreak is identified, this does not mean that it will be easy or even possible to prosecute spammers in other countries. There are numerous safe harbours for spammers. Some analysts argue that anti-spam laws will only be effective if and when they are implemented across the globe.

Attempts and results

‘A CSO (Chief Security Officer) cannot wait for government regulations to take effect. Add anti-spam products or services to your messaging architecture.’

Eric Ogren, analyst, *Yankee Group*.

In fact, numerous attempts have been made to leverage this global consensus to eliminate spam. To date, nearly 100 anti-spam legislation attempts have been made in the western world (see Figure 3 for a geographical breakdown), and not a single one can be called successful. So far, not a single anti-spam law has affected the bottom line, and our inboxes are as full of spam as ever.

It is not impossible that the combined effect of legislation and enforcement will bring better results in the future. Yet it is already clear that it will be years before anti-spam laws make a significant impact. It is worth noting that legislation can be risky; in some cases it can even become counterproductive. For example, the much-discussed CAN-SPAM law has had

Number of anti-spam laws passed to date*	
European Union	10
Austria	3
Belgium	1
Czech Republic	3
Denmark	4
Finland	4
France	2
Germany	3
Greece	1
Ireland	1
Italy	3
Luxemburg	1
Netherlands	1
Norway	3
Portugal	1
Spain	1
Sweden	3
UK	3
Total European anti-spam laws	48
US Federal laws	4 proposals 1 enacted
US State laws	over 40
Total legislation attempts:	over 92

*References, including the full text of most laws: <http://www.spamlaws.com/>

Figure 3: Worldwide anti-spam legislation attempts.

some detrimental effects: some spammers have complied to the technical legalistic requirements, while their spamming business continues as usual. For them, the law provided a cloak of legitimacy. On the other hand, some spammers who have been identified and prosecuted have simply renounced the pretence of authenticity, adopting more aggressive distribution techniques and criminal content.

Arresting Mafiosos for jaywalking

The main difficulty facing these methods is simple: the most serious, professional spammers are already living outside the law – they are involved in much more severe criminal activity (spreading computer viruses, information theft, international fraud, or trading in illegal substances).

More than likely, legislators will conceive of creative laws against spam, some of which may be more effective than CAN-SPAM and other existing laws. Yet in essence, all legislation aims at the ‘soft’ spammers, the low-hanging fruit. Generally speaking, legislative efforts do not even presume to influence the behaviour of hard-core criminals who are already using viruses to distribute illegal messages.

Sadly enough, most (nearly 80% of) spam traffic today is attributed to the hard-core type. At the end of the day, anti-spam laws may be as effective at blocking spammers as speed limits are at stopping bank robbers.

COMMERCIAL APPROACHES TO FIGHTING SPAM

Traditional methods, first generations

'Bayesian filtering has proved to be unreliable for spam filtering in the enterprise environment.'

Arabella Hallawell, Research Director, *Gartner*, 2004.

Commercial approaches to fighting spam have developed rapidly, going through three or four cycles of evolution in the past four years.

One of the first approaches to fighting spam was the blacklist approach. By pooling data on spam sources from users' reports of IP misuse, such services compile blacklists of suspect sources (IPs). However, blacklists rapidly became outdated when spammers moved to temporary distribution sources, jumping from source to source far ahead of blacklists' update time.

Keywords filters, which identify spam according to its textual content, were developed in parallel – employing the content filters concept to spam-fighting. Alas, spammers quickly learned how to foil content filters. Using elaborate methods to constantly and automatically alter the content of their messages, today's spammers sidestep most content-based filters.

Heuristic filters (including the well-known Bayesian filters) are far more sophisticated than naïve content filters. They use various techniques to estimate the probability that a given message is spam. Since heuristic filters do not positively identify a message before blocking it, high false positive ratios can be a problem.

Each of these methods had its 'day in the sun' when it was able to block high ratios of spam. Yet at the end of the day, they all failed in delivering robustness over time, in terms of both effectiveness and accuracy (see Figure 4).

Spammers have a strong financial motivation, and they have learned to overcome blacklists, content-based filters and heuristic approaches. Again and again spammers have demonstrated their flexibility, adapting to each new anti-spam method by rapidly disguising the characteristic it uses to identify spam.

Advanced anti-spam techniques take into consideration the fact that spammers will make tremendous efforts to circumvent them. Rather than focusing on spam elements that can be modified relatively easily (look and feel, source, content), advanced technologies analyse factors that are

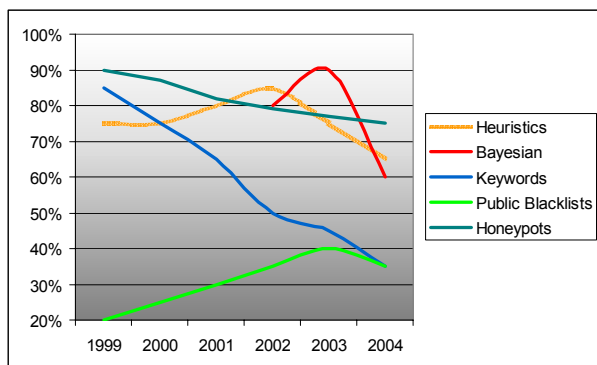


Figure 4: Long-term effectiveness of anti-spam methods.

fundamental to spam such as distribution patterns, distribution volume and speed, and the URLs of spammers' websites.

The challenge: response time

- As spam becomes increasingly polymorphous and dynamic, it becomes more difficult to identify it by analysing individual messages.

To maintain a high level of robustness over time, any anti-spam technology must respond rapidly to the dynamic spam environment.

Given the low, and rapidly decreasing effectiveness of traditional anti-spam methods, industry analysts have delineated a roadmap to effectively combat the problem of spam. According to analysts, the anti-spam methods of the future will not categorize spam by analysing individual messages. Rather, they will be based on network-wide vigilance and control.

The more polymorphous and dynamic spam becomes, the more difficult it is to identify and block it by analysing individual messages. Analysts, anti-spam vendors and service providers are looking for methods that:

- Have a bird's-eye view of Internet messaging: any method based on analysis of individual messages risks high false positive levels and a short shelf-life.
- Identify new outbreaks within minutes or seconds: in the world of mega-attacks, distributing 100 and even 200 million messages in a few hours, adequate response time is measured on a scale of minutes, not hours.
- Identify spam according to its fundamental characteristics: spammers constantly adapt to changing definitions of spam, so effective methods must focus on characteristics that spammers cannot change.
- Maintain extremely high accuracy levels (no false positives).

Network-based anti-spam approaches

'Although there is a greater volume of spam being sent, the technology used by anti-spam companies is now able to confidently identify unwanted emails.'

Megan Dahlgren, Senior Analyst for Software, *IDC*, June 2005.

Reaching high levels of effectiveness (spam capture rate), along with accuracy (low false positive rate), and maintaining it for years is far from easy. Not easy, and yet feasible.

Network-based solutions prove significantly more successful than standalone enterprise solutions that lack a network-based perspective for constant updating. Effective network-based solutions include managed services (e.g. *Postini*, *MessageLabs*), as well as hybrid models combining network-based detection centres with enterprise-side gateways (e.g. *Brightmail-Symantec*, and *CommTouch's Recurrent Pattern Detection* technology).

- Managed service detection centres: these detection centres process high volumes of messages. By analysing this traffic, they can identify emerging trends (e.g. the same message being sent repeatedly, or to various enterprises) and tag suspect messages.
- Recurrent Pattern Detection (RPD): this approach relies on the fact that all spam is sent in massive

outbreaks over relatively short periods of time. By analysing global email traffic in real time, RPD identifies new outbreaks as soon as they emerge. RPD's instant blocking is the key to its high detection rates, because messages are blocked in the first few minutes of the outbreak, before reaching users' mailboxes.

commercial solutions for near-absolute protection on the enterprise level.

In an independent survey of 23 leading anti-spam products conducted by *Osterman Research* in 2004, several commercial vendors achieved spam detection rates of 90 and above – see Figure 5.

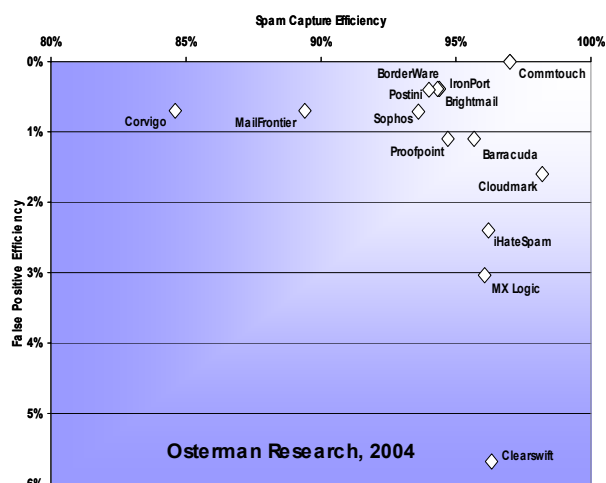


Figure 5: Anti-spam solution performance: capture rate vs. false positive rate.

CONCLUSION

Spammers are tenacious. As new anti-spam methods are implemented, spammers adapt rapidly, changing the content, format or distribution method of their spam.

During the first few years in the development of the anti-spam industry, each new method was effective for only a few months. Since then, the industry has become more sophisticated, forecasting which approaches will be most difficult to circumvent.

So far, the most effective techniques to detect and block spam have been commercial methods. Today, the top 10 commercial solutions on the market have an average detection rate of 92–97%. This is a far rosier picture than that of 2002, when the top five solutions covered barely 80–85% of spam. Standards of tolerance toward false positives have changed significantly as well: whereas 3–5% was once considered good enough, now users will not accept even 0.5–1% false positive rates.

While non-commercial approaches (mainly Internet standardization and legislation) serve important functions in and of themselves, they are simply not enough to solve the problem of spam. We should continue using them to make the Internet a less spam-friendly environment, and a more reliable environment for legitimate users. Doubtless these are important endeavours.

In parallel, just as with viruses and other malware threats, IT managers will continue relying only on themselves. For the foreseeable future, they will continue to deploy and optimize