

## Pattern-based Messaging Security for Hosting Providers

Email hosting is a key offering for Web hosting providers, and many deployments seem straightforward at first glance. Yet any email hosting offering must employ email filtering tools, or users will be overrun with spam, malware and phishing messages, and the provider's network will be overburdened with unwanted email, leaving fewer resources for handling legitimate email.

There are many low-cost and open-source tools to assist with filtering email, yet drawbacks of these solutions include:

- Unexpected upgrade or add-on costs or additional costs to ensure compatibility with existing infrastructure
- Nearly constant tuning by the provider to deal with new outbreaks
- Over-burdening of network resources, often requiring multiple servers to deal with unwanted mail
- Poor detection levels with a higher-than-acceptable level of false positives.

As email-borne threats continue to grow in magnitude and sophistication, effective solutions must deliver high detection rates with minimal mistakes. Hosting providers need to protect their end-customers from spam, phishing and malware, yet still provision their networks in a cost-efficient manner.

This White Paper outlines current message threats and introduces a messaging security suite based on unique, patented pattern-recognition technology deployed in the cloud. These offerings – Commtouch Anti-Spam, Zero-Hour™ Virus Outbreak Protection and GlobalView™ Mail Reputation – are available in easy-to-deploy integration packages designed specifically for Web hosting providers, enabling the provision of highly accurate, efficient and cost-effective email filtering. With more than 100 technology licensees around the globe, Commtouch provides industry-leading messaging security technology to many medium- and large-sized hosting providers globally, through a variety of integration tools for Plesk, Qmail, SpamAssassin, Milter and others.

Commtouch RPD integrates easily with

- Plesk
- Qmail
- SpamAssassin
- Milter
- Multiple other platforms

### Challenges in Email filtering

Filtering of unwanted email provides various challenges in the realms of security and network efficiency.

### Security Challenges

There are three main types of unwanted email that need to be filtered out by email service providers: spam, phishing, and malware messages. Challenges of each include:



- **Spam:** When composing spam messages, spammers use sophisticated tactics to evade existing spam detection applications. This includes masking the originator or sending machine of the spammers, manipulating or hiding commercial URLs, use of non-English words and phrases and a host of other methods. Typically a massive spam outbreak will only last a few hours and be launched from a network of “zombie” or “bot” machines. To complicate the detection process, each message within the massive spam outbreak can be composed differently and employ multiple evasion techniques. Spammers often unleash new techniques specifically designed to bypass anti-spam methods and take advantage of the window of opportunity to send messages in the new format until anti-spam technologies catch up. Examples of these include spam sent as images, PDF files or other types of attachments.
- **Phishing or password harvesting messages,** are sent for the single purpose of identity theft. There is an underlying intent within each message to violate the privacy of the recipient and commit fraud. The goal of a phishing message is to fool users to navigate to a Web site which will prompt them to enter private information. By using highly effective social engineering methods, these messages target users, often with a sense of urgency, to believe they have arrived at a legitimate site such as their bank or a recognized online vendor to enter their private information. The sender is then able to gather personal information such as credit card numbers, passwords, social security or identification numbers. Phishing messages appear to be from genuine or credible sources, and like spam, they messages can be sent in any language or format in attacks that typically last only a few hours and are usually launched from an army of zombie machines on the Internet.
- **Email-borne malware outbreaks** are created and released for malicious purposes. Like spam and phishing messages, each virus message can be different in terms of its content and the characteristics of the associated or attached executable files that contain the virus. However, email-borne viruses and worms, in particular, can be received from legitimate and trusted email sources that might have been infected and are unintentionally distributing the virus to others. Also like spam and phishing, email-borne virus attacks often last for very short time periods. In the case of viruses, users are exposed and unprotected during the first hours of the attack because most anti-virus defenses depend heavily on the use of a database of signatures that identify the threat by matching it with already-known characteristics. Recently, virus writers have become even more sophisticated by distributing multiple instances of the same virus within the same outbreak to evade heuristic systems and to maximize the impact before new signatures are propagated.

## Resource Management Challenges

Filtering the various types of malicious and unwanted messages provides unique challenges to the network administrator:

- **Even filtered threats still penetrate providers’ networks.** Because most email defense solutions do not protect at the perimeter, many email-borne security threats eventually enter the provider’s network, risking internal damage such as infection or the creation of backdoors for illicitly accessing information.
- **Storage:** Quantities of unwanted email are vast and growing. Because spammers and malware distributors can activate an unlimited number of zombies to send their messages, they can afford to send massive quantities of email at a negligible cost. Service providers bear the brunt of this in their own networks, being forced to waste vast amounts of storage space on this virtual garbage.



- **Network efficiency:** Besides storage, other resources such as bandwidth and server processing capacity are wasted on inefficient filtering technologies.
- **Network administrator resources** are unnecessarily consumed on managing numerous email filtering servers and technologies, and having to keep them constantly up-to-date with new patches and rules.

## Message Patterns

Massive outbreaks which distribute spam, phishing, and email-borne viruses or worms, consist of many millions of messages intentionally composed differently in order to evade commonly-used filters. Nonetheless, all messages within the same outbreak share at least one and often more than one unique, identifiable value which can be used to distinguish the outbreak.

For example, in the case of spam, the objective is to lead the recipient to groups of commercial Web sites to entice them to buy a product. In doing so, different spam attacks are often launched from identifiable zombie machines that can be blocked on the basis of their origin. In the case of phishing, recipients are lured to voluntarily disclose personal and confidential information via clever social engineering methods and the objective is often to lead the victims to the same fake URLs. Email-borne viruses always contain the same malicious code (otherwise it is a different virus or another instance of the same virus). These are all recurring values of typical outbreaks, called the outbreak 'message patterns.' Any message containing one or more of these unique patterns can be assumed with a great deal of certainty to be part of the same outbreak.

Message patterns can be extracted from the message envelope, headers, and body with no connection to the meaning of the message content. Thus, pattern analysis can be used to identify outbreaks in any language, message format, and encoding type. Message patterns can be divided into *distribution patterns*, which determine if the message is 'good' or 'bad' by analyzing the way it is distributed to the recipients, and *structure patterns*, which determine the volume of the distribution.

The challenges of message pattern classification include determining which message patterns identify outbreaks without generating cases of false positives, and determining how to extract and analyze these patterns before the outbreak wanes. Most outbreaks have a relatively short lifecycle measured in only a few hours. Therefore, any solution that does not detect and classify messages in real-time will only be effective towards the end of the outbreak, when most of the damage has already been done. All malicious outbreaks attempt to disguise messages as legitimate email correspondence pretending to arrive from trusted sources, and therefore solutions that are based on pattern analysis must be able to differentiate between 'good' and 'bad' patterns and avoid making mistakes.

The challenges are made more complex by the fact that each new outbreak usually introduces completely new patterns that were not previously analyzed and are therefore unknown to the pattern analyzer. Because tactics for distributing spam, phishing, and email-borne viruses and worms are constantly evolving, it is necessary to proactively identify new and unique patterns in real-time in order to determine new outbreaks as they are released to the Internet and begin targeting recipients. Pattern detection represents a highly accurate method for identifying email-borne threats, even those that were previously unknown.



## Commtouch® Recurrent Pattern Detection Technology

Commtouch has developed a unique and highly successful response to these challenges with its Recurrent Pattern Detection (RPD™) technology, which focuses on detecting recurrent message patterns in outbreaks, rather than on an analysis of the contents of individual email messages. The aim of RPD is to identify and classify all types of email-borne threats and outbreaks. It is content-agnostic and can therefore detect spam and phishing in any language, format or encoding method, while identifying patterns of new email-borne viruses and worms for which signatures have not been made.

RPD is a patented technology (based on Commtouch's patent #6,330,590) that extracts and then analyzes relevant message patterns which are used to identify massive email-borne outbreaks. RPD classifies both distribution patterns and structure patterns and the analysis results are stored in a vast repository of classifications. In addition to identifying new threat patterns, RPD is also used to confirm and enhance the classification of already-identified message patterns.

Commtouch uses RPD technology in a highly scalable environment to deliver extremely high performance rates by automatically analyzing patterns extracted from billions of email messages each day (24x7x365). On average, new outbreaks are identified within minutes from the time they are launched globally on the Internet. RPD technology was designed to be fully automated and requires no human intervention. To ensure maximum privacy and business confidentiality, RPD was designed to analyze encrypted values of message patterns and not the 'open' values nor the message content.

RPD is designed to distinguish between distribution patterns of solicited bulk emails which represent legitimate business correspondence, from those of unsolicited bulk emails by applying a reverse analysis. The results of this analysis are 'bleached' message patterns belonging to 'good' messages such as popular newsletters and mailing lists.

RPD technology is hosted at the Commtouch Global Data Center to proactively analyze massive amounts of Internet traffic in real-time to classify message patterns. The Commtouch Data Center also hosts the Commtouch Signature Repository, a vast warehouse of threat patterns. The Commtouch detection capabilities are distributed among several locations for full redundancy in the cloud.

The Commtouch Data Center is highly scalable, includes provisioning for redundancy and load balancing, and is highly secured from external attacks. New patterns are added to the Commtouch Data Center's vast signature repository in real-time and made available immediately to all Commtouch partners and customers worldwide.

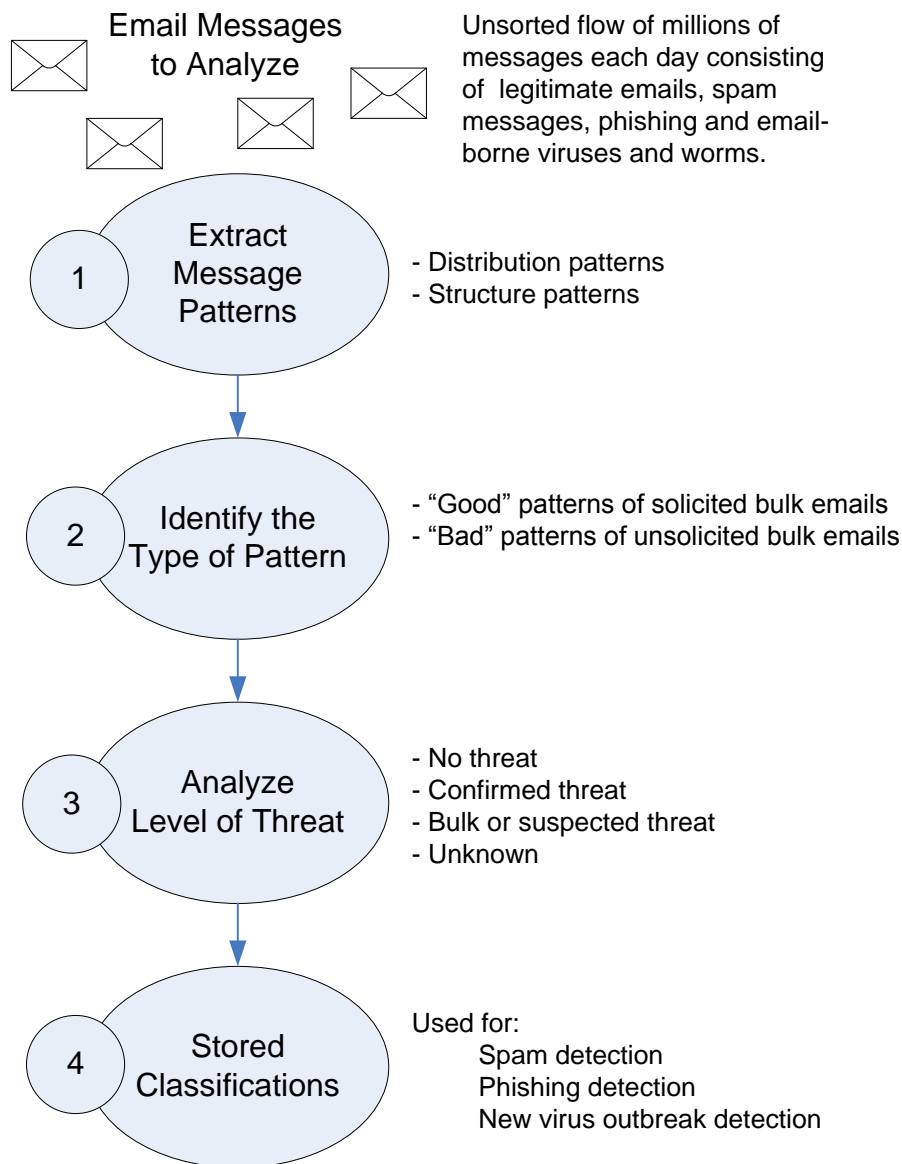
Commtouch RPD technology has several major benefits that carry over to the product offerings built on this technology platform:

- **Real-time detection and blocking** because most attacks last only a few hours and involve the release of tens of millions of emails containing malicious content or spam. If the solution does not work in real-time to detect and block outbreaks proactively, it will simply be reacting towards the end of the attack, when most of the damage has already been done.
- **Language and content-agnostic** because damaging threats are often found in emails that do not contain English characters or messages that use images rather than text. Solutions that focus on the content of a message are likely to miss these threats, decreasing their overall effectiveness and increasing instances of false negatives received by recipients.



- **Adaptive to new methods and tricks** because malicious hackers and sophisticated spammers constantly change their tactics of distribution and infiltration in order to fool and evade current technologies.

### Recurrent Pattern Detection Technology Flow





## GlobalView IP Reputation

RPD technology provides the strong foundation for Commtouch's GlobalView IP reputation technology. Taking the RPD classification, together with an ongoing analysis of multiple attributes, Commtouch can provide an accurate reputation score for each sender. Analyzed attributes include: spam and malware volumes over time, DNS and whois attributes such as domain age, geography, dynamic/static IP addresses.

Commtouch classifies the reputation of each source IP address with a series of meaningful and measurable values both dynamically and in real-time. Real-time classification is critical for blocking the hit-and-run approach of zombie-based attacks. IP values are constantly updated to include the most current information and reflect any change in the credibility of the source.

Having accurate sender reputation data enables the provision of an additional layer of email filtering, done on-session at the perimeter of the network, before full messages are even received. This helps weed out spam messages and email-borne malware at the entry point before these messages enter the messaging network, thereby relieving the need for resource-consuming downstream filtering.

Nowadays, many public services already deliver free information about blacklisted sources; typically, these lists are static with limited commitment for accuracy by the owners. Other commercial services also enable setting central and personal policies for whitelisted sources during the SMTP session. However, these solutions are not designed to handle the vast number of dynamically changing reputations of computers that forced into botnets and armies of zombies.

## Commtouch Messaging Security Suite

### Commtouch Anti-Spam

Based on RPD technology, Commtouch spam-filtering technology differentiates between confirmed spam messages, suspected spam messages and non-spam messages. Non-spam includes both private one-to-one messages and legitimate newsletters or other mass-distributed correspondence. RPD-based Anti-Spam offers these and more immediate benefits:

- Spam detection rates of over 98%
- Almost no cases of false positives
- Protection against phishing attempts
- Content-agnostic threat protection
- Multi-language spam detection
- Multi-format spam detection (e.g. images, PDF, etc.)

### Zero-Hour Virus Outbreak Protection

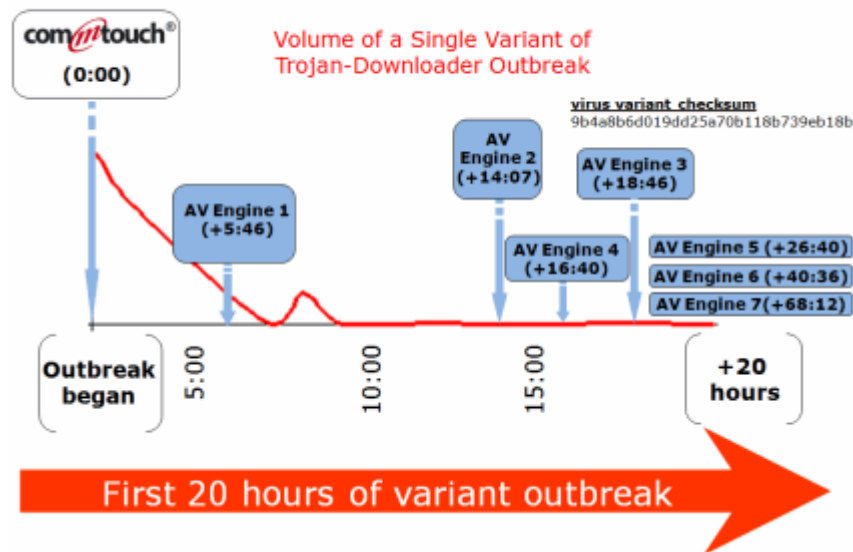
Commtouch Zero-Hour Virus Outbreak Protection (Zero-Hour AV) takes a different approach to malware defense than traditional signature-based or heuristic anti-virus technologies. Instead of focusing on hunting for new viruses and racing to catch them with a signature or heuristic, Commtouch RPD-based Zero-Hour AV identifies and blocks



email-borne malware in real-time, providing immediate protection against new variants, in the first critical hours of an outbreak. It can be used either in conjunction with traditional AV solutions, or on its own to protect against email-borne malware. Benefits of Zero-Hour Virus Outbreak Protection include:

- Real-time blocking
- No reliance on updating signatures or inaccurate heuristics
- Reduction of vulnerability window from hours to 1-2 minutes

**Protects in the first moments of outbreak**



Patent #6-330-590

**GlobalView Mail Reputation**

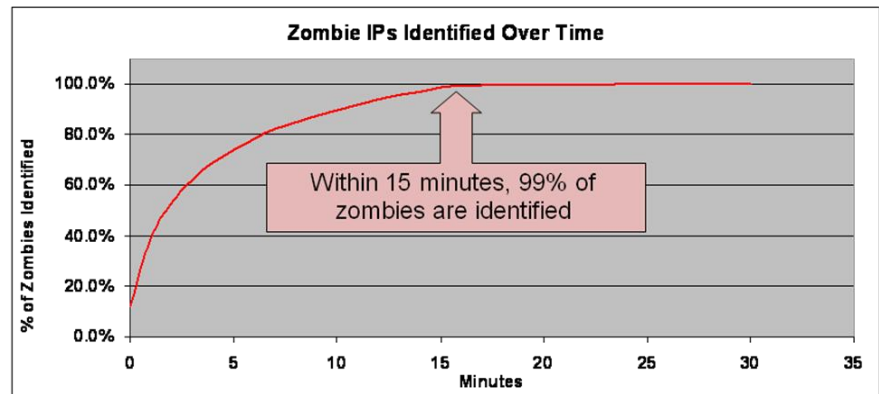
GlobalView Mail Reputation enables on-session blocking of unwanted email senders through the use of rejecting senders either temporarily (“tempfail”) or permanently (“permfail”) and accepting those that have positive reputations. This can be used as part of an overall strategy to optimize network accessibility so the network’s messaging processes are efficient and focused on allowing legitimate sources full and uninterrupted access. At the same time, GlobalView Mail Reputation also makes access more difficult for unauthorized sources with bad reputations attempting to abuse the network since service providers can throttle their connection attempts based on reputation data.

GlobalView Mail Reputation delivers organizational benefits including:

- Reduce IT resources such as server count, CPU load, storage, etc.
- Eliminate multiple security risks
- Reduce the level of false positives
- Minimize the cost of downstream filtering
- Lower overall bandwidth consumption
- Optimize IT labor required to manage the overall messaging process



Commtouch's GlobalView reputation solution responds in real time to the ever-changing botnet landscape, and can identify activated and de-activated zombies within mere minutes. In fact, within 15 minutes of an outbreak, 99% of all the zombies are identified.



## Feedback Loop For All Three Security Layers

One of the prominent advantages of the "in the cloud" messaging security topology is that it enables Commtouch to host multiple services in one central place while leveraging each service's unique data outputs as additional sources of information for other Commtouch solutions. A feedback loop is created between data center services as each service's analysis of the more than two billion transactions per day is adjusted to meet that service's specific needs. Each service was designed to answer a different problem; however, while each service engine is generating its own output, this output can be a valuable asset to other product engines. Sharing these outputs among various services allows Commtouch to enrich the information and to significantly improve detection and false positive rates.

Blended threats are becoming a common tactic for malicious activity, i.e. spam emails directing to a phishing site or Web sites infecting machines with malicious software. Zombies are a main source of blended threats and cyber criminals use the same botnets for multiple fraudulent activities including distribution of spam, malware and phishing. Since multiple threats originate from repeating sources, only a system that sees the "big picture" can provide a holistic and robust solution. Integrating this feedback loop among the various services allows Commtouch to analyze information from all different angles, ensuring best coverage and accuracy results.

## Integration Tools for Hosting Providers

Commtouch provides each of the messaging security suite offerings in various easy-to-integrate formats, including:

- Milter (for Sendmail and/or Postfix)
- SpamAssassin plug-in
- Integration options for Plesk, Qmail, and QpSMTPd
- RBL interface available for GlobalView Mail Reputation

There are many other plug-ins and integration options; please speak with a Commtouch representative to determine if your hosting platform is supported, or write to [nospam@commtouch.com](mailto:nospam@commtouch.com).



## Conclusion

The Commtouch approach to threat detection and protection is based on an understanding of the fundamental challenges constantly posed by today's sophisticated spammers, phishers and malware distributors. More importantly, the Commtouch solution is prepared for future tactics and methods these groups will develop in the future. Because of its modular nature, it is easily adapted to future industry requirements and developments.

Commtouch developed its real-time detection solution in response to the realization that the majority of threat outbreaks cause the most damage in a relatively short period of time from release. Typically, within the first minutes of a release, Commtouch has already proactively identified the outbreak and is able to classify and instruct the host application to block any messages from the outbreak before they reach recipients.

Because the Commtouch approach does not focus on content analysis, it is completely irrelevant whether malicious hackers vary the contents of messages, use images, non-English characters, or single or double byte encoding, etc. Even when virus authors release multiple instances of the same virus within an attack, Commtouch is able to track and block all variations.

## About Commtouch

Commtouch® (NASDAQ: CTCH) is the source of proven messaging and Web security technology for scores of security companies and service providers, founded on a unique cloud-based datacenter approach. Commtouch's expertise in building efficient, massive-scale security services has resulted in its patented technology mitigating Internet threats for thousands of organizations and hundreds of millions of users in more than 100 countries. Commtouch technology automatically analyzes billions of Internet transactions in real-time to identify new threats as they are initiated, protecting email infrastructures and enabling safe, compliant browsing. The unmatched suite of Commtouch security offerings is based on patented Recurrent Pattern Detection (RPD™) and GlobalView™ technologies, which work together in a comprehensive feedback loop and offer equally effective protection for all languages and formats. Commtouch was founded in 1991, is headquartered in Netanya, Israel, and has a subsidiary in Sunnyvale, Calif.

Stay abreast of the latest news at the Commtouch Café: <http://blog.commtouch.com>. For more information about enhancing security offerings with Commtouch technology, see <http://www.commtouch.com> or write to [info@commtouch.com](mailto:info@commtouch.com).

© 2009 Commtouch Software Ltd. All rights reserved. Recurrent Pattern Detection, RPD, Zero-Hour and GlobalView are trademarks, and Commtouch is a registered trademark, of Commtouch Software Ltd. U.S. Patent No. 6,330,590 is owned by Commtouch. All other trademarks and registered trademarks are the property of their respective owners.