

THE MARRIAGE OF SPAM AND MALWARE: IMPLICATIONS FOR SMTP MALWARE DEFENCE

Amir Lev

CommTouch Software Ltd, 4A Hatzoran St. PO
Box 8511, Netanya 42504, Israel

Tel +972 9 8636816

Email amir.lev@commtouch.com

ABSTRACT

Once considered two distinct entities, spam and email-borne malware are becoming increasingly similar. Both take advantage of email as a primary means of mass-distribution, and both use botnets to launch outbreaks and evade detection. As attacks continue to get bigger, faster and more severe, the industry must reassess the strategy of defending against each type of threat individually.

THE COMMON GROUND: EMAIL-SENDING BOTNETS

Spam and malware are two of the most menacing Internet plagues. Historically the industry has dealt with them separately, but today a common denominator has emerged – email-sending botnets. Zombie PCs are responsible for sending an estimated 85% of global spam [1].

Trojan malware is used to hijack machines, turn them into zombies and add them to botnets. Though not perceived by many IT managers as a serious threat [2], botnets can be used to carry out an array of malicious activity, including further propagation of all types of virus. Spam and malware also share a common distribution vector – email. Email is obviously the main delivery mechanism for spam, but it is also the leading infection source for malware, accounting for 27% of malware infections overall [3]. Given that a large proportion of malware is sent via email, the question arises, ‘Are spam-sending bots also sending virus-infected email?’

For the past few years bots have been successfully involved in several types of malicious activity and sophisticated blended attacks. A notable example was the attack started by the seeding of the Storm/Nuwar trojan in early 2007 [4].

This explorative research focuses specifically on the sender correlation between spam and email-borne malware. The inspiration for this study is the assumption that if a high percentage of email-borne malware is found to be sent by bots that also send spam, it is possible that a spam-based IP reputation solution could be used to help identify, classify and block a comparable percentage of email-borne malware. For this study a practical test was designed to discover the extent to which spam-sending bots also send malware.

RESULTS AND IMPLICATIONS

Primary result: 65% of email-borne malware is sent from spam-sending IPs.

Our research showed that 65% of email-borne malware was sent by IPs known to be sending spam. While it is not

surprising to see proof that a single bot can be used for several types of malicious activity, this experiment provides quantitative evidence of the extent of the correlation. The primary practical implication of this high correlation is that a currently existing spam botnet detection system can be utilized immediately to block approximately 65% of incoming SMTP malware on-session, at the network perimeter. Real-time blocking of a large percentage of malware-infected emails during the SMTP session can decrease the time it takes to protect customers from new outbreaks, since no signature, heuristic or prior knowledge of the virus variant is needed. A secondary benefit of proactively identifying malicious SMTP connections is the potential to expedite the collection of new malware outbreak samples.

STUDY OUTLINE

Data sources

Spam bot detection system

The *CommTouch* botnet detection system dynamically detects spam and malware sending sources using Recurrent Pattern Detection (RPD) technology. RPD analyses billions of messages of global email traffic per week, and identifies and classifies attack patterns to detect spam email outbreaks. The massive outbreaks which distribute spam consist of 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. These values, called ‘attack patterns,’ are detected by RPD within the first moments of a new outbreak. Because tactics for distributing spam are constantly evolving, RPD proactively identifies new and unique patterns in real time in order to determine new outbreaks as they are released to the Internet and begin targeting recipients.

Once a new attack pattern is identified, all source IPs sending that pattern are identified. The source IPs can be assumed with a high degree of certainty to be currently participating in the same attack, though they are not necessarily a distinct botnet. This pattern identification enables real-time mapping of IPs actively participating in each spam outbreak.

Zero-hour malware detection

The Zero-Hour Malware Detection system applies the RPD technology to detect email-borne malware outbreaks. A series of patterns are extracted from the executable content. Recurring instances of existing patterns are recognized and new patterns are also detected. This allows the system to constantly learn about new outbreaks the moment they begin. The real-time identification of these values is considered in the final decision about whether the executable is malware or benign. Risk level is determined very quickly since distribution behaviour alone indicates the level of suspicion and there is no need for time-consuming analysis of the binary. System accuracy is 100% for malware classified as ‘Virus’ and approaching 100% for malware classified as ‘High Risk’. Over 99.7% of all infected traffic falls into these two categories.

Test sample

Over the course of an 11-day period in May 2007, a daily sample of 5,000 virus-infected email messages was taken from

the Zero-Hour Malware Detection Centre database. The malware sender IPs were then aggregated and duplicates removed. The distinct IPs were sent to query the dynamic botnet detection system to map overlap with spam-sending sources.

It should be noted that the test period was relatively quiet in terms of email-borne malware outbreaks. If the experiment were to be repeated in the midst of a massive attack, the results might be markedly different. Also worth mentioning is the discovery of a few email messages carrying both malware and spam content. Analysis of these double-carrier malicious email messages was beyond the scope of this study.

Methodology

To test the sender correlation, daily samples of new global virus outbreaks were taken from the zero-hour malware detection centre. The true sender IPs were extracted from the infected messages. The malware source IPs were sent to query the spam botnet detection system to determine if they were known as spam-sending bots. This shows the source correlation between malware and spam.

Results

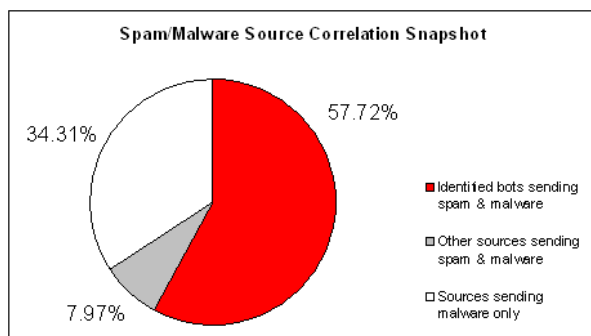
IP category descriptions*:

- **Identified bots sending spam and malware.** This category represents IPs that actively sent both spam and malware during the test period. Had a spam-oriented IP reputation system been in use, these malware-infected messages could have been blocked during the SMTP session.
- **Sources sending malware only.** This category represents IPs that actively sent only malware during the test period.
- **Other sources sending spam and malware.** This category represents IPs that actively sent both spam and malware during the test period. Though these IPs did send malicious email, the data was not sufficient to prove they are indeed bots, and as such are classified as 'other sources'.

**Note: Although the botnet detection system is designed to return a variety of classifications per source IP, to simplify the test as a proof-of-concept these were aggregated into the above three classes.*

Practical implications for SMTP malware defence:

1. 57% of malware was sent by machines known as spam-sending bots.



2. An additional 8% of malware was sent from spam-sending sources that are not confirmed as bots, although some of them may very well be bots.
3. Because of this high correlation, blocking known spam sources has the potential to prevent penetration of approximately 65% of email-borne malware at the network perimeter, during the SMTP session.
4. Blocking spam-sending IPs serves as a first-line of defence against email-borne malware and reduces the time and resources required to protect customers against new variant outbreaks.
5. IT resources are saved since half of the malicious messages never enter the network.
6. Messages rejected during the SMTP session do not require the MTA to generate bounce-back messages, improving networking and processing utilization.

Notes on results

1. The data presented above provides a snapshot view of the malicious SMTP activity of botnets. The real population data is in constant flux. Therefore, results should not be interpreted as constant.
2. The botnet detection system utilized in the study is a dynamic system that identifies and blocks bots in real time. When a bot is inactive for a sufficient period of time it expires in the system. Therefore, the data presented here provide a snapshot of spam activity in the period directly prior to the malware detection. IPs that appear as 'sources sending malware only' may indeed have sent spam in the past, but have since become dormant. IPs that sent spam in the period after the test will appear here as 'sources sending malware only'. If a longer time period were considered, the overlap percentage might actually be higher.
3. For the purposes of this study, no differentiation was made between the types of malware sent. It is possible that if each type of malware were tested separately the results would be dramatically different. For example, if a certain malware type is emailed out only by botnets, the 'spam and malware zombie' value would be 100%. Conversely, malware types that are distributed only by non-botnet methods would show 0%.

CONCLUSION

The test performed in this study showed the practical potential of spam-based bot detection systems to better defend against email-borne malware. These results strongly suggest that the integration of new bot-detecting anti-spam technologies could:

- Block over 50% of malware-sending sources .
- Significantly reduce the volume of new incoming malware variants.
- Lead to improvements in AV protection rates.

Multi-purpose botnets are increasing the pressure on anti-virus and anti-spam communities alike. Spam and malware have already united to carry out malicious activities. Perhaps the time has also come for the anti-spam and anti-virus industries to join forces.

REFERENCES

- [1] Leyden, J. Zombie PCs Spew Out 80% of Spam. June 2004. http://www.theregister.co.uk/2004/06/04/trojan_spam_study/.
- [2] Types of Malware Threats Ranked by Seriousness. Computer Economics: 2007 Malware Report, p.17.
- [3] Malware: vectors for infection. Computer Economics: 2007 Malware Report, p.15.
- [4] Stewart, J. Storm Worm DDoS Attack. Secure Networks. February 8, 2007.