Recent research results concerning email spam threats and mitigation

Christian J. Dietrich
dietrich@internet-sicherheit.de

Christian Rossow
rossow@internet-sicherheit.de

Prof. Dr. Norbert Pohlmann
norbert.pohlmann@informatik.fh-gelsenkirchen.de

Institute for Internet Security
University of Applied Sciences Gelsenkirchen
Germany
http://www.internet-sicherheit.de

Recent spam development

Spam has been an IT security topic of high importance throughout 2007 and still it poses a significant threat to email usage on the internet. ENISA's antispam efforts during 2007 enlightened in detail the security measures taken by European Email and Internet Service Providers. Especially the comparison statistics of filtering methods shows a significant increase in the use of all kinds of ingress as well as egress filtering methods from 2006 to 2007. Inevitably, the question arises: As the antispam measures spread, can we see an improvement concerning the spam development on the Internet? Is it getting any better?

By the help of sensors that measure a representative fraction of European Internet email traffic, we are able to get an insight into the spam traffic development. The evaluation of our measurement system is based on email traffic from at least 1.3 million different sending email servers per day. On the recipient side we cover a daily set of about 10,000 unique email receivers, corresponding to more than 10,000 different target domains.

As often, there are two sides of the story. From the email user's perspective, our evaluation shows that less spam is actually being stored in users' mailboxes, i.e. users are generally prevented from receiving more spam. This result is probably mainly achieved due to widespread filtering measures of Email Service Providers.

However, on the downside, we have detected an ever-increasing spam volume since mid 2007, measured before spam hits a mailbox. From July 2007 until beginning of March 2008, spammers tried to deliver more and more spam. In the first week of March 2008 the spam volume peaked, having increased by a factor of 9-10 over the last 9 months in comparison to July 2007.

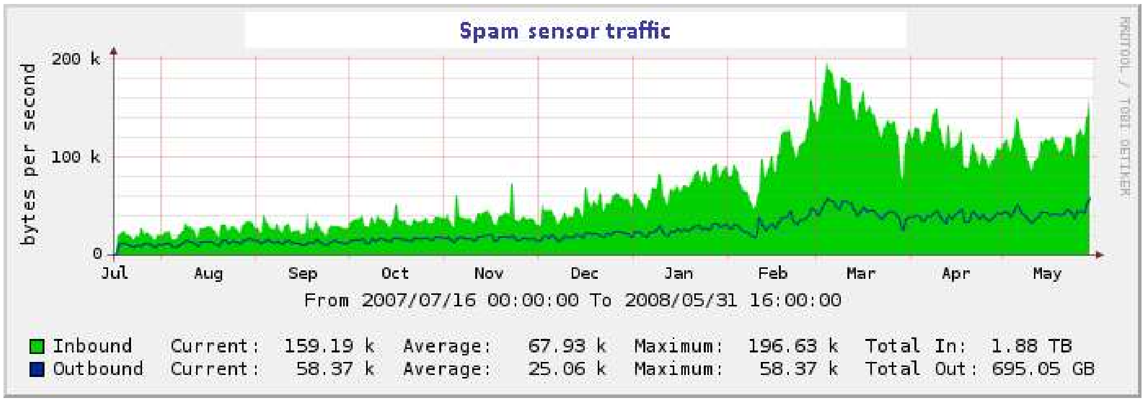

Figure 1: Spam sensor traffic from July 2007 to May 2008.

At first glance, the high spam volume on the Internet seems to contradict the comparably low amount of spam messages in mailboxes. However, the explanation of these two facts is simple: The high spam volume on the Internet reflects the connection attempts and the delivery efforts of spammers. Actually, due to the wide spread use of spam filtering mechanisms and the improved detection rates of spam filters, only few of the spam messages make it through and actually reach the mailboxes. To further elucidate: an antispam method that correctly detects 99% of all spam messages requires a spammer on average to deliver 100 spam messages in order for one spam message to make it through, statistically. Although this example is simplified, it shows the extent to which spammers adapt to the enhancements made in antispam techniques.

Spam sources

Furthermore, not only the volume of spam increases, but also the distribution of spam has changed. Nowadays, botnets are responsible for most of the spam sent on the Internet. The fact of abusing someone else's computer makes it easy for a spammer to cover the tracks. In addition, botnets are difficult to detect before any abusive behavior occurs. And even if a botnet is detected, quite an effort is needed and several parties are involved in order to take it down. Botnets are highly dynamic to evade countermeasures.

Unfortunately, botnets are not only used for spamming, but also for other kinds of misuses, such as hosting phishing sites, mounting distributed denial of service attacks or spying confidential data from the infected computers. Among those abusive actions, spam is the most obvious one because the bots play an active role – they send out messages that can be detected by spam filters. Thus, detecting spam sources can not only be used to protect from spam in the future, but could also be useful to protect from possible other threats such as phishing or denial of service attacks. All in all, a reputation service could summarize and collect different kinds of abuse and make that information available to others.

Challenges

A couple of challenges remain in order for such a reputation service concept to actually work. Spam is sent in a very distributed fashion. On the one hand, this could make it difficult to detect a sufficiently large number of spam sources from only a small number of vantage points. This means that, from the perspective of effectiveness and benefit, many different vantage points are required. On the other hand, the lack of one sole vantage point, where all mail is screened, is a big advantage in terms of data protection. Thus, a reputation service should be designed and setup in a distributed fashion. A distributed system prevents a single entity from having control over all data. Moreover, the distributed system enables many reputation sources to feed and store information, which solves the problem of having one single point of data storage.

Another challenge is posed by the fact that spammers may switch to sending spam at a low volume in order to evade spam detection techniques. In fact, even today, few spammers already send only very few spam messages from one source. Fortunately, as the quality of spam filters improves, sending spam at low volume will probably render spamming unprofitable due to its very nature. Apart from that, spammers need web space where they can setup their online stores and where people can buy the spamvertised products. Nowadays, these hosts are part of a botnet, too. Thus, detecting botnets by spam sources could also mitigate falling for fake online stores – another advantage of a distributed reputation service.

Both challenges, the distributed nature of spam as well as an adaption of spammers tactics, can be mitigated by strong cooperation in the distributed system. If a sufficient number of service providers exchanged information about abused sources, for example spam sending hosts, such a system could be put in place and it would actually help to avoid other abuses.