The Anti-Social Behavior of Spam

Farnaz Moradi, Tomas Olovsson, Philippas Tsigas
Computer Science and Engineering
Chalmers University of Technology
Gothenburg, Sweden
Email: moradi,tomasol,tsigas@chalmers.se

***Abstract*—Spam mitigation strategies that aim at detecting spam on the network level, should classify email senders based on their sending behavior rather than the content of what they send. To achieve this goal, we have performed a social network analysis on a network of email communications. Such a network captures the social communication patterns of email senders and receivers. Our social network analysis on email traffic have revealed that structural properties of networks of email communications differ from other types of interaction and social networks such as online social networks, the web, Internet AS topology, and phone call graphs. The difference is caused by extensive amount of unsolicited email traffic which therefore can be used to discriminate spam senders from legitimate users. Deployment of such social network-based spam detection strategy on a small network device makes it possible to stop spam closer to its source and without inspecting email contents. In this presentation, we will look at the anti-social behavior of spam and how it can be used for detection of spam senders.**

SUMMARY

We have studied and analyzed the behavior of email traffic by examining network traffic captured on an Internet backbone link of a national university network. From the collected packet level data, we have generated a number of *email networks*. An email network is an explicit social network with email sender and receiver addresses as nodes and the exchanged emails as edges.

The structural and temporal properties of the generated email networks have been studied and compared with the properties of other social networks such as online social networks, the Internet topology, the web graph, phone call graphs, and other email networks. Our study have shown that the structure of email networks generated from real traffic containing unsolicited email (*spam*) cannot be accurately modeled similar to other social networks .

We have shown that the difference between the structure of our email networks and the structure of other types of social networks is caused by the anti-social behavior of the spam traffic. The legitimate email (*ham*) traffic exhibits the same structural and temporal properties that have been observed in other social networks, which means that a ham network can be modeled as a scale-free small-world network. This observation is interesting since the generated email networks do not contain a complete set of email communications between all the email addresses, however, the social behavior of legitimate email communications can still be observed. Furthermore, a spam network containing only unsolicited email traffic exhibits anti-social behavior, which is not hidden behind the social behavior of legitimate traffic. Therefore, the structural properties of email networks containing both type of email deviates from existing models for social networks.

As a future work, it would be interesting to investigate the deployment of these distinctive structural properties in a network device monitoring traffic to detect nodes with anti-social behavior without even inspecting the content of what they are sending. Such an approach can potentially be used as a complement to existing anti-spam tools to stop spam closer to its source.

ACKNOWLEDGMENTS

This work was supported by .SE – The Internet Infrastructure Foundation and SUNET. The research leading to these results has also received funding from the European Union Seventh Framework Programme (FP7/ 2007-2013) under grant agreement no. 257007.