Metrics for Measuring ISP Badness: The Case of Spam (Short Paper)

Benjamin Johnson[1], John Chuang[2], Jens Grossklags[3], and Nicolas Christin[4]

[1] Department of Mathematics, University of California, Berkeley
[2] School of Information, University of California, Berkeley
[3] College of Information Sciences and Technology, The Pennsylvania State University
[4] Information Networking Institute and Cylab, Carnegie Mellon University

Abstract

We consider the problem of ISP targeting for spam prevention through disconnection. Any such endeavor has to rely on adequate metrics that consider both the badness of an ISP as well as the risk of collateral damage. We propose a set of metrics that combines the two. Specifically, the metrics compare each ISP's "spamcount" with its "disconnectability". We offer a concrete methodological approach to compute these metrics, and then illustrate the methodology using datasets involving spam statistics and autonomous system relationships. This analysis represents the first step in a broader program to assess the viability of economic countermeasures to spam and other types of malicious activity on the Internet.

1 Introduction and Related Work

Recent studies have shown that a large percentage of all spam on the Internet is attributable to sources from a small percentage of the Internet's autonomous systems [3, 8, 24]. Thus in considering the spam-prevention problem, it makes sense to concentrate attention on those few systems who contribute to the problem the most. Evidence suggests that targeting an especially bad player can be effective. For example, the November 2008 takedown of McColo [6, 9] resulted in a significant decline in the global volume of spam (by estimates as much as 70%) [19].[5]

In this work, we address the targetability question by defining metrics to determine when an autonomous system is doing substantially more harm than good. These metrics can then be used not just for assessing the feasibility of targeting an ISP, but also for recognizing which ISP's may be susceptible to economic incentive structures designed to elicit implementation of outbound preventative mechanisms. The set of metrics we propose are based on ratios between "badness" measures which quantify the ill effects an autonomous system (AS) poses to the rest of the network, and "disconnectability" measures that quantify the collateral damage that would result from the AS's disconnection from the Internet graph.

[5] There are also bad registrars but that is a different story [18].

This research was partially funded by the National Science Foundation under ITR award CCF-0424422 (TRUST). This paper appeared in the Proceedings of the 2012 Financial Cryptography and Data Security Conference. This is an author preprint for personal use only.

1.1 Spam Measurement Studies and Mitigation

There is a diverse and growing literature on the measurement of spam and other network threats. Ramachandran and Feamster showed that network level properties can reveal decisive cues in the fight against spam [24]. Other contributions are the development of behavior-driven or signature-driven spam classification systems which are desirable given the transient nature of spam origins [12, 14, 25, 27, 33].

A number of studies have investigated the problem of botnet identification [11, 23, 34]. Ehrlich *et al.* proceed with a two-step methodology. First, they identify individual botnet nodes. Second, they continue their investigation to determine the command-and-control infrastructure of the botnet. In related projects, research groups have worked on building infrastructures to identify rogue networks [15, 28].

Finally, a growing number of research studies is concerned with the better understanding of spam economics, by studying large-scale spam campaigns [1, 4, 16], and by tracing click trajectories to better understand the spam value chain [20].

1.2 Economics of Service Provisioning and Security

Several reports have explored the economic incentives of Internet Service Providers to invest in security measures [3, 32]. A key observation is that ISPs respond differently to emerging threats leading to varying degrees of botnet infestations in their user population [3]. Some ISPs may act vigorously, while others appear to be slacking [5]. Finally, a residual group aims to derive a profit from providing a safe harbor for undesirable activities [8].

ISPs are in an excellent position to address security problems [2]. However, it is an open debate whether or to what degree liability can be assigned to them for insufficient or even detrimental behaviors [21].

But even from the perspective of well-motivated ISPs it is not obvious how to address security threats in a cost-efficient manner [26]. ISPs can incentivise users to higher security vigilance, but there are tradeoffs. Some incentive schemes target higher individual security effort levels [29], while others focus more on group-level security outcomes [13].

Another approach is to reduce the autonomy of individual users by installing security client software that monitors and controls network access. However, the majority of consumer-oriented ISPs shy away from direct technical intervention involving access to the users' home resources. Some argue this to be a government's role [7]. We are only aware of one US consumer ISP experimentally testing a similar approach [22]. However, several ISPs utilize redirection and quarantining techniques to encourage users to engage in clean-up efforts [17].

The rest of the paper is organized as follows. In Section 2, we define several metrics for use in ranking autonomous systems according to their miscreant behavior and discuss their key properties. In Section 3, we briefly describe our methodology for computing the proposed metrics on real data. Section 4 contains examples and illustrations pertaining to the ASes responsible for the most spam volume in our dataset. We discuss plans for future work and conclude in Section 5.

2 Proposed Metrics

In this section, we introduce metrics to formally quantify the *badness* and *disconnectability* of autonomous systems, along with the *cost-benefit tradeoff* that can be used as part of a formal basis for decisions involving AS targeting.

With respect to a given set of connectivity relations between autonomous systems, we define the following associated set of ASes. The ***exclusive customer cone*** of an autonomous system X is the set of autonomous systems that would become disconnected from the network if X were completely disconnected from the rest of the network. Note that every AS is a member of its own exclusive customer cone. The exclusive customer cone of X is a subset of the *customer cone* of X, which was defined by Dimitropoulos *et al.* as the set of customers of X together with those customers' customers, and so forth [10]. By way of contrast, the exclusive customer cone does not include those customers and subcustomers of X with a connection to at least one additional provider that can move traffic to the core of the network while avoiding X. The customer cone is a reasonable measure of the importance of X, but the exclusive customer cone has a more direct bearing on the question of whether to target X.

We next define a set of three metrics related to the exclusive customer cone. The ***exclusive customer cone size*** of X is the number of ASes in the exclusive customer cone of X. The ***exclusive customer cone prefix size*** of X is the number of distinct /24 prefixes assigned to ASes in the exclusive customer cone of X. The ***exclusive customer cone address size*** of X is the number of IP addresses assigned to ASes in the exclusive customer cone of X.

With respect to a given set of data attributing spam to IP addresses, we define the following two metrics. The ***spamcount*** of an autonomous system X is the number of spam messages attributable to IP addresses directly assigned to X. The ***spamipcount*** of X is the number of distinct IP addresses directly assigned to X that are responsible for sending at least one spam message. We can obviously extend these definitions to negative attributes other than spam. For example, we could analogously define the ***badness*** of X and the ***ipbadness*** of X relative to any data set that associates a measure of badness to certain IP addresses.

Lastly with respect to both a badness measure on IP addresses and a set of connectivity relations among autonomous systems, we define a set of ratio metrics comparing the two. For example, the ***spamipcount to exclusive customer cone prefix size ratio*** of an autonomous system X is the ratio of the spamipcount of X to the exclusive customer cone prefix size of X.

3 Methodology

In this section, we briefly describe a methodological approach to computing these metrics on real data.

The metrics require three types of data for autonomous systems: a notion of badness for each AS; a measure of size for each AS; and the AS customer/peer relationship structure. The current state of the art in relationship structure is publicly available via the Cooperative Association for Internet Data Analysis (CAIDA) [31]. CAIDA also publishes size measures for autonomous systems.

Measures of badness come in many forms, and good datasets are more difficult to obtain. We carried out an illustration of our methodology using a data source consisting of about 3.4 million spam email messages collected by Ramachandran and Feamster at Georgia Tech over a period of about 17 months starting in July 2008. The methodology of that data collection was previously used and is described in an earlier publication [24]. We processed these emails to obtain a best guess source IP for each email; and then used the bulk Whois query tools provided by Team Cymru [30] to associate IP addresses obtained from the spam dataset with their associated autonomous system numbers.

4 Examples and Observations

The purpose of this section is to exemplify our methodological approach through illustration.

4.1 Exclusive Customer Cone Properties

Figure 1 in the appendix offers visualizations of the exclusive customer cones for two of the five most spammy ASes in our dataset: 4134 and 4837. The edges reflect provider-to-customer relationships. The diameter of each node increases linearly with the node's spamcount, and the color of the node moves from green to red as the spamipcount to exclusive customer cone prefix size ratio increases.

4.2 Aggregated Spam Characteristics

Figure 2 in the appendix shows the distributions of spamcount and spamipcount across the top 300 most spam-facilitating ASes in our dataset, as well as the cumulative distribution of these same badness measures over our entire dataset. The graphs confirm that most of the spam in our dataset is generated by IP addresses from a small percentage of autonomous systems, consistent with many other measurement studies involving spam [11, 23–25, 34]. This feature offers justification for our focus on only a few of the worst offenders. The graphs also give a sense of the relationship between spamcount and spamipcount, showing that both measures follow similar distributions and are strongly correlated.

4.3 Targeting ASes with High Spamcount

Table 1 below gives the computed value of several metrics for each of the top five highest-ranked ASes by spamcount.

The spam ip to exclusive customer cone prefix size ratio is given in the table's last column. As mentioned previously, this is our preferred metric for quantifying targetability. We see from the table that one particular autonomous system, ASN 45899, VPNT Corp, has a high score under this metric. It is a stub AS, with no customer ASes; it has a high badness score using both spamcount and spamipcount – in fact the spamcount is more than double that of any other AS in our entire dataset; and it does not have any

Table 1. ASes with the highest spamcount in our dataset

ASN	Organization Name	Spam Count	Percent of Spam	Spam IP Count	Customer Cone Prefix Size	Exclusive Cust Cone Prefix Size	Spam IP to Prefix Size Ratio
45899	VNPT Corp	132918	7.9%	30222	12	12	2518
4766	Korea Telecom	54237	3.2%	6833	203306	2537	2.7
4134	Chinanet Backbone	44482	2.7%	10360	66172	1361	7.6
7738	Telecomunicacoes da Bahia S.A.	39570	2.4%	7904	111324	518	15.3
4837	China169 Backbone	36617	2.2%	8269	6438	1664	5.0

counterbalancing large number of IP addresses or /24 prefixes in its exclusive customer cone.

The next few autonomous systems on the list have more exclusive customers and are generally much higher up in the internet tier hierarchy. The ratio metrics applied to these ASes is not nearly so high, by several orders of magnitude; and it would be more difficult to justify a targeting strategy for any of the next four ASes based on this metric.

Due to current limitations of our data on the badness side, our true proscription towards actually targeting VPNT Corp is perhaps not very strong. The intent of this paper is not to make targeting proscriptions, but rather to introduce and illustrate a useful methodology. We have shown that well-motivated targetability metrics can be computed and applied to real ASes, using plausible data, with interesting and highly variable results. As our badness measurements become better quantified with higher quality data, proscriptive targeting arguments can be supported on this basis.

4.4 Discussion

What are the relative advantages of the different metrics? Our view is that it depends on the application and the quality of the datasets involved. The metrics invoking badness only at the IP level are more robust in the sense that they are less likely to change quickly over time, (since, for example, it may be easier to send additional spam messages from the same compromised IP address, then to compromise an additional IP address), and are also more applicable to the problem of diagnosing the source of badness. On the other hand, metrics equating badness with direct aggregate features such as raw spam volume are in better correspondence with the negative effects imposed on the rest of the network. In a similar fashion, metrics involving the *exclusive customer cone size* and *exclusive customer cone prefix size*, are more robust in the sense that they are less likely to change significantly if connectivity relations change by not too much. Further, if the data being used is accurate and not likely to change, then metrics involving the exclusive customer cone address size most directly correspond to the aggregate benefit to the network that would be lost if an autonomous system were targeted.

We want to emphasize that each of these metrics, based on exclusive customer cones, is a conservative metric, in the sense that the metric will generally paint an ISP as "less targetable" than it would be painted if more edges were included. This feature is important, as the connectivity data we use is approximate and is known to err on the side of omitting edges from the graph [10]. A practitioner might worry that targeting a particular AS might cause more collateral damage than expected. This would likely be the case if we used a metric based on customer cone. In such circumstance, an ISP who might otherwise be subject to targeting would have an incentive to simply add more customer edges to lower their disconnectability measure. But because we are using the exclusive customer cone, such a strategy would not be effective. Moreover, the customers of a bad AS would have the ability to adopt an alternative provider, resulting in a simultaneous decrease in target risk for the customer, and an increase in the disconnectability of the bad provider. The idea is that by publishing a metric, we create an incentive structure that tends to isolate the bad guys from the good guys. As miscreant behavior of an AS increases, the good guys have more and more of an incentive to line up an alternate provider, to avoid becoming collateral damage.

5 Conclusions and Future Work

Today's spam problem involves many players with competing interests; and any solution requires numerous tradeoffs. In our study of this problem, we have focused our attention on the players we see as having the most power to resolve the problem, namely the ISPs. This approach has lead us to an investigation of metrics applicable to ISP targeting through disconnection. Policy makers must consider a wide variety of technical, policy, and incentive-relevant challenges to realizing ISP disconnections in practice. Our contribution to this effort has involved demonstrating the tractability of developing an objective framework for addressing the problem.

Our goal is to continue advancing research on the prevention of Internet-related misbehavior through the publication of metrics that can affect an ISP's reputation. We consider our program as parallel to more direct technological approaches for combatting spam, such as botnet targeting and spam filtering. By developing an objective framework for considering ISP targeting through disconnection, we advance the tools available to economic researchers for use in the modeling of the Internet ecosystem, and help to foster a better understanding of the problem from crucial perspectives.

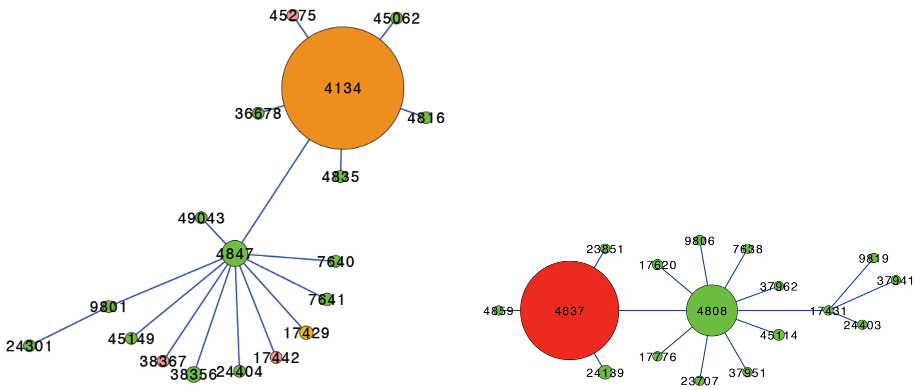
References

1. ANDERSON, D., FLEIZACH, C., SAVAGE, S., AND VOELKER, G. Spamscatter: Characterizing internet scam hosting infrastructure. In *Proceedings of 16th USENIX Security Symposium* (Boston, MA, Aug. 2007), pp. 135–148.
2. ANDERSON, R. Why information security is hard - an economic perspective. In *Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC'01)* (New Orleans, LA, Dec. 2001).
3. ASGHARI, H. Botnet mitigation and the role of ISPs: A quantitative study into the role and incentives of internet service providers in combating botnet propagation and activity, Jan. 2010. Master Thesis, Delft University of Technology.

4. BÖHME, R., AND HOLZ, T. The effect of stock spam on financial markets. In *Proceedings of the Fifth Annual Workshop on Economics and Information Security WEIS'06* (Cambridge, UK, June 2006).
5. CLAYTON, R. Using early results from the 'spamHINTS' project to estimate an ISP Abuse Team's task. In *Proceedings of the Conference on E-Mail and Anti-Spam (CEAS)* (Mountain View, CA, July 2006).
6. CLAYTON, R. How much did shutting down McColo help? In *Proceedings of the Conference on E-Mail and Anti-Spam (CEAS)* (Mountain View, CA, July 2009).
7. CLAYTON, R. Might governments clean-up malware? In *Proceedings of the Ninth Annual Workshop on Economics and Information Security WEIS'10* (Cambridge, MA, May 2010).
8. DANCHEV, D. Bad, bad, cybercrime-friendly ISPs. `http://blogs.zdnet.com/security/?p=2764`, March 4 2009.
9. DIBENEDETTO, S., MASSEY, D., PAPADOPOULOS, C., AND WALSH, P. Analyzing the aftermath of the McColo shutdown. In *Proceedings of the Ninth Annual International Symposium on Applications and the Internet (SAINT)* (Seattle, WA, July 2009), pp. 157–160.
10. DIMITROPOULOS, X., KRIOUKOV, D., FOMENKOV, M., HUFFAKER, B., HYUN, Y., CLAFFY, K., AND RILEY, G. AS relationships: Inference and validation. *ACM Computer Communication Review 37*, 1 (Jan. 2007), 29–40.
11. EHRLICH, W., KARASARIDIS, A., LIU, D., AND HOEFLIN, D. Detection of spam hosts and spam bots using network flow traffic modeling. In *Proceedings of the 3rd USENIX Workshop on Large-scale Exploits and Emergent Threats (LEET)* (San Jose, CA, Apr. 2010).
12. ESQUIVEL, H., MORI, T., AND AKELLA, A. Router-level spam filtering using TCP fingerprints: Architecture and measurement-based evaluation. In *Proceedings of the Conference on E-Mail and Anti-Spam (CEAS)* (Mountain View, CA, July 2009).
13. GROSSKLAGS, J., RADOSAVAC, S., CÁRDENAS, A., AND CHUANG, J. Nudge: Intermediaries' role in interdependent network security. In *Proceedings of the Third International Conference on Trust and Trustworthy Computing (TRUST 2010)* (Berlin, Germany, June 2010), pp. 323–336.
14. HAO, S., SYED, N. A., FEAMSTER, N., GRAY, A. G., AND KRASSER, S. Detecting spammers with snare: Spatio-temporal network-level automatic reputation engine. In *USENIX Security Symposium* (2009), USENIX Association, pp. 101–118.
15. KALAFUT, A., SHUE, C., AND GUPTA, M. Malicious hubs: Detecting abnormally malicious autonomous systems. In *Proceedings of the 29th IEEE International Conference on Computer Communications (INFOCOM)* (San Diego, CA, Mar. 2010), pp. 326–330.
16. KANICH, C., KREIBICH, C., LEVCHENKO, K., ENRIGHT, B., VOELKER, G., PAXSON, V., AND SAVAGE, S. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the Conference on Computer and Communications Security (CCS)* (Alexandria, VA, Oct. 2008), pp. 3–14.
17. KIRK, J. ISPs report success in fighting malware-infected PCs, June 2009. `http://www.pcworld.com/businesscenter/article/166444/isps_report_success_in_fighting_malwareinfected_pcs.html`.
18. KNUJON. 2009 Registrar Report. `http://knujon.com/registrars/#feb09RegistrarReport`, Feb. 2009.
19. KREBS, B. Takedowns: The shuns and stuns that take the fight to the enemy. *McAfee Security Journal 6* (Fall 2010), 5–8.
20. LEVCHENKO, K., CHACHRA, N., ENRIGHT, B., FELEGYHAZI, M., GRIER, C., HALVORSON, T., KANICH, C., KREIBICH, C., LIU, H., MCCOY, D., PITSILLIDIS, A., WEAVER, N., PAXSON, V., VOELKER, G. M., AND SAVAGE, S. Click Trajectories: End-to-End Analysis of the Spam Value Chain. In *Proceedings of 32nd annual Symposium on Security and Privacy* (May 2011).

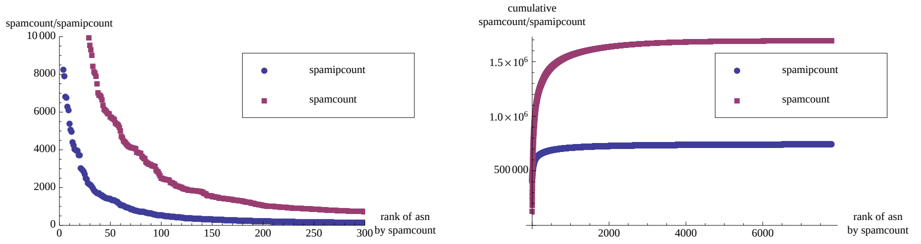
21. LICHTMAN, D., AND POSNER, E. Holding Internet Service Providers accountable. *Supreme Court Economic Review 14* (Apr. 2006), 221–259.
22. MILLS, E. Comcast pop-ups alert customers to PC infections. *CNet* (Oct. 2009). `http://news.cnet.com/8301-27080_3-10370996-245.html`.
23. MORI, T., ESQUIVEL, H., AKELLA, A., SHIMODA, A., AND GOTO, S. Understanding large-scale spamming botnets from internet edge sites. In *Proceedings of the Conference on E-Mail and Anti-Spam (CEAS)* (Redmond, WA, July 2010).
24. RAMACHANDRAN, A., AND FEAMSTER, N. Understanding the network-level behavior of spammers. In *Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM 2006)* (Pisa, Italy, Sept. 2006), pp. 291–302.
25. RAMACHANDRAN, A., FEAMSTER, N., AND VEMPALA, S. Filtering spam with behavioral blacklisting. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS 2007)* (Alexandria, VA, Oct. 2007), pp. 342–351.
26. ROWE, B., REEVES, D., AND GALLAHER, M. The role of Internet Service Providers in cyber security, June 2009. Available from the Institute for Homeland Security Solutions.
27. SHIN, Y., GUPTA, M., AND MYERS, S. The Nuts and Bolts of a Forum Spam Automator. In *Proceedings of the 4th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)* (Mar. 2011).
28. STONE-GROSS, B., KRUEGEL, C., ALMEROTH, K., MOSER, A., AND KIRDA, E. FIRE: FInding Rogue nEtworks. In *Proceedings of the 25th Annual Computer Security Applications Conference (ACSAC)* (Honolulu, HI, Dec. 2009), pp. 231–240.
29. TAKAHASHI, Y., AND ISHIBASHI, K. Incentive Mechanism for Prompting ISPs to Implement Outbound Filtering of Unwanted Traffic. In *NetGCOOP 2011 : International conference on NETwork Games, COntrol and OPtimization* (Paris, France, Oct. 2011).
30. TEAM CYMRU RESEARCH NFP. IP to ASN mapping. Available at: `http://www.team-cymru.org/Services/ip-to-asn.html`.
31. THE COOPERATIVE ASSOCIATION FOR INTERNET DATA ANALYSIS. The CAIDA AS relationships dataset. Available at: `http://www.caida.org/data/active/as-relationships/`.
32. VAN EETEN, M., AND BAUER, J. M. Economics of malware: Security decisions, incentives and externalities. STI Working Paper, May 2008.
33. VENKATARAMAN, S., SEN, S., SPATSCHECK, O., HAFFNER, P., AND SONG, D. Exploiting network structure for proactive spam mitigation. In *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium* (Berkeley, CA, USA, 2007), USENIX Association, pp. 11:1–11:18.
34. ZHAO, Y., XIE, Y., YU, F., KE, Q., YU, Y., CHEN, Y., AND GILLUM, E. BotGraph: Large scale spamming botnet detection. In *Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation (NSDI)* (Boston, MA, Apr. 2009), pp. 321–334.

A Appendix



(a). The exclusive customer cone of ASN 4134 (b). The exclusive customer cone of ASN 4837

Fig. 1. The size of each node relates linearly to its spamcount. The color of each node relates to the ratio of its spamipcount to its individual prefix size.



(a.) This graph shows the raw number of spam messages (spamcount) sent by the ASN with the given rank, and the raw number of distinct IP addresses (spamipcount) sending at least one spam message from the ASN with the given rank.

(b.) This graph shows the cumulative number of spam messages (spamcount) sent by all ISPs below the given rank, and number of distinct IP addresses sending at least one spam message (spamipcount) from an ASN below the given rank.

Fig. 2. The ASNs in order by spamcount.