Measuring, Characterizing, and Avoiding Spam Traffic Costs

Osvaldo Fonseca, Elverton Fazzion, Ítalo Cunha, Pedro Las-Casas, Dorgival Guedes, Wagner Meira Jr.  
Department of Computer Science  
Universidade Federal de Minas Gerais, UFMG  
Belo Horizonte, Brazil  
Email: {osvaldo.morais, elverton, cunha, pedro.lascasas, dorgival, meira}@dcc.ufmg.br

Cristine Hoepers, Klaus Steding-Jessen, Marcelo H. P. Chaves  
Brazilian Emergency Responde Team (CERT.br)  
Brazilian Network Information Center (NIC.br)  
Brazil  
Email: {cristine, jessen, mhp}@cert.br

Abstract—Spam messages are often used to propagate malware, to disseminate phishing exploits, and to advertise illegal products. Those messages generate costs for users and network operators, but it is hard to measure how much of their costs are associated with spam traffic, and who actually pays for it. In this work, we provide a method to quantify the transit costs of spam traffic. We identify the routes traversed by spam messages collected at five honeypots. Combining the volume of spam traffic with traceroute measurements and a database of inter-network business relationships, we show that stub networks are systematically subject to high spam traffic costs. Furthermore, we show that some networks profit from spam traffic and might not be interested in filtering spam. Finally, we present a simple but effective algorithm to identify the networks that would benefit in cooperating to filter spam traffic at the origin to reduce transit costs.

Keywords—Electronic mail, Network topology, Measurement techniques.

I. INTRODUCTION

Spam messages accounted for 90% of all e-mail messages and generated approximately 216 TB of traffic per day in 2013 [1]. The war against spammers is fought on multiple fronts. Recently, several proposals have focused on filtering spam at its origin, to prevent spam messages from reaching the AS and reduce network bandwidth consumption [2], [3]. However, in practice, spam is usually treated only at the destination e-mail server, by filtering content just before it is delivered to the end user. Although the volume of traffic created by spam may be small if compared with other sources, such as streaming video, spam is still an important problem for network administrators [4].

An Autonomous System (AS) in the Internet is an entity registered with Internet resource allocation authorities. Each AS operates its own network, with end-hosts, routers, and interconnecting links. To achieve global reachability, networks establish peering relationships to exchange traffic. Inter-AS peering relationships may be paid, e.g., when a regional AS buys transit from a global AS, or settlement-free, when two ASes agree to exchange traffic free of charge. Because of the nature of such peering relationships, sending and receiving spam messages may result in direct costs for ASes that pay for transit. In this work we evaluate the cost of spam traffic at the granularity of individual ASes. We combine the volume of spam messages, the paths they traverse, and inter-AS relationships to estimate the cost of spam traffic. Since inter-AS peering contracts are private, we cannot know exactly how much each byte costs for each AS, and cannot compute the absolute cost of spam traffic. Instead, we consider the relative cost of spam traffic based on inferred AS business relations and the volume of spam traffic.

Our measurement approach (Section II) allows us to understand which ASes pay for and which ASes profit from spam traffic. Using five honeypots deployed in different countries, in three continents, during 31 days, we observed the traffic generated by 133 million spam messages that were delivered to those honeypots. We measured 57,419 routes traversed by spam messages using traceroutes issued from RIPE Atlas daily. We mapped IP addresses observed in traces to the AS that originated the IP prefix and post-processed the resulting AS-level paths to remove Internet Exchange Points [5]. Finally, we estimated traffic costs using CAIDA’s database of AS relationships [6], [7], which tells whether an AS pair has paid peering or settlement-free peering.

Our data shows (Section III) that large, global ASes profit from spam traffic as they exchange traffic with paying customers and with settlement-free peer ASes. Medium, regional ASes often lack a settlement-free peer AS to forward spam messages toward their destination, so at least some of the messages will be forwarded over links with paid relationships. Small, border ASes pay for the entirety of their spam traffic, as they rely on their providers for connectivity. Interestingly, ASes that originate large amounts of spam have more limited connectivity (less peering ASes) than ASes that receive spam, increasing overall spam traffic costs.

Finally, we propose an algorithm (Section IV) to identify pairs of ASes that would mutually benefit and save costs by filtering spam traffic close to its source. Our algorithm uses only information publicly
available to ASes and could be executed by the ASes or as a service for them. Our methodology applies not only to spam, but also to other sources of unsolicited traffic, e.g., high-bandwidth DDoS attacks.

Our evaluation shows that filtering can significantly reduce spam traffic costs, but only when an AS uses our algorithm to identify the few other ASes that could also benefit from such filtering and therefore would be willing to act on such traffic. Our characterization indicates that global initiatives against spam may waste their efforts on ASes that are not ultimately interested in filtering spam traffic. Our contributions are applicable today and may be useful in existing spam filtering services.

II. DATA SETS AND METHODOLOGY

We collect spam messages, measure routes that spam messages traverse, identify Autonomous Systems (ASes) on each route, and infer spam traffic costs based on AS peering relationships. We now describe how we collect and combine our data sets. We report statistics and results on data collected between Sep. 8th and Oct. 8th, 2015. Figure 1 illustrates our measurement infrastructure.

A. Spam messages and global spam traffic

We collect spam messages from five honeypots, machines that pass as e-mail open relays and proxies [8]. As honeypots are never publicly announced, we assume (and manual inspection indicates) the only e-mail messages they receive are from spammers that scan for e-mail open relays and proxies. Our honeypots never forward the received spam messages, except for messages whose content indicate they are test messages that spammers send to verify whether open relays and proxies work. Our five honeypots are hosted at educational and commercial networks in five different countries in Europe, North America, and South America.

We log all spam messages at each honeypot, and collect all messages to a central server daily. In the analyzed period, we collected 133 million spam messages from 56,051 IP addresses in 879 distinct ASes registered in 115 countries. We find most IP addresses (82.02%) sending spam use the honeypots as open relays and send few messages (26.75% of the total), behavior previously observed in stealthy botnets. The remaining IP addresses (17.98%) use the honeypots as proxies to send a large number of messages (73.25% of the total), behavior consistent with that of dedicated spam servers.

Figure 1 shows a dedicated spam server in AS9 (bottom center) using our honeypot in AS7 as a proxy to send an e-mail message to a recipient whose e-mail domain is in AS10. The red curve shows the path that the spam message would take if it were forwarded by our honeypot.

Although our honeypots do not forward spam messages, we do consider the traffic that would be generated if the messages were sent. This outgoing traffic would be 3.08 times larger than incoming traffic as spam messages have recipients in multiple domains. Although our data amounts to a small fraction of spam traffic in the Internet and are collected at proxies and open relays, we believe our conclusions generalize to global spam traffic disseminated through these mechanisms. When reporting spam traffic volumes below, we also parenthesize a rough estimate of what our observations would amount to when scaled to global spam volume. In particular, we multiply the spam volume we observe by 6,250, the ratio between Symantec’s estimation of global spam volume [1] and spam volume in our data. Our goal is not to provide accurate global traffic volumes, just approximate the order of magnitude of what ASes are likely to observe in the Internet. For better coverage of global spam traffic, our methodology can be applied to other spam datasets.

B. Measuring Internet routes traversed by spam messages

We estimate the routes traversed by spam messages with traceroute measurements. We issue traceroutes from RIPE Atlas vantage points in the same ASes as our honeypots toward each destination domain observed in spam messages. We do not issue traceroutes from the honeypots directly to preserve their identity. RIPE Atlas is a distributed measurement platform with more than 8,000 vantage points and provides vantage points in all ASes that host our honeypots.

To measure routes from spammers to honeypots, we first identify the AS that hosts each IP address that sent spam to one of our honeypots. We then identify RIPE Atlas vantage points on these ASes and issue traceroutes from these vantage points. Again, to preserve the identity of the honeypots, we do not issue traceroutes to the honeypots directly; we instead issue traceroutes to a RIPE Atlas vantage point (or its first reachable IP hop) in the same AS as the honeypot. RIPE Atlas provides good (40.05%) coverage of the 879 ASes from which we received spam messages, allowing us to measure routes for most (92.26%) spam messages our honeypots receive. This avoid uncertainties due to violations of destination-based routing or asymmetric routes.

Considering the infrastructure depicted in Figure 1, we issue traceroutes from the RIPE Atlas vantage point in AS9 to the RIPE Atlas vantage point in AS7, then from the RIPE Atlas node in AS7 to the recipient’s e-mail server. The red curve shows the measured routes.

One limitation is that RIPE Atlas enforces strict rate limits on measurements, which prevents us from measuring all routes from spammers to our honeypots and from honeypots to all destinations. We operate with a budget of 500 traceroutes per honeypot per day. To maximize the usefulness of our
measured budget, we issue traceroutes to each honeypot from RIPE Atlas vantage points located in the 50 ASes that send the most spam to that honeypot and are covered by RIPE Atlas. From each honeypot, we also pick the 450 destinations that receive the most spam from that honeypot. We recompute the set of ASes and destination domains that send and receive the most spam daily based on the spam we collected on the previous day.

As the distributions of the number of spam messages from each AS and to each destination are heavily skewed, our daily budget still allows for expressive message coverage. Over the analyzed period, we can cover 92.25% of messages from spammers to honeypots on average, with insignificant variation as the set of spammers is stable over the one-month period. We can also cover 69.53% of messages from honeypots to destinations, with a standard deviation of 24.69% across different days as spam campaigns and the set of destination domains change throughout the month.

C. Mapping traceroutes to AS-level paths

We map IP-level traceroute measurements to AS-level paths using IP-to-AS mapping data from iPlane [9], which maps an IP prefix to the set of ASes that originate the prefix. If a router is unresponsive or an IP address is not mapped to any AS but is surrounded by responsive routers with IP addresses that map to the same AS, e.g., \( \{ \ldots, AS_1, x, AS_1, \ldots \} \), we map the IP address to \( AS_1 \).

If a router is unresponsive or an IP address is not mapped to any AS but is surrounded by responsive routers with IP addresses that map to different ASes, e.g., \( \{ \ldots, AS_1, x, AS_2, \ldots \} \), we assume traffic flows from \( AS_1 \) to \( AS_2 \). This happens on 4.59% of paths. This assumption may impact the completeness of our results (when \( x \) is not in \( AS_1 \) or \( AS_2 \)), but it never impacts the correctness of our results as data flows, even if indirectly, between \( AS_1 \) and \( AS_2 \). If a traceroute does not reach the destination, we consider the route up to the last measured hop.

In the example shown in Figure 1, we convert the traceroute measurement from the RIPE Atlas vantage point in AS7 to the AS-path \([AS_7, IXP_1, AS_3, AS_8, AS_10]\). If all routers in \( AS_8 \) are unresponsive or have unmapped IP addresses, we would obtain \([\ldots, AS_3, AS_10]\) instead.

We use a BGP routing table dump obtained from one of the ASes that host one of the honeypots to verify the correctness of our mapping method. We compared the AS-paths obtained with the mapping process above for routes from one of our honeypots and with the (true) AS-paths in the BGP routing table used by the AS hosting the honeypot. We found that 89.16% of AS-paths from IP-to-AS mapping were identical to the BGP AS-paths and that 99.39% had at most one different AS. Although IP-to-AS mapping may have errors, isolated wrong mappings do not impact our ability to identify the overall flow of money.

D. Computing spam traffic costs

We identify which ASes pay for or profit from spam traffic using CAIDA’s inter-AS peering relationship database [6], [7]. CAIDA AS relationships database is known to have errors, but is the most accurate and complete database on AS relationships we are aware of. CAIDA’s database classifies inter-AS peering relationships as either customer-to-provider or peer-to-peer. We assume customers pay providers to achieve global connectivity, and that peers exchange traffic free of charge. These assumptions are common in the literature [6].

CAIDA’s database does not list inter-AS relationships between transit ASes and Internet exchange
points (IXPs). Reasons include that IXPs do not provide transit themselves and that IXPs only provide connectivity between ASes. In this work, we are interested in the peering relationship between ASes that exchange traffic at IXPs. We built a list of IP prefixes and AS numbers used by Internet exchange points (IXPs) combining data from PeeringDB and previous work [6], [10]. We remove these IP prefixes from traceroutes and AS numbers from AS-paths. We have also identified ten peerings with content providers (Google, Amazon, and Microsoft) that were not present in CAIDA’s relationship database. These peerings appear in 2.30% of AS-level paths (8.14% of messages). We have manually labelled these peerings as settlement-free peer-to-peer relationships. With these modifications, CAIDA’s relationship database contains 97.85% of all AS relationships in our AS-level paths and can resolve all relationships in 93.73% of AS-level paths. Our modifications allow us to compute all AS relationships traversed by 99.33% of the messages for which we measured routes.

In the example topology in Figure 1, we place providers above customers and show provider-to-customer relationships with solid lines. We show settlement-free peering relationships with dashed horizontal lines. We remove IXP to AS paths before computing traffic costs. Finally, payments would flow from AS to AS to AS, and from AS to AS to AS. We show content provider networks peering with Verizon and AS.

III. SPAM TRAFFIC COSTS IN THE INTERNET

We now present our findings on spam traffic costs and provide representative examples.

We estimate spam traffic costs for each AS. As peering contracts between different ASes are private, we cannot know how much transferring each byte costs each AS. We are conservative and make no assumptions on the cost of traffic. We instead estimate spam traffic costs as the spam traffic volume exchanged with providers and customers. ASes profit from spam traffic exchanged with customers, and pay for spam traffic exchanged with providers. We study the net spam traffic for each AS, the volume of traffic exchanged with customers minus the volume of traffic exchanged with providers. A positive net spam traffic means the AS profits from spam traffic, and vice-versa. Although we find spam traffic is often unbalanced across settlement-free peering links, we consider traffic on these links does not incur costs and ignore it when computing net spam traffic.

Figure 2 shows the median and quartiles (boxes), 5th and 95th percentiles (dashed lines), as well as outliers of net spam traffic for ASes in our data. ASes with positive net spam traffic are in the ‘profit’ column, and vice-versa. Note the logarithmic scale on the y axis. We classify ASes by their customer cone sizes, the number of other ASes they can reach without using a provider (i.e., free of charge) [6]. We find most ASes pay for or profit from a small volume of spam traffic. To focus on ASes more significantly impacted by spam traffic, Figure 2 does not include ASes with average net spam traffic rate between ±16 Kbps (±100 Mbps after multiplying by 6250 to scale to global spam volumes [1]).

Large ASes, with customer cones including more than 100 ASes, rarely pay for and often profit from spam traffic. This is because these ASes profit from exchanging spam traffic with their customers, and forward spam traffic through peering ASes free of charge. Level3 (AS3356) transits 28.17% of the spam messages in our data, and profits from a net spam traffic rate of 2.77 Mbps (17.36 Gbps). Notice that large ASes profit twice from spam traffic whenever they receive a spam message from a customer and forward it to another customer.

Medium ASes, with customer cones including tens of ASes, rarely pay for spam traffic and often pay for spam traffic. Figure 2 shows one medium AS with a profit, and all other medium ASes with losses. Medium ASes pay for spam traffic in two situations:

1) Medium ASes pay a provider for outgoing spam traffic whenever they cannot forward it to a peering ASes free of charge. This amounts to 49.37% of total spam traffic costs, on average, for medium ASes in our data. We find forwarding traffic through peering ASes free of charge reduces spam costs by traffic by 34.68%, on average, for ASes in our data.

2) Medium ASes pay for incoming spam traffic received from their providers. This accounts for the remaining 50.73% of total costs for medium ASes in our data. We note that ASes that originate spam have worse connectivity than ASes that receive spam (their average number of provider plus peer ASes is 1.38 and 4.31, respectively). This causes spam traffic to go high in the Internet’s transit hierarchy and incurs costs on all ASes on the downstream portion of the path.

Medium ASes profit from exchanging spam traffic with their customers. However, the costs above may result in negative net spam traffic when medium ASes originate or receive spam, as originated and received spam traffic is not exchanged with any customer and do not generate revenue.

The only medium AS we observe with a profit is HINET (AS9680), customer cone with 24 ASes), which profits from receiving an average 689 Kbps (4.3 Gbps) of outgoing spam traffic from its customers and forwards 80.44% of it through peering ASes. Further investigation shows all of HINET’s spam comes from its Taiwan-based subsidiary (AS3462). HINET has
been reported before as a spam haven, and our results explains why it remains so: HINET can charge spammers and forward most spam for free.

**Small and stub ASes,** with customer cones smaller than ten ASes, pay for the bulk of spam traffic costs. These ASes originate and receive spam traffic, incurring losses from exchanging spam traffic with their providers. Although the three outliers are ASes hosting our honeypots (customer cone sizes of 1, 3, and 9), we observe that 25% of ASes pay for spam traffic rates higher than 237 Kbps (1.48 Gbps).

IV. **Filtering Spam Traffic**

In this section we use our findings to propose an algorithm to identify ASes interested in filtering spam traffic. We note that some ASes profit from spam traffic and would not be interested in spending human and computational resources to filter spam traffic. We also note that even ASes that pay for spam traffic would not be interested in filtering spam traffic on the downstream path, i.e., when they would forward spam traffic to one of their clients. For example, in Figure 1, AS7 would be willing to filter spam coming from AS9 and destined to AS10, on behalf of AS10, since it would receive payment for this traffic from AS9 but avoid its own payment to AS3. However, AS7 would be unwilling to filter spam from AS10 to AS9 on behalf of AS9 (after having already incurred the payment to AS3 for delivering this traffic) for less than the profit from delivering this traffic to AS10. A practical solution would need to filter spam traffic at or close to its originating AS.

Algorithm 1 summarizes our proposal for identifying possible spam traffic filtering partnerships. Each AS \( x \) in the Internet can collect spam messages it receives from a provider, i.e., that it pays for. AS \( x \) can then use our methodology to identify which other ASes pay to send \( x \)'s spam messages to a provider and would profit from filtering these messages. AS \( x \) can then contact these ASes in an attempt to establish spam traffic filtering agreements. Similarly, each AS \( x \) can collect spam messages it forwards to a provider, then use our methodology to identify which ASes hosting the destination domains pay to receive the messages. AS \( x \) can then contact these ASes and offer to filter spam traffic destined to them. Filtering all traffic from a source (spammer) to a destination (SMTP server) is possible at the packet level based on IP addresses. Selective filtering of spam messages among legitimate messages requires more complex (e.g., stateful) processing at the intermediate AS.

The solid blue curve in Figure 3 shows the distribution of potential savings in spam traffic costs for all medium and small ASes in our data. We compute potential savings for each AS as the fraction of spam traffic exchanged with providers that can be filtered. More precisely, we compute potential savings for an AS \( x \) as the ratio of \( x \)'s spam traffic costs when all ASes agree to filter all messages that incur costs by \( x \)'s total spam traffic costs without filtering. The curve shows that filtering, even when driven by profit, can mitigate all spam traffic costs for 60.14% of ASes and reduce spam traffic costs by at least half for 88.66% of ASes.

The red dashed curve in Figure 3 shows the distribution of spam traffic cost savings when each AS establishes filtering agreements with ten ASes, as computed by our methodology. We consider each AS establishes agreements with the ten ASes that lead to the highest reduction in spam traffic costs, as computed by our methodology. After choosing an AS to establish an agreement, we recompute spam
Algorithm 1 Choose partners for filtering spam

**Input:** Spam messages, AS-level route measurements, inter-AS business relationships  
**Output:** $C[x][y]$: amount of traffic $x$ and $y$ can filter to reduce spam traffic costs

1: for each AS $x$ do  
2:  $S_{in} \leftarrow$ set of spam messages destined to $x$ received from providers  
3:  for each AS $y$ on the paths traversed by messages in $S_{in}$ do  
4:    if $y$ reaches $x$ through a provider then  
5:      $C[x][y] \leftarrow$ volume of spam traffic between $x$ and $y$  
6:  $S_{out} \leftarrow$ set of spam messages through $x$ sent to providers  
7:  for each AS $y$ hosting destination domains for messages in $S_{out}$ do  
8:    if $y$ receives messages through $x$ from a provider then  
9:      $C[x][y] \leftarrow C[x][y] +$ volume of spam traffic between $x$ and $y$

traffic costs, i.e., re-run Algorithm 1, before picking the next. We observe that cooperating with as few as ten ASes is enough to achieve significant savings when ASes are chosen intelligently. Cooperating with the top ten ASes mitigates all spam traffic costs for 37.00% of ASes and reduces spam traffic costs by at least half for 77.06% of ASes.

Finally, the dotted black curve close to the $y$ axis in Figure 3 shows the distribution of expected savings when AS $x$ establishes filtering agreements with 15 ASes at random. We approximate expected savings as the average savings over 100 random samples of 15 filtering agreements. The curve shows that picking partnerships at random does not allow for significant savings. This happens because most (90.28%) ASes either forward no spam for $x$ or are uninterested in filtering $x$’s spam traffic.

V. RELATED WORK

Our data collection methodology draws heavily on previous work on topology measurements [9], converting traceroute measurements to AS-paths [5], and identifying inter-AS business relationships [6], [7]. Our measurements using RIPE Atlas closely follow routes traversed by spam messages, avoiding errors due to commonplace asymmetric routing and rare violations of per-destination routing.

Researchers have studied the properties of spam traffic for more than a decade [11]. More related to our work, researchers have studied the cost of the spam sending infrastructure (servers and botnets) [12] and how to filter spam at the origin AS [2]. Researchers have also studied how to identify suspicious ASes in the Internet based on their networking practices [13]. Our work is complementary and advances the state of the art by investigating the cost of spam traffic at the granularity of ASes, giving insight on incurred costs and how to mitigate it.

VI. CONCLUSIONS

We have presented a measurement methodology to estimate the cost of spam traffic at the granularity of Autonomous Systems (ASes). We have used our data to characterize the cost of spam traffic in the Internet today; we found that large ASes profit from spam traffic while medium and small ASes pay. We used this insight to propose an algorithm to identify pairs of ASes that might benefit from cooperating to filter spam traffic close to the sources. Our algorithm uses only data that is readily available and could be run by any AS or be provided as a service. Our evaluation shows that our algorithm can help ASes find other ASes that could benefit and therefore might be willing to cooperate in filtering spam. Results show that it performs significantly better than choosing ASes to cooperate with at random. Our contributions are applicable in spam filtering efforts and may ultimately lead to reduced spam traffic in the Internet.

ACKNOWLEDGEMENTS

This work was partially funded by NIC.BR, Fapemig, CNPq, CAPES, and by projects InWeb (MCT/CNPq 573871/2008-6), MASWeb (FAPEMIG-PRONEX APQ-01400-14), and EUBra-BIGSEA (H2020-EU.2.1.1 690116, Brazil/MCTI/RNP GA-000650/04).

REFERENCES


