Gaining Insight into AS-level Outages through Analysis of Internet Background Radiation

Karyn Benson, Alberto Dainotti, KC Claffy
CAIDA, UC San Diego
Email: {karyn, alberto, kc}@caida.org

Emile Aben
RIPE NCC
Email: emile.aben@ripe.net

***Abstract*—Internet Background Radiation (IBR) is unsolicited network traffic mostly generated by malicious software, e.g., worms, scans. In previous work, we extracted a signal from IBR traffic arriving at a large (/8) segment of unassigned IPv4 address space to identify large-scale disruptions of connectivity at an Autonomous System (AS) granularity, and used our technique to study episodes of government censorship and natural disasters [1]. Here we explore other IBR-derived metrics that may provide insights into the causes of macroscopic connectivity disruptions. We propose metrics indicating packet loss (e.g., due to link congestion) along a path from a specific AS to our observation point. We use three case studies to illustrate how our metrics can help identify packet loss characteristics of an outage. These metrics could be used in the diagnostic component of a semi-automated system for detecting and characterizing large-scale outages.**

I. Introduction

Internet Background Radiation (IBR) is a mix of unsolicited network traffic mostly generated by malicious activity, e.g., worms, scans, most easily observed by instrumenting large segments of unassigned address space (or "darknets") [2], [3]. Researchers have analyzed IBR traffic to detect and analyze worm spreading [4], [5] and denial-of-service attacks [6]. Casado et al. [7] proposed techniques to use IBR traffic to opportunistically measure network properties unrelated to malware, e.g., local link bandwidth, host uptime, NAT usage. Dainotti et al. [1], [8] also applied the opportunistic measurement idea to IBR traffic, extracting signals that revealed interesting dynamics of large-scale connectivity disruptions, e.g., caused by country-level censorship or large earthquakes. IP addresses from disconnected networks will not send IBR traffic, and mapping such addresses to their geolocation enables an estimation of the geographic impact of a natural disaster on communications infrastructure [1]. Comparing IBR traffic levels to BGP observations allows one to distinguish between control plane and data plane (e.g., packet filtering) disruptions, as when governments experiment with different approaches to Internet censorship [8].

We investigate new IBR-derived metrics that can provide insights into the causes of macroscopic connectivity disruptions. These metrics can indicate whether an outage involves packet loss, e.g. due to link congestion. The vast majority of IBR traffic is composed of TCP SYNs probing the Internet trying to establish connections to vulnerable (usually Windows) hosts. Because a darknet is completely passive (it does not respond to any packets), sources sending these SYNs must re-transmit them. TCP retransmit behavior (such as how many retransmits per connection attempt, and how much time between them) is typically a function of the host operating system or application, which means it is consistent across large enough populations of hosts to constitute a reliable signal. We derive two metrics from two different dimensions of this signal: the number of SYN retransmit per TCP flow; and the distribution of inter-packet times (IPT) between them. We show that both metrics can reflect packet loss, providing additional insight compared to metrics that only indicate reachability. We apply this metric to three case studies where either route leaks caused link congestion for an entire AS (and ultimately a complete outage in one case) or packet filtering imposed by a regime almost entirely isolated a country from the rest of the Internet.

Traditional passive approaches to inferring packet loss use attributes such as the congestion window [9], RTT [9], [10], and TCP acknowledgments [11] – all of which require bidirectional communication. In contrast, we: (i) observe unidirectional traffic from a darknet, and (ii) use retransmit packets as opportunistic probes that measure large-scale Internet events. We are not aware of similar studies and we consider this work a first attempt to investigate this approach.

II. Data source and signal extraction

We analyze IBR traffic captured at the UCSD Network Telescope, a /8 darknet of unassigned IP addresses [12]. A darknet receives but does not respond to traffic, so all flows (defined as the traditional 5-tuple) are unidirectional. When an external host attempts to open a TCP connection, the resulting flow carries only SYN retransmits, which we call a *SYN flow*.

To derive IBR metrics that correlate with packet loss, we need attributes that are normally consistent yet change during connectivity disruptions. The ideal signal would be strong (statistically significant), stable (low noise), and globally pervasive (seen in most networks). But IBR includes diverse types of traffic [2], [3], so we selected two subsets of IBR that have consistent and predictable enough behavior to use as signals for packet loss:

- Conficker-like traffic, i.e., SYN flows to TCP port 445, widely publicized during the Conficker episode in 2008 but a target of scanning activity for years; it constitutes a large percentage of the packets observed at the UCSD

Attribute	Extracted Metric	Reference	What significant change we can infer
traffic volume	packets per second	[8]	connectivity disruption
number of sources	θ	[1]	connectivity disruption and its intensity
average number of SYN retransmissions	γ	Section III-A	packet loss (e.g., congestion)
inter-packet times between SYN retransmits	η	Section III-B	packet loss (e.g., congestion)

TABLE I: Attributes and Metrics used to infer outages and their characteristics.

telescope (more than 40%), is globally pervasive, and consistent [13], [14].

- The default configuration of Windows machines is to send at most 3 SYN packets [15] when attempting to establish a connection, which makes SYN flows from such machines a consistent signal.

To infer packet loss, we selected two attributes of SYN flows – number of retransmissions and IPT – that follow consistent patterns. Since the darknet never responds with an ACK, the number and timing of SYN retransmits is determined by the application or the OS originating such traffic. The consistency of these attributes depends on the conditions of the path traversed by the packets, so substantial drops in SYN retransmits or substantial variation in the IPT may reflect network-induced packet loss.

We did a preliminary analysis of SYN flows collected at the UCSD Network Telescope for January 2012[1], during which no large-scale outages occurred. We used Corsaro [16], a publicly available software tool that supports plugins for the analysis and tagging of darknet traffic. We implemented a passive OS fingerprinting plugin, based on *p0f* [17], to analyze our two selected traffic classes.

Operating System	Number of Flows
Windows XP	2299144254
Windows NT	229961989
Windows 7 or 8	53445230
Other (Linux, BSD, Solaris, . . .)	394731

TABLE II: Conficker-like SYN flows observed per OS (Jan. 2012)

Port	OS	%3-Flows	Num 3-Flows	Num /8
80	Windows 7 or 8	0.850	6107763	184
443	Windows 7 or 8	0.775	2656821	170
443	Windows NT	0.828	2602825	169
1433	Windows XP	0.814	39476702	114
3260	Windows 7 or 8	0.987	293572	85
4661	Windows NT	0.984	183551	97
4899	Windows 7 or 8	0.993	16108965	73
28931	Windows 7 or 8	0.984	25961	71
22292	Windows XP	0.804	11433398	174

TABLE III: The top four OS-port combination flows of the following categories: at least 25,000 3-packet SYN flows; originating from 64 or more /8 networks; 3-packet SYN flows comprise more than 75% of the SYN flows from the specified OS and port. There are 9 listed in the table because of overlap in the top-four lists. (Jan. 2012)

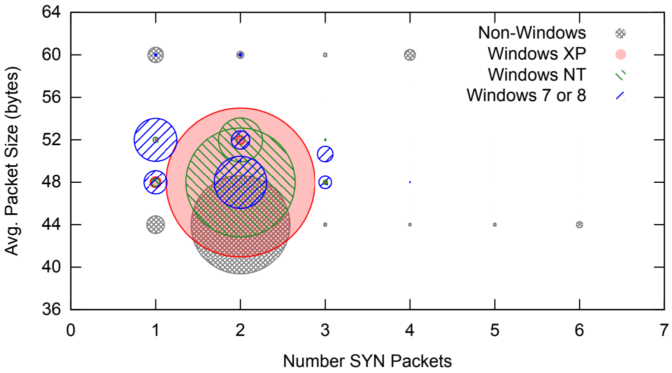

Fig. 1: Packet size vs. number of packets per flow by OS (Jan. 2012) for Conficker-like traffic. The radius of each circle is proportional to the percentage of SYN flows from the given OS having that packet size and number of packets. Windows 7 or 8 and Non-Windows hosts send a variety of packets per flow and packet sizes - making it hard to extract a stable signal. Conversely, Windows XP and NT consistently send flows with 2 packets of size 48 or 52, creating a stable signal.

A. *Conficker-like Traffic*

Figure 1 shows the distribution of SYN flows destined to TCP port 445 as a function of packet size, number of retransmits, and OS. Most of these SYN flows contain only two SYN packets, consistent with the behavior of Conficker-infected hosts [14]. To obtain a strong and stable signal for a retransmit-based metric, we tried to isolate such behavior (i.e., 2-packet SYN flows) by selecting only flows from Windows XP and Windows NT (about 89% and 9% of the total flows in Table II) with packet sizes of either 48 or 52 bytes. The IPT metric is not usable with the Conficker-like traffic since the flows only have two packets; loss of one of them prevents a valid IPT calculation.

B. *Default Windows Behavior*

To build a second signal usable for packet loss inference, from IBR observed at the UCSD Network Telescope in January 2012 we selected the port-OS combinations satisfying all of the following criteria:

- more than 75% of their SYN flows carry 3 packets (aiming at a stable signal);
- more than 25000 3-packet SYN flows (strong signal);
- their 3-packet SYN flows originate from more than 25% of the total number of /8 IPv4 networks (globally pervasive signal).

We selected 100 port-OS pairs that met these criteria, including 56 “Windows 7 or 8” ports, 29 “Windows XP”, and 15 “Windows NT”. Table III lists the top four port-OS pairs for each separate criterion.

Although the total number of 3-SYN flows that we select is two orders of magnitude smaller than Conficker-like flows (about 156M vs. 13B in January 2012) and the number of sources generating 3-pkt SYN flows is smaller by a factor of 7 (an average of 14K hosts/hour compared to 100K Conficker hosts/hour in January), having a second traffic signal is still useful, especially to validate findings. Also, the 3-SYN flows

[1] The telescope was down for about 40 hours starting on 2012-01-14 and for about 120 hours starting on 2012-01-19.

metrics are not malware-specific, which is especially important as machines are upgraded and patched, limiting the spread of the Conficker-like traffic. The 3-SYN flows are also amenable to IPT calculations when one of the packets is lost, unlike the Conficker-like (mostly two-packet) flows.

III. Definition of Metrics

A. Number of Packets per SYN Flow: γ

We first considered simply the average number of packets per SYN flow from the selected traffic (either Conficker-like or 3-pkt SYN flows). When considering flows sent by only a subset of source IPs, such as the AS-level interpretation (that is, computing such metric only for IBR originating from a specific AS), this value could be significantly skewed as a result of the increased influence of a single host or flow. For example, when a single host conducts a horizontal scan by sending one packet to every IP address in the darknet, the AS-level average is approximately 1 packet per flow regardless of other host activity from that AS. Similarly, a single flow consisting of a large number of SYN packets significantly increases the overall average. The following improvements reduce the impact of such anomalies:

- we exclude all the SYN flows with more than a given number of packets: *three* for Conficker-like SYN flows (97% of SYN flows had three or fewer packets in our reference dataset of January 2012); *four* for the Default Windows SYN flows;
- we calculate the average number of packets per SYN flow for each distinct source IP, and then take the average (mean) of this distribution, thus limiting the influence of a single source IP sending packets to the darknet.

If the set of all source IPs is S, F_s denotes the set of flows matching our criteria with source IP s, and the function $\mathsf{packets}(f)$ returns the number of packets in a flow f then our metric is

$$\gamma = \frac{1}{|S|}\sum_{s\in S}\frac{\sum_{f\in F_s}\mathsf{packets}(f)}{|F_s|} \tag{1}$$

If there are no sources matching our criteria, then γ is undefined. We call the metric γ_C for the Conficker-like traffic and γ_3 for the flows that are expected to have three packets per SYN flow. We do not combine the two metrics γ_C and γ_3, as the ratio of hosts contributing to each metric is not constant.

Figure 2 shows this metric across all ASes for January 2012, calculated in hourly bins. The number of source IPs and γ approximately follow a sinusoidal pattern with a phase of one day. The value of γ_C is always between 1.98 and 2.02. The value of γ_3 is always between 2.59 and 2.78. The large drop in γ_3 seems to be related to traffic on BitTorrent and HTTPS ports.

Outages are likely to affect only a subset of the Internet hosts. Grouping by AS provides a natural way to divide the IP address space. We used CAIDA's Prefix to AS Mapping Dataset and RouteViews BGP data [18]. Figure 3 shows γ_C calculated for three ASes of different size. As expected, when calculating γ_C for a single AS, there is higher variance for ASes with fewer infected hosts, typically proportional to their size. Increasing the size of the time bins would reduce such measurement variance, but at the expense of precision in when a connectivity disruption occurred.

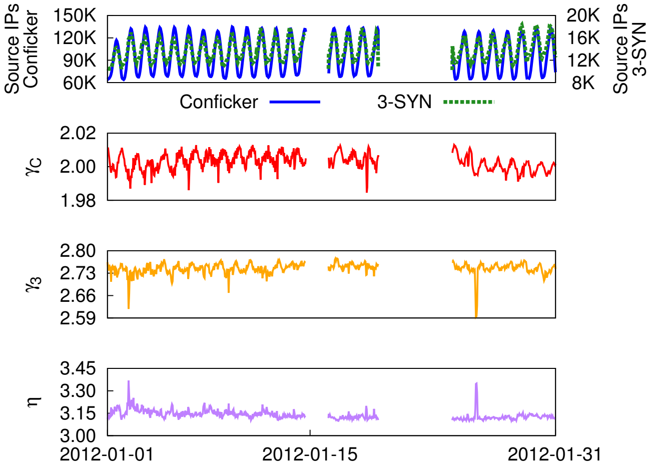

Fig. 2: Number of source IPs and new metrics for all ASes (Jan. 2012) with time bins of 1 hour. The number of source IPs sending Conficker-like and Default-Windows traffic is significant and follows a sinusoidal pattern. γ_C is always between 1.98 and 2.02; γ_3 is always between 2.59 and 2.78; η is always between 3.09 and 3.37. Under normal circumstances, each metric has a small range - which is necessary to identify deviations associated with outages. The telescope was down for about 40 hours starting on 2012-01-14 and for about 120 hours starting on 2012-01-19.

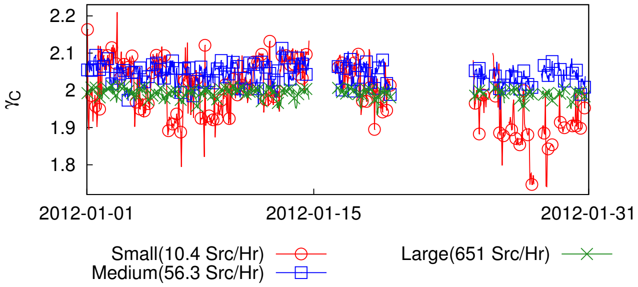

Fig. 3: AS-level γ_C and number of source IPs for three ASes with different number of infected hosts (Jan. 2012, hourly bins). When there are more infected hosts γ_C is more consistent – meaning it is more useful for discerning abnormalities.

B. Inter-Packet Times: η

Hosts following RFC 6298 [19] should wait at least one second before retransmitting the initial SYN packet; subsequent retransmission timeouts (RTOs) should back off exponentially by a factor of two. The convention is to use 3 seconds as the initial RTO (i.e., the RTOs are normally 3, 6, 12, 24, . . . seconds).[2] The average of the first IPT can be used to verify the findings of the γ metric. In flows with three packets, if a

[2] Although RFC 6298, states that the RTO should be 1 second, we observe in the darknet that the RTO is still ∼3 seconds for more than 99% of flows from Windows hosts. If an RTO of 1 second is more widely adopted, we can identify the RTO typically used by each source and normalize the metric.

single packet is lost, the first IPT is either (approximately) 6, 9, or 3 seconds corresponding to loss of the first, second, or third packet, respectively. We can only calculate η on expected 3-pkt SYN flows.

As in the calculation of γ, we consider the possibility of skew from a few deviant hosts. If F_s denotes the set of flows with source IP s that are expected to have 3 packets and have at least 2 packets, $S = \{s \in \text{seen source IPs} | F_s \neq \emptyset\}$, and the function $\mathsf{IPT}(f)$ returns the first IPT of a flow f, then our metric is

$$\eta = \frac{1}{|S|} \sum_{s \in S} \frac{\sum_{f \in F_s} \mathsf{IPT}(f)}{|F_s|} \qquad (2)$$

If $|S| = 0$, then η is undefined.

η is a less precise metric than γ, since it uses fewer flows during connectivity disruptions, allowing it to skew more easily. However, the combination of η and γ allows for strong inference. A decrease in γ may also mean that fewer packets than expected are actually being sent for a traffic class instead of being lost along the path, but η can help us distinguish between the two cases (i.e., assuming RFC-compliant behavior, η can distinguish between sending only two packets and a random loss of one of three packets). Figure 2 shows η calculated across all ASes for January 2012.

IV. Case Studies

In this section, we evaluate our metric using three different service-disruption case studies. The first two outages – the "Dodo-Telstra" and the "Bell-Dery" case – had network-induced packet loss as a result of BGP route leaks [20], [21]. The third one – the Libyan Internet blackout – was the result of packet filtering. If effective, our metrics will reflect packet loss in the first two case studies but not in the last.

For each of the case studies, we only use metrics which were stable throughout the entire month preceding the outage.

A. "Dodo-Telstra" Routing Leakage

On February 23, 2012, around 2:40 UTC, the multi-homed network operator *Dodo* announced internal BGP routes to its provider *Telstra*, a major ISP in Australia, which erroneously accepted them. As a result, Telstra sent all of its traffic to the small network, Dodo, instead of a large transit provider, inducing a bottleneck leading to a complete outage [20]. The effect was massive: most Australians were left without Internet connectivity for about half an hour [22].

Figure 4 plots our metrics for IBR traffic originating from AS1221 (Telstra) calculated in 5-minute bins. The figure shows significant drops of both γ_C and γ_3 during the first phase (20 minutes) of the episode, meaning that far fewer packets per flow were reaching the darknet than normal. However, when γ_C and γ_3 first drop, η increases from about 3 to 5 seconds, which corroborates packet loss (assuming individual hosts did not change their retransmission patterns). This spike was calculated using the 7 distinct source IPs observed from this region at the darknet. In the following three 5-minute time bins the number of sources (0, 1, 2 respectively) contributing to the calculation of η was not statistically significant. Such a significant drop in γ_C and γ_3 and the increase in η are a direct consequence of congestion on the affected links. Routers started dropping packets, including some of the SYN packets destined to the darknet. Eventually, this congestion deteriorated to a complete outage (lasting another 20 minutes), during which the telescope did not observe any sources sending SYN packets from Telstra (so our metrics cannot be calculated).

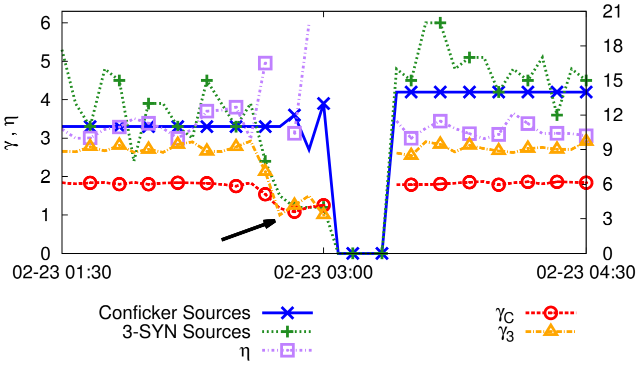

Fig. 4: Our packet-loss metrics plotted in 5-min. bins for traffic originating from AS1221 during the Dodo-Telstra routing leak in February 2012. The arrow points at the first phase (20 minutes) of the outage, where the metric values indicate a bottleneck, i.e., packet loss: γ_C and γ_3 decreased, and η increased. The number of IPs sending Conficker traffic remained the same, while the number of IPs sending 3-SYN flows decreased – an artifact of the frequency at which each type of host contacts the darknet. In the second phase, no flows were observed in the darknet traffic, implying a complete outage.

B. "Bell-Dery" Routing Leakage

On August 8, 2012, at 17:27 UTC, dual-homed provider *Dery Telecom* started to leak a full BGP table to the major Canadian ISP *Bell*. These routes were accepted and propagated to Bell's peers [21]. Our analysis shows that the biggest disruption lasted about half an hour.

Figure 5 plots our metrics calculated for traffic coming from AS577 (Bell) surrounding the outage. The Bell network never was completely offline, but the plot indicates a severe disruption (∼17:30-17:45) followed by slight improvement (∼17:45-17:55) before restoration. During this time period, the total number of Conficker and 3-SYN source IPs dropped from about 12 and 20 to 2 and 6, respectively. Both γ_3 and η indicate significant packet loss during this time period. Strangely, γ_C stayed close to 2 during the worst part of the disruption, decreasing slightly when conditions appeared to improve (number of Conficker sources rose from 2 to 11).

To determine the reason behind the differences in γ_C and the other two metrics, we broke down the traffic from AS577 by network prefix and inspected the TTL header fields in the collected packets. In the 90 minutes surrounding the outage, packets from AS577 originated from 63 distinct /16 prefixes, of which 38 sent traffic in at least 9 of 18 5-minute time bins, and all but one experienced a considerable volume drop. Upon further inspection, two IP addresses in this prefix continued to

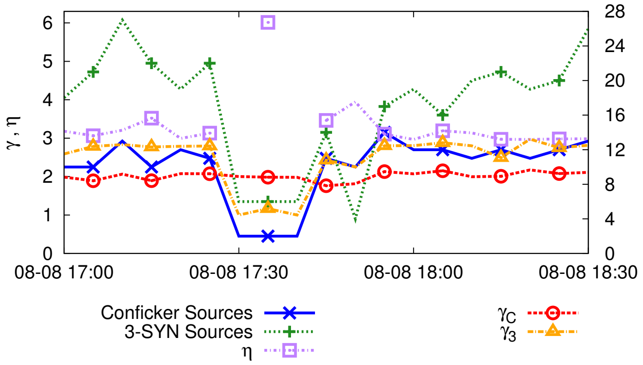

Fig. 5: During the Bell-Dery routing leak of August 2012, we observed traffic from AS577 during every 5-minute bin. The number of Conficker and 3-SYN source IPs dropped drastically. Two of our metrics, γ_3 and η indicated packet loss, but the γ_C metric did not, which we later discovered was because one network was unaffected by the BGP leak.

transmit Conficker-like traffic at their pre-outage rate, depicted in Figure 6.

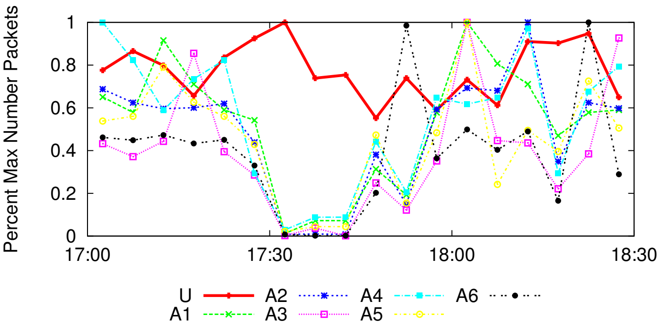

Fig. 6: Looking at SYN traffic volume by prefix reveals that one prefix (U) did not experience loss during the outage. Each point represents, for the given prefix, the fraction of packets sent during a 5-minute bin normalized to the time bin with the most packets. We show only prefixes observably active during all 18 time bins.

Since the Bell-Dery event was caused by a route leak, it is possible to observe changes in the way packets were routed by looking at the TTL value, reflecting a different number of hops in the path to the telescope. We discovered that the only two IP addresses whose packet rate at the telescope was not affected by the disruption were also the only two IP addresses whose packets carried a constant TTL both outside and during the disruption (one such IP address depicted in top graph of Figure 7). We suspect that traffic from this prefix was re-routed through a different path that was unaffected by the route leak.

C. Libyan Internet Blackout

Our third case study applies our metrics to the Libyan Internet blackout happened in February and March 2011, when the Libyan government used BGP disconnection and later packet filtering to implement nationwide censorship [8]. There were three outages, lasting approximately 7 hours (the first two) and 3.7 days (the last one). We examine the second one, when the state telecom (AS21003) isolated most of the country through packet filtering [8]. This case study illustrates that our metrics effectively distinguish large-scale outages that are characterized by some packet loss from those that are not.

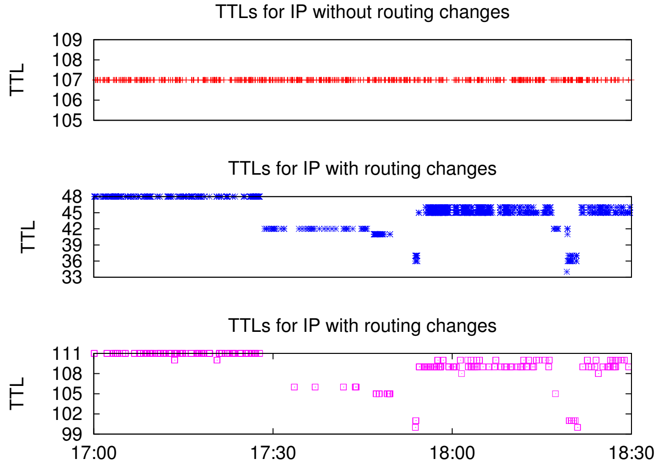

Fig. 7: TTLs of an IP address with and without routing changes. During congestion, most packets reaching the darknet have a lower TTL than before the outage, indicating that they took a longer path. The top graph plots TTLs of packets from an IP address in the unaffected network, whose path to the telescope presumably does not change. Not until approximately 2012-08-08 21:20 (not shown) does the TTL of packets sent by the second and third IP address return to their pre-outage value.

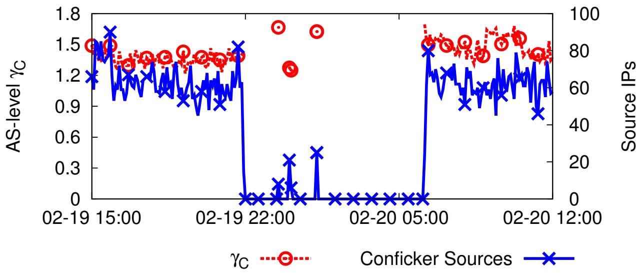

Fig. 8: Libya's second 2011 outage used packet filtering as a method of censorship. Although there were fewer source IPs during the censorship, the hosts that did send Conficker-like flows sent approximately the same number of packets per flow as prior to the outage, indicated by the similar values of γ_C (calculated in 5 minute time bins). γ_3 and η are excluded as there were not enough hosts (2 or less) to accurately make inferences.

Figure 8 shows that when a subset of hosts can communicate through the filtering system, γ_C remains near pre-censorship values, despite fewer sources sending traffic. Thus we can infer that the outage was not caused by an event inducing network packet loss. We excluded γ_3 and η from this measurement, since there were not enough hosts sending 3-packet SYN flows to accurately infer anything from these metrics.

D. Utility of Metrics

In all three case studies, the metrics γ and η provided insight into the nature of the outage. In the "Dodo-Telstra" case study, network congestion preceded the complete outage. In response to congestion, the network dropped packets, decreasing the number of packets per flow, which reduced the values of γ_C and γ_3 and increased η. In the "Bell-Dery" case study, the metrics extracted from the Conficker signal implied network-induced packet loss (e.g congestion). However, γ_C initially painted a different picture of packet loss: sources able to send Conficker-like traffic were unaffected. A deeper exploration of traffic volume by prefix and TTLs revealed that the connectivity disruption was more severe for some subnets than others. This result demonstrated that multiple data classes and metrics can strengthen the quality of inferences and provide a starting point for further investigation. In the Libyan Internet Blackout example, although the traffic volume was smaller, γ_C remained at pre-censorship levels whenever Conficker-like traffic was observed. This behavior is consistent with filtering packets by subnet: the number of traffic sources decreases but per-flow characteristics will not change.

V. Conclusion

To augment the binary signal of presence or absence of traffic flows from a particular network, we explored IBR-derived metrics that help characterize connectivity disruptions that induce packet loss, e.g., link congestion. Our metrics are based on SYN retransmits in unsolicited Internet background radiation, visible from passive darknet instrumentation. Because these retransmits typically follow consistent patterns that are a function of operating system or application implementation, we can infer packet loss if some retransmits are not observed by the darknet.

We used three case studies to demonstrate that our γ and η metrics can distinguish a transit bottleneck-induced outage from an intentional nation-wide disconnection caused by packet filtering. One unexpected finding was that in the Bell-Dery route leak incident, different parts of the affected AS reacted differently to the route leak, confirmed by examination of TTL values on a per-prefix level. This analysis provided hints on how to group parts of a specific AS into finer-grained units that may be affected differently by a disruption.

Our method has several limitations: it only measures packet loss between a given source and our darknet. It also relies on the presence of Conficker-like or IBR TCP traffic in general. But our simple metrics applied to a large darknet traffic segment enable us to continually monitor one aspect of network connectivity (i.e., reachability to our darknet) from all over the world.

Our method is complementary to techniques using active probes to discover outages. Due to the large number of prospective IP addresses to probe, current active techniques do not ensure that the outage will be captured. For example, [23] only monitors /24s with a large number of responsive IPs - which is about 9% of the allocated address space. Alternatively, [24] covers 89% of the Internet's edge address space but the focus is on failures lasting longer than 15 minutes.

In the future we would like to test this metric on other connectivity scenarios and other darknet traffic, explore other IBR-related metrics that can characterize network disruptions, and integrate such metrics into a system for comprehensive detection and diagnosis of such disruptions.

VI. Acknowledgements

This material is based upon work supported by the National Science Foundation under grants CNS-1059439 and CNS-1228994. Support for the UCSD network telescope operations and data collection, curation, analysis, and sharing is provided by DHS cooperative agreement FA8750-12-2-0326, NSF grant CNS-1059439, and UCSD.

References

[1] A. Dainotti, R. Amman, E. Aben, and K. C. Claffy, "Extracting Benefit from Harm: Using Malware Pollution to Analyze the Impact of Political and Geophysical Events on the Internet," *SIGCOMM Comput. Commun. Rev.*, 2012.

[2] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson, "Characteristics of Internet Background Radiation," in *Internet Measurement Conference (IMC 2004)*, 2004.

[3] E. Wustrow, M. Karir, M. Bailey, F. Jahanian, and G. Huston, "Internet Background Radiation Revisited," in *Internet Measurement Conference (IMC 2010)*, 2010.

[4] D. Moore, C. Shannon, and J. Brown, "Code-Red: a case study on the spread and victims of an Internet worm," in *Internet Measurement Workshop (IMW 2002)*, 2002.

[5] C. Shannon and D. Moore, "The Spread of the Witty Worm," *IEEE Security and Privacy*, 2004.

[6] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage, "Inferring Internet Denial-of-Service Activity," *ACM Trans. Comput. Syst.*, 2006.

[7] M. Casado, T. Garfinkel, W. Cui, V. Paxon, and S. Savage, "Opportunistic Measurement: Spurious Network Events as a Light in the Darkness," in *ACM Fourth Workshop on Hot Topics in Networks (HotNets-IV)*, 2005.

[8] A. Dainotti, C. Squarcella, E. Aben, K. C. Claffy, M. Chiesa, M. Russo, and A. Pescapé, "Analysis of Country-wide Internet Outages Caused by Censorship," in *Internet Measurement Conference (IMC 2011)*, 2011.

[9] S. Jaiswal, G. Iannaccone, C. Diot, J. Kurose, and D. Towsley, "Inferring TCP Connection Characteristics Through Passive Measurements," in *23rd Annual Joint Conference of the IEEE Computer and Communication Societies. (INFOCOM 2004)*, 2004.

[10] N. Fonseca and M. Crovella, "Bayesian Packet Loss Detection for TCP," in *24th Annual Joint Conference of the IEEE Computer and Communications Societies. (INFOCOM 2005)*, 2005.

[11] S. Katti, D. Katabi, E. Kohler, and J. Strauss, "M&M: A Passive Toolkit for Measuring, Correlating, and Tracking Path Characteristics," MIT CS and AI Lab, Tech. Rep., 2004.

[12] "UCSD Network Telescope," 2010, www.caida.org/data/passive/network_telescope.xml.

[13] UCSD Network Telescope Global Attack Traffic (current). www.caida.org/data/realtime/telescope/.

[14] E. Aben. (2008) Conficker/Conflicker/Downadup as seen from the UCSD Network Telescope. www.caida.org/research/security/ms08-067/conficker.xml.

[15] "TcpMaxConnectRetransmissions," technet.microsoft.com/en-us/library/cc938209.aspx.

[16] A. King and A. Dainotti, "Corsaro," www.caida.org/tools/measurement/corsaro/, 2012.

[17] M. Zalewski, "p0f v3," lcamtuf.coredump.cx/p0f3/, 2012.

[18] "University of Oregon Route Views Project," www.routeviews.org/.

[19] V. Paxson, M. Allman, J. Chu, and M. Sargent, "Computing TCP's Retransmission Timer," RFC 6298, Internet Engineering Task Force, 2011.

[20] "How the Internet in Australia went down under," bgpmon.net/blog/?p=554, Feb. 2012.

[21] A. Toonk, "A BGP Leak Made in Canada," www.bgpmon.net/a-bgp-leak-made-in-canada/, Aug. 2012.

[22] G. Huston, "Leaking Routes," www.potaroo.net/ispcol/2012-03/leaks.html, 2012.

[23] L. Quan, J. Heidemann, and Y. Pradkin, "Detecting Internet Outages with Precise Active Probing (extended)," USC/Information Sciences Institute, Tech. Rep., 2012.

[24] E. Katz-Bassett, H. V. Madhyastha, J. P. John, A. Krishnamurthy, D. Wetherall, and T. Anderson, "Studying Black Holes in the Internet with Hubble," in *USENIX Networked Systems Design & Implementation (NSDI 2008)*, 2008.