A Taxonomy of Anomalies in Backbone Network Traffic

Johan Mazel
NII/JFLI
johanmazel@nii.ac.jp

Romain Fontugne
NII/JFLI
romain@nii.ac.jp

Kensuke Fukuda
NII
kensuke@nii.ac.jp

***Abstract*—The potential threat of network anomalies on Internet has led to a constant effort by the research community to design reliable detection methods. Detection is not enough, however, because network administrators need additional information on the nature of events occurring in a network. Several works try to classify detected events or establish a taxonomy of known events. But, these works are non-overlapping in terms of anomaly type coverage. On the one hand, existing classification methods use a limited set of labels. On the other hand, taxonomies often target a single type of anomaly or, when they have wider scope, fail to present the full spectrum of what really happens in the wild.**

We thus present a new taxonomy of network anomalies with wide coverage of existing work. We also provide a set of signatures that assign taxonomy labels to events. We present a preliminary study applying this taxonomy with six years of real network traffic from the MAWI repository. We classify previously documented anomalous events and draw to main conclusions. First, the taxonomy-based analysis provides new insights regarding events previous classified by heuristic rule labeling. For example, some RST events are now classified as network scan response and the majority of ICMP events are split into network scans and network scan responses. Moreover, some previously unknown events now account for a substantial number of all UDP network scans, network scan responses and port scans. Second, the number of unknown events decreases from 20 to 10% of all events with the proposed taxonomy as compared to the heuristic approach.

I. Introduction

Network anomalies often represent a threat to networks, since they have a potentially detrimental effect on users' Internet access. Network anomaly detection is thus a critical task in network management. This research field attracted a lot of attention during the last decade and many proposals have been made. Detection techniques rely on statistical methods such as wavelets [1], Kalman filters [2], hash projection [3]–[5], principal component analysis (PCA) [6], [7], and pattern recognition [8]. Because, these techniques only target event detection, however, they provide limited or no information regarding anomaly characteristics. Event analysis is thus required in order to understand the nature of occurring events. This is an extremely tedious manual task that should be automated.

Our goal is to ease the network administrator's task of anomaly monitoring by providing a framework that classifies events into precise, meaningful categories. Several previous works address network anomaly classification. Works that target classification of detected events [9]–[13]. all use a limited set of signatures (less than 10). On the other hand, several detailed taxonomies have also been proposed. These either focus on a specific type of anomaly like distributed denial of service (DDoS) attacks [14] or scans [15], or consider network anomalies in general [16], [17]. These taxonomies provide diverse, non-overlapping and thus incomplete coverage of all known network anomalies. Furthermore, none of these works provide any material that would allow third parties to easily reproduce the results.

In this work, we propose a network anomaly taxonomy, consisting of a set of anomaly labels (e.g., scan, outage, etc.) and corresponding signatures (i.e., a set of rules characterizing associated network traffic). Our contributions are that: (1) we propose a new taxonomy that widely covers previous taxonomies; (2) we provide it to the research community to confirm reproducibility; and (3) we apply it to real network traffic from the MAWI repository. Our study using real traffic show that our results improve on previous classification results by reducing the proportion of unknown events and providing new insights in terms of anomaly occurrence.

The paper is structured as follows. Related work is discussed in Section II. Our new taxonomy and signatures are described in Section III. We then present a longitudinal study on 6 years of real network traffic in Section IV. In Section V, we discuss our results and potential future work before concluding in Section VI.

II. Related work

There is a large literature related to network anomaly detection. Several methods have been proposed to automatically classify events. Lakhina et al. [9] revisit their PCA-based method [6] but use several entropy values based on source and destination address distributions and source and destination port distributions. They then propose to reuse these entropy values to classify anomalies. They use two categories of unknowns: unknown events in which there is a slight concentration in source and destination addresses, and false alarms for unknown behavior. Xu et al. [10] apply clustering to entropy values similar to [9] in order to build a traffic model and then classify events. Fernandes et al. [11] present NADA, a signature-based tool that classifies anomalies into six categories. Similarly, Silveira et al. [12] propose URCA, a method to identify the root causes of anomalous events. Tellenbach et al. [13] present a Tsallis entropy-based traffic entropy spectrum (TES). Their classification scheme uses simulated anomalies (whose models are provided) and a support vector machine (SVM) to train their classification algorithm. Fontugne et al. [18] use simple heuristic rules to classify events into three main categories.

978-1-4799-0959-9/14/$31.00 ©2014 IEEE

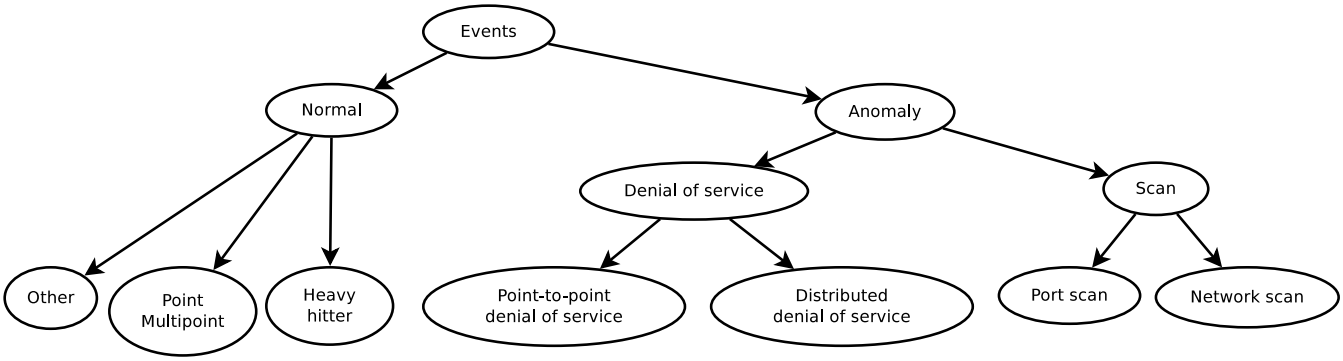


Figure 1. General view of our taxonomy.

Table I. ANOMALY TYPE COVERAGE OF EXISTING WORK (·FOR ANOMALY TYPE, AND ⊙FOR PROVIDED SIGNATURE). THE COVERAGE OF OUR WORK IS NOT COMPLETELY PRESENTED IN THIS TABLE.

	Classification						Taxonomy				
	[9] Entropy	[10] Entropy/clustering	[11] Signatures	[12] Graph clustering	[13] Entropy/SVM	[18] MAWILab	[14] Mirkovic	[15] Brattner	[16] Plonka	[17] CAPEC	Our work
Network scan	⊙	⊙	⊙	⊙	·			·	·	·	⊙
Distributed network scan				·						·	⊙
Port scan	⊙	⊙	⊙	·	·			·	·	·	⊙
DoS	⊙				·				⊙	·	⊙
DDoS	⊙		⊙	·	·		·		⊙	·	⊙
DDoS reflection					·						⊙
Attack response			⊙								⊙
Heavy hitter	⊙			·							⊙
Flash crowd	⊙										
Routing change				⊙					·		
Outage	⊙			⊙	·				·		⊙
Measurement									·		
Point-multipoint	⊙										⊙
Unknown/false alarm or normal	⊙/⊙		/⊙	/·							⊙/
Flag, port, and protocol labels						⊙					

Other network anomaly classification approaches have been proposed. Treurniet [19] targets stub network monitoring. That work relies on network traffic's distributed nature analysis and protocol behavior characterization based on a finite state machine. Such behavior analysis requires complete availability of bidirectional flow traffic. Because of asymmetric routing, however, backbone traffic often contains unidirectional flows [20]. This aspect prevents us from applying a similar approach. On the other hand, Brownlee [21] and Glatz et al. [22] aim at analyze one-way traffic in the context of darknet traffic.

There have also been several proposed anomaly taxonomies. Mirkovic et al. [14] propose a classification of DDoS attacks and defense mechanisms according to several criteria (e.g., IP address spoofing, exploited weakness). Barnett et al. [15] present a taxonomy of scanning events, while Plonka et al. [16] present a taxonomy that covers a wide range of anomalies. CAPEC [17] provides an online database for host attack patterns.

Table I summarizes the coverage of previously proposed anomaly classification methods and taxonomies. This summary clearly shows the non-overlapping coverage of these proposals, along with the lack of available rules and signatures.

In addition, few works have been published on longitudinal studies of anomaly occurrence. Borgnat et al. [23] study seven years of traffic and analyze its long range dependency (LRD). They provide an embryonic analysis of anomalies in the MAWI dataset [1]. Allman et al. [24] study 13 years of scanning activity. To our knowledge, those are the only longitudinal studies of network anomaly occurrence in the wild.

III. TAXONOMY AND SIGNATURES FOR NETWORK ANOMALY CLASSIFICATION

Network anomalies are extremely diverse, since there are many behaviors that should be considered anomalous. These behaviors can be characterized by using criteria on network traffic. Here, we propose an anomaly taxonomy together with associated signatures. The goal is to fully characterize these anomalous behaviors through both the structure of the taxonomy, and the provided signatures.

We first describe the methodology that we followed to create the taxonomy. We then examine the general structure of the taxonomy and the classification process in detail. Finally, we give two signature examples to help the reader understand how signatures are built.

A. Methodology

We build the taxonomy through an iterative process that we bootstrap by applying expert knowledge on network anomalies. We then iteratively refine our anomaly descriptions by carefully examining events that are flagged by detectors but not classified in the taxonomy. Such events are carefully analyzed, and appropriate signatures are built if an interesting and previously uncharacterized behavior is observed.

B. General structure

Figure 1 shows a general view of the structure of our taxonomy. It is separated into two main categories of events: anomalous and normal. Anomalous events comprise denial of service events and scans, while normal events include heavy hitter (also called alpha flows), point-multipoint behaviors, and other kinds of events (outages, tunnels, small point-to-point flows). Some events may be considered either legitimate

[1] http://mawi.wide.ad.jp/mawi/

or illegitimate depending on the context or event magnitude. For example, scans can be research activities [25] or attack precursors. In this work, we follow a conservative, pessimistic approach that considers scans as anomalies.

The taxonomy is built as a tree in which each node contains an anomaly label. The closer a label is to the root of the tree, the more general it is. Each label *may* be associated with a signature. A signature is a set of rules specifying detailed traffic features representing the nature of an event. We actually use more than 80 different signatures. For lack of space, we do not give them all here but we provide two examples in Section III-E.

Regarding label-signature association, some labels have very broad meanings, and it is thus very difficult or even not possible to define such labels through a traffic pattern description. In this case, a label is not associated with a signature. On the other hand, a label further from the root has a more precise behavior and is accordingly linked to a signature.

C. Events labeling and signature matching

We assign a *single* label to each event. We first try to match an event with a label belonging to the subtree whose root is the node labeled "anomaly" in Figure 1. If there is no match, we repeat this process with the "normal" subtree. If there is still no match, the event is labeled as "unknown". One event can match several signatures inside each of the two subtrees. Since the degree of signature specialization increases with distance from the root (cf. Section III-B), however, signatures that match a single event must be on the same path to the root. We also require that such signatures occur consecutively in the path, in order to avoid potential strange behavior during classification because of a faulty taxonomy. When several signatures match an event, we choose the most specialized one, i.e., the one furthest from the root. This ensures that we always choose the most accurate label.

D. Detailed structure

We next address the structure of the taxonomy in more detail.

1) Scans: Scans are events in which hosts want to acquire knowledge about certain targets. We characterize this type of event through two axes representing the scanning pattern and traffic characteristics.

The scanning pattern is determined by the target, which can either be a single machine or several hosts. The first case corresponds to a port scan: one host tries to find open services on a single machine. The probing host will thus send many probe packets to determine whether the target allows connection on a particular port. The second case corresponds to a network scan. This type of scan aims to find either alive hosts or hosts with one or several (usually a small number) of open services or ports. The ultimate goal of this type of event is either to map a network or to identify specific services running or protocol versions (possibly through fingerprinting). In this second case, the attacker will then try to exploit vulnerable machines. We also consider distributed scans, in which several hosts target a great number of machines.

Table II. SIGNATURE EXAMPLES (INDENTATION AFFECTS THE EXPRESSION EVALUATION)

UDP network scan	$nb\_src\_addr < 5$ $\wedge\ nb\_dst\_addr >= 20$ $\wedge \frac{nb\_packets}{nb\_dst\_addr} < 5$ $\wedge \frac{nb\_udp\_packets}{nb\_packets} >= 0.8$
UDP network scan ICMP response	$nb\_dst\_addr < 5.$ $\wedge \frac{nb\_icmp\_packets}{nb\_packets} >= 0.8$ $\wedge \frac{nb\_destination\_unreachable\_packets}{nb\_icmp\_packets} >= 0.8$ $\wedge$ $nb\_src\_addr < 20$ $\wedge$ $\frac{nb\_network\_host\_unreachable\_packets}{nb\_destination\_unreachable\_packets} >= 0.8$ $\vee \frac{nb\_prohibited\_unreachable\_packets}{nb\_destination\_unreachable\_packets} >= 0.8$ $\vee$ $nb\_src\_addr >= 20$ $\wedge \frac{nb\_packets}{nb\_src\_addr} < 20$ $\wedge \frac{nb\_protocol\_port\_unreachable\_packets}{nb\_destination\_unreachable\_packets} >= 0.8$ $\wedge\ nb\_icmp\_src\_addr < 5.$ $\wedge\ nb\_icmp\_dst\_addr >= 20.$ $\wedge \frac{nb\_icmp\_du\_udp\_packets}{nb\_destination\_unreachable\_packets} >= 0.8$

Scanning events exploit special traffic characteristics. ICMP probe packets use the ICMP types "echo request", "timestamp request", and "address mask request". TCP scans use many different flags or combinations of flags. SYN is used to initiate a connection. ACK can map a filtered port (filtered ports answer via ICMP, while open or closed ports answer with RST packet). Combinations of no flag and FIN/PUSH/URG flags force TCP to answer with RST when the targeted port is closed [26]. UDP scans do not exhibit special transport protocol patterns. We thus leverage the scanning pattern of a UDP scan where only a small number of packets is sent to each destination.

2) Scan response: For each type of scan previously identified, we create an associated label corresponding to its response. This label is actually composed of three elements: a main generic label and two sub-labels. The first sub-label targets ICMP error messages (we only use network/host unreachable and prohibited types but plan to add others such as time exceeded, redirect and source quench) that answer a scan. In this case, we assume that the scan is mostly unsuccessful. The second sub-label corresponds to scans that mostly work, i.e. a majority of targets answer the scan. The generic label targets scans that are more successful than those captured by the first sub-label but less successful those corresponding to the second sub-label.

3) Denial of service: This label characterizes denial of services attacks. We target both point-to-point and distributed denial of service. We use header information, such as the SYN flag and ICMP type, to identify DDoS. We identify UDP DDoS as massive UDP traffic toward a single host and port, where the mean packet size is above a threshold. This allows us to isolate anomalous behavior from legitimate UDP traffic exhibiting a communication pattern similar to that of DDoS (such as DNS traffic).

4) Normal: Normal events are subdivided into three sub-labels. Heavy hitters correspond to point-to-point traffic with more than 1000 packets. Point-multipoint events represent server traffic, i.e. point to multipoint for a source server or multipoint to point for a destination server. Finally, "other" events include ICMP errors (outages, expired time-to-live,

Table III. HEURISTIC RULES FOR LABELING TRAFFIC, CORRESPONDING TO A SET OF ALARMS, IN THREE CATEGORIES ("ATTACK", "SPECIAL", AND "UNKNOWN"). THESE RULES ORIGINATE FROM ANOMALIES PREVIOUSLY REPORTED [4], [18] AND MANUAL INSPECTION OF MAWI.

Label	Category	Details
Attack	Sasser worm	Traffic on ports 1023/tcp, 5554/tcp or 9898/tcp
Attack	NetBIOS	Traffic on ports 137/udp or 139/tcp
Attack	RPC	Traffic on port 135/tcp
Attack	SMB	Traffic on port 445/tcp
Attack	Ping	High ICMP traffic
Attack	Other TCP attacks	Traffic with more than 7 packets and: SYN, RST or FIN flag $\geq 50\%$
Attack	Other attacks	FTP, SSH, HTTP, HTTPS traffic with SYN flag $\geq 30\%$
Special	FTP/SSH/HTTP/HTTPS	Traffic on ports 20/tcp, 21/tcp, 22/tcp, 80/tcp and 8080/tcp, 443/tcp with SYN flag $\leq 30\%$
Unknown	Unknown	Traffic that does not match other heuristics

etc.), point-to-point traffic due to tunnels (GRE or IPv4-IPv6) and small-volume point-to-point traffic (less than 1000 packets).

E. Anomaly signature examples

Each signature is composed of one or more rules covering attributes that describe the nature of traffic. For example, we use the number of source hosts and the number of destination hosts to characterize the distributed or point-to-point behavior of an event. This is similar to previous work in traffic classification [27], [28]. We also define indexes that convey the distribution of port numbers in a similar way as entropy does. Finally, we use proportions of packets that fit a certain pattern (e.g., proportion of ICMP packets, proportion of "destination unreachable" ICMP packets among ICMP packets, proportion of TCP packets with SYN flag set among TCP packets, etc.). These principles are similar to those applied in [11].

Table II lists two examples of an anomaly signature. The first example is a UDP network scan. We characterize a UDP network scan as a set of packets with a small number of sources, a high number of destinations, a small number of packets for each destination (which is consistent with probing activity), and a high proportion of UDP packets. The second example is a UDP network scan response and is actually the signature for the first sub-label given in Section III-D2 regarding UDP network scans, i.e., an ICMP answer to an unsuccessful UDP network scan. The number of destinations is small because we consider that the scan source was a single host. The majority of the packets is ICMP, and among those, the majority have the ICMP type "destination unreachable". The signature then takes into account two cases: either the network/hosts are unreachable (ICMP code 0/1) or forbidden by firewall or security rules (ICMP code 9/10/13); or the host is online but the protocol/port is not available/open (ICMP code 2/3). In the first case (shown in red in the table), the gateway is sending the ICMP packets, and there is thus a single source. In the second case (blue), each target answers and the number of sources is thus higher. The last three lines of the signature describe the original UDP packets encapsulated in the ICMP "destination unreachable" packets and ensure that they actually constitute a scan.

Here, the thresholds for the numbers of sources and destinations are not strict (e.g. the threshold for number of source hosts for a scan is 5 and thus does not enforce limitation to a single source). This results from our iterative signature building process. In the original use case of our taxonomy, we classify events documented in MAWILab (cf. Section IV-A). These events are flagged by detectors and combined together. We notice that, unfortunately, events of similar nature are sometime merged and thus generate unorthodox traffic. For example, two scans targeting the same network may be grouped into a single event. The relaxed thresholds allow us to account for such behavior.

IV. NETWORK TRAFFIC ANALYSIS

We next apply our new taxonomy to previously documented anomalies found in raw network traffic. We first give the technical background of the network traffic used for this work. We then compare our new classification results with those obtained by previously used heuristic rules. Finally, we discuss a longitudinal study of anomaly occurrence.

A. Background

The general context of this work is the study of the MAWI repository. MAWI is a public collection of 15-minute network traffic traces captured every day on a backbone link between Japan and the USA since 2001. Building on this repository, the MAWILab project [18] dataset[2] aims to identify anomalies present in MAWI traces. MAWILab uses a combination of four anomaly detectors based on different theoretical backgrounds [4], [5], [7], [8]. The studied traffic spans six years, from 2001 to 2006.

In this paper, we intend to classify events from the MAWILab repository. We use alarm reports containing host IP addresses to extract associated network traffic and capture various traffic features related to our signatures (cf. Section III-E). We then match these features against signatures associated with our taxonomy. Here, we only consider events classified as anomalous or suspicious by MAWILab.

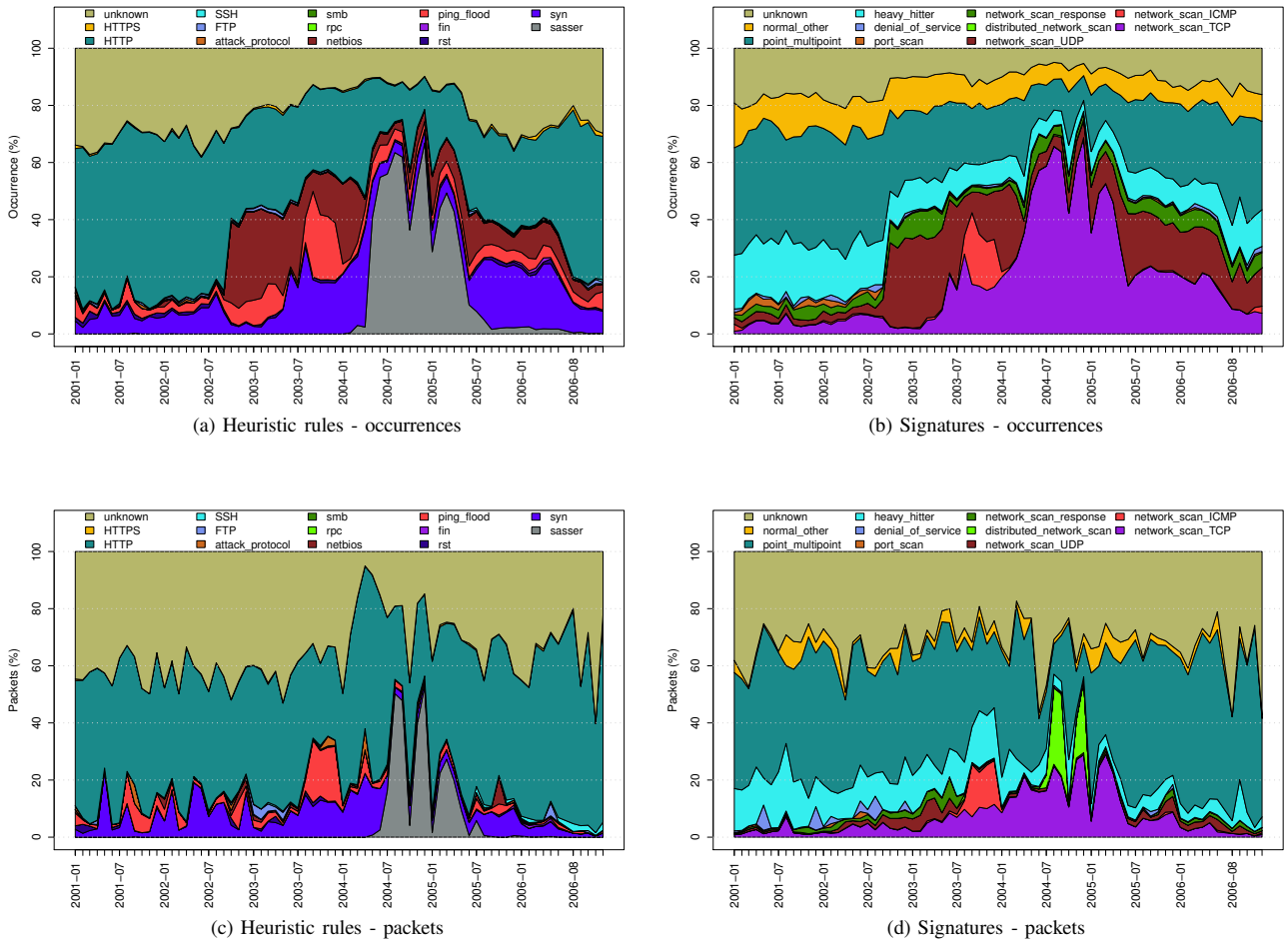
B. Comparison between heuristic- and signature-based classification

We here compare taxonomy-based classification results with results obtained using the heuristic rules listed in Table III. Those rules have previously been used for classification in the context of MAWI and MAWILab [18], [23]. Table IV shows the confusion matrix between the heuristic-based results and the new results. Although the results vastly differ in terms of event labels, some results are consistent and allow us to cross-validate our results. For example, the overwhelming majority of sasser (Sasser is a computer worm that emerged in April 2004 [29]) and syn events are classified as TCP scans by signatures. Further breakdown of TCP scans by port-based signatures allows us to identify Sasser-linked activity. For the lack of space, however, we do not present that level of detail in this paper. Furthermore, NetBIOS events are now massively classified as UDP scans. These examples show that some results are consistent across the two classification methods.

[2]http://www.fukuda-lab.org/mawilab/

Table IV. CONFUSION MATRIX BETWEEN HEURISTIC-RULES AND TAXONOMY-BASED RESULTS. BOLD NUMBERS INDICATE INTERESTING OVERLAPS BETWEEN EVENTS OBTAINED HEURISTICALLY AND NEW EVENTS EXTRACTED BY SIGNATURES.

Heuristic \ Taxonomy	network scan TCP	network scan ICMP	network scan UDP	distributed network scan	network scan response	port scan	denial of service	heavy hitter	point multipoint	normal other	unknown	total
sasser	**46254**	0	0	216	566	0	4	5	203	3	104	47355
syn	**17521**	0	0	3	4	117	964	143	1659	172	452	21035
rst	0	0	0	0	**380**	4	0	162	887	193	124	1750
fin	86	0	0	0	0	5	0	8	43	8	4	154
ping_flood	0	**2868**	0	0	**5129**	0	244	5	2	2816	439	11503
netbios	9	0	**20433**	0	1	0	0	51	406	36	328	21264
rpc	0	0	7	0	0	0	0	1	6	1	7	22
smb	14	0	0	0	0	0	0	13	46	12	10	95
attack_protocol	0	1	0	0	0	0	0	120	183	141	77	522
FTP	2	0	0	0	1	0	0	**289**	**301**	**317**	35	945
SSH	2	0	5	0	0	0	0	**82**	**207**	**59**	8	363
HTTP	104	0	7	0	3	**388**	0	**12885**	**26556**	**10627**	12610	63180
HTTPS	2	0	0	0	0	0	0	**270**	**367**	**194**	103	936
unknown	213	0	**7243**	0	**2881**	**413**	**13**	**5576**	**15353**	**3207**	6735	41634
total	64207	2869	27695	219	8965	927	1225	19610	46219	17786	21036	210758



(a) Heuristic rules - occurrences

(b) Signatures - occurrences

(c) Heuristic rules - packets

(d) Signatures - packets

Figure 2. Events occurrence (a,b) and packet number (c,d) from the MAWILab repository detected by heuristic rules (cf. Table IV) (a,c) and the newly created taxonomy (b,d).

The advanced taxonomy-linked signatures also allow a precise breakdown of events classified heuristically. For example, almost half of the rst events obtained heuristically are now classified as network scan responses, representing 4% of all such responses. Likewise, detailed analysis of ICMP traffic shows that ping flood events constitute almost all new ICMP scans and 57% of network scan responses. The new signatures can also extract many new event types for events

whose nature was previously unknown. These newly extracted events account for 26% of all UDP network scans, 32% of all network scan responses, and 45% of all port scans. On a more general note, FTP, SSH, HTTP, and HTTPS heuristic labels are now largely classified in the categories of heavy hitter (69% of all heavy hitters), point-multipoint (59% of all point-multipoints events), and other (63% of all other).

Regarding unknown events, the heuristic rules cannot classify 20% of all such occurrences. The taxonomy-linked signatures exhibit better results by reducing the proportion of unknowns to 10% of all events. In other words, our signatures give 10 percentage points fewer unknown events than with the heuristic rules. On a side note, we observe that unknown events typically exhibit multiple sources and destinations.

It is also important to note that the signatures use a more conservative threshold than do the heuristic rules. For example, in our taxonomy the threshold for the proportion of SYN packets among TCP packets is 80% for all SYN-related signatures. This threshold is greater than any threshold used in the heuristic rules (cf. Table III). The fact that signatures give better results even with using more conservative thresholds further demonstrates the improved performances of this approach in relation to heuristic rules.

C. Longitudinal comparison between heuristic- and signature-based classification

We now discuss a longitudinal study of anomaly occurrences over six years, from 2001 to 2006, as shown in Figure 2 in terms of the proportions of both occurrences and packets. We only cover six years because of page limitations and also because these years are well known [18]. We intend to study the remaining years, from 2007 and onward, in future works. In the heuristic-based rules, sasser events represent communication on known Sasser backdoors. Sasser's main activity spans from May 2004 to June 2005 in Figure 2a. This event surge is also visible in Figure 2b as a surge of TCP scans in May 2004. As explained in the previous Section IV-B, TCP scans in the new taxonomy are actually composed of sasser and syn events obtained heuristically. This is consistent with what Figure 2 shows. This surge of scans is also visible in the packet-related figures (2c and 2d). It is interesting to note that a very small number of distributed scans account for a significant quantity of probing packets at the end of 2004: 27% in August, 29% in September, 13% in November, and 25% in December.

Regarding ping flood events found heuristically, there is a surge that starts in September 2003 and lasts until December 2003. This rise is also visible in Figures 2b and 2d. The new classification approach, however, allows us to understand that this rise is actually linked to ICMP network scans. This breakdown of ping flood events is very interesting because it allows us to understand the actual nature of a particular surge whose nature was hidden.

Another new result obtained through taxonomy-based classification (already listed in Table IV) is the breakdown of traffic heuristically labeled as "special". Figure 2 clearly shows that the new signatures separate FTP/SSH/HTTP/HTTPS events into (in decreasing proportion): point-multipoint events, heavy hitter and other events (which, in fact, are mainly "light hitters", i.e., point-to-point traffic of less than 1000 packets).

Finally, the taxonomy-based classification also generate fewer "unknown" events across the six analyzed years, as compared to the heuristic-based classification.

V. Discussions and future work

Although we intend the taxonomy to be exhaustive, this is a complicated task. Our advanced signatures leverage communication patterns and header information analysis to provide detailed insights into the nature of events. These insights are impossible to acquire through simple analysis using heuristic rules based on flags or protocol use. The new anomaly signatures thus provide better coverage than that of the previously used heuristic rules and the existing state-of-the-art. We know that taxonomy building always remains a work in progress. We will thus continue to follow our iterative process in order to adapt our taxonomy to events detected in more recent years. For example, we intend to add signatures for DNS and NTP reflection attacks since several high-profile attacks of this type happened in 2013 and 2014.

In this paper, we consider traffic captured at a single point. We thus cannot find network-wide anomalies, as Lakhina et al. [9] proceed for gaps, failures, or routing changes. URCA [12] showed, however, that by analyzing consecutive time windows with volume-based metrics it is possible to find outage and routing changes. Our current detection of outages (both host and network) relies on analysis of ICMP packets of the "destination unreachable" type. One possible improvement in our method would be to perform a similar kind of time-based analysis and correlate the results with ICMP traffic analysis.

VI. Conclusion

We propose a new taxonomy for accurate classification of network anomaly in backbone traffic. Our taxonomy supersedes existing anomaly classification work. We also define a set of taxonomy-associated signatures, which rely on traffic features to correctly classify anomalous network events.

We apply our taxonomy over six years of events obtained by state-of-the-art detectors in the MAWI repository. Our results are consistent with previous classification done through simple heuristic rules, but we also provide a deeper understanding of several previous event types. For example, some RST events are actually network scan responses, and the majority of ICMP events are actually network scans and network scan responses. Moreover, some previously unknown events are now classified as UDP network scans, network scan responses, and port scans. We also show that the new taxonomy reduces the proportion of unknown events from 20 to 10% of all events.

Reproducibility and comparison of results are paramount to scientific progress. We thus make our taxonomy available to researchers [30], and we intend to release an associated classification tool. We welcome any feedback or suggestions.

Acknowledgments

This research has been supported by the Strategic International Collaborative R&D Promotion Project, of the Ministry

of Internal Affairs and Communication, Japan, and by the European Union Seventh Framework Programme (FP7/ 2007-2013) under grant agreement No. 608533 (NECOMA).

REFERENCES

[1] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," ser. IMW '02, pp. 71–82.

[2] A. Soule, K. Salamatian, and N. Taft, "Combining filtering and statistical methods for anomaly detection," ser. IMC '05, pp. 31–31.

[3] B. Krishnamurthy, S. Sen, Y. Zhang, and Y. Chen, "Sketch-based change detection: methods, evaluation, and applications," ser. IMC '03, pp. 234–247.

[4] G. Dewaele, K. Fukuda, P. Borgnat, P. Abry, and K. Cho, "Extracting hidden anomalies using sketch and non gaussian multiresolution statistical detection procedures," ser. LSAD '07, pp. 145–152.

[5] D. Brauckhoff, X. Dimitropoulos, A. Wagner, and K. Salamatian, "Anomaly extraction in backbone networks using association rules," *IEEE/ACM Transactions on Networking*, 2012.

[6] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," *ACM SIGCOMM Computer Communication Review*, 2004.

[7] Y. Kanda, R. Fontugne, K. Fukuda, and T. Sugawara, "ADMIRE: Anomaly detection method using entropy-based pca with three-step sketches," *Computer Communications*, 2013.

[8] R. Fontugne and K. Fukuda, "A Hough-transform-based anomaly detector with an adaptive time interval," *ACM SIGAPP Applied Computing Review*, 2011.

[9] A. Lakhina, M. Crovella, and C. Diot, "Mining anomalies using traffic feature distributions," *ACM SIGCOMM Computer Communication Review*, 2005.

[10] K. Xu, Z.-L. Zhang, and S. Bhattacharyya, "Internet traffic behavior profiling for network security monitoring," *IEEE/ACM Transactions on Networking*, 2008.

[11] G. Fernandes and P. Owezarski, "Automated classification of network traffic anomalies," ser. SecureComm '09, pp. 91–100.

[12] F. Silveira and C. Diot, "URCA: Pulling out anomalies by their root causes," ser. INFOCOM '10, pp. 1–9.

[13] B. Tellenbach, M. Burkhart, D. Schatzmann, D. Gugelmann, and D. Sornette, "Accurate network anomaly classification with generalized entropy metrics," *Computer Networks*, 2011.

[14] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Computer Communication Review*, 2004.

[15] R. J. Barnett and B. Irwin, "Towards a taxonomy of network scanning techniques," ser. SAICSIT '08, pp. 1–7.

[16] D. Plonka and P. Barford, "Network anomaly confirmation, diagnosis and remediation," ser. CCC '09, pp. 128–135.

[17] "Capec," http://capec.mitre.org/.

[18] R. Fontugne, P. Borgnat, P. Abry, and K. Fukuda, "MAWILab: combining diverse anomaly detectors for automated anomaly labeling and performance benchmarking," ser. Co-NEXT '10, pp. 1–12.

[19] J. Treurniet, "A network activity classification schema and its application to scan detection," *IEEE/ACM Transaction on Networking*, 2011.

[20] W. John, M. Dusi, and K. C. Claffy, "Estimating routing symmetry on single links by passive flow measurements," ser. IWCMC '10, pp. 473–478.

[21] N. Brownlee, "One-way traffic monitoring with iatmon," ser. PAM'12.

[22] E. Glatz and X. Dimitropoulos, "Classifying internet one-way traffic," ser. IMC '12.

[23] P. Borgnat, G. Dewaele, K. Fukuda, P. Abry, and K. Cho, "Seven years and one day: Sketching the evolution of internet traffic," ser. INFOCOM '09, pp. 711 –719.

[24] M. Allman, V. Paxson, and J. Terrell, "A brief history of scanning," ser. IMC '07, pp. 77–82.

[25] L. Quan, J. Heidemann, and Y. Pradkin, "Trinocular: Understanding internet reliability through adaptive probing," *ACM SIGCOMM Computer Communication Review*, 2013.

[26] J. Postel, "Transmission Control Protocol," RFC 793.

[27] T. Karagiannis, K. Papagiannaki, and M. Faloutsos, "BLINC: Multilevel traffic classification in the dark," *ACM SIGCOMM Computer Communication Review*, 2005.

[28] Y. Himura, K. Fukuda, K. Cho, P. Borgnat, P. Abry, and H. Esaki, "Synoptic graphlet: Bridging the gap between supervised and unsupervised profiling of host-level network traffic," *IEEE/ACM Transactions on Networking*, 2013.

[29] "Sasser," http://www.cve.mitre.org/cgi-bin/cvename.cgi?name=CAN-2003-0533.

[30] "Classification," http://www.fukuda-lab.org/mawilab/classification/.