An Empirical Evaluation of Entropy-based Traffic Anomaly Detection

George Nychis, Vyas Sekar, David G. Andersen, Hyong Kim, Hui Zhang
Carnegie Mellon University

ABSTRACT

Entropy-based approaches for anomaly detection are appealing since they provide more fine-grained insights than traditional traffic volume analysis. While previous work has demonstrated the benefits of entropy-based anomaly detection, there has been little effort to comprehensively understand the detection power of using entropy-based analysis of multiple traffic distributions in conjunction with each other. We consider two classes of distributions: flow-header features (IP addresses, ports, and flow-sizes), and behavioral features (degree distributions measuring the number of distinct destination/source IPs that each host communicates with). We observe that the timeseries of entropy values of the address and port distributions are strongly correlated with each other and provide very similar anomaly detection capabilities. The behavioral and flow size distributions are less correlated and detect incidents that do not show up as anomalies in the port and address distributions. Further analysis using synthetically generated anomalies also suggests that the port and address distributions have limited utility in detecting scan and bandwidth flood anomalies. Based on our analysis, we discuss important implications for entropy-based anomaly detection.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication-Networks**]: Network Operations—*network management, network monitoring*

General Terms

Management, Measurement

Keywords

Entropy, Anomaly Detection

1. INTRODUCTION

There has been recent interest in the use of entropy-based metrics for traffic analysis [20] and anomaly detection [10, 4, 7, 5, 12, 18]. The goal of such analysis is to capture fine-grained patterns in traffic distributions that simple volume based metrics cannot identify [10].

Several traffic features (e.g., flow size, ports, addresses) have been suggested as candidates for entropy based anomaly detection. However, there has been little work in understanding the analysis capabilities provided by a set of entropy metrics used in conjunction with one another. For example, it is unknown whether the different features complement each other, or if they detect the same anomalies and are redundant.

The goal of this paper is to provide a better understanding of the use of entropy-based methods in anomaly detection. We consider two types of distributions based on *flow-header* features and *behavioral* features. The flow-header features are addresses (source and destination), ports (source and destination), and the flow size distribution (FSD) [10, 5, 9]. The behavioral features are the in and out-degree distributions (degree of an end-host X is the number of distinct IP addresses that X communicates with) that capture the structure of end-host communication patterns.

The key results from our measurement study are:

• Port and address distributions are highly correlated, with pairwise correlation scores greater than 0.95. The degree distributions and FSD are weakly correlated with each other and with the port/address distributions.

• The correlation between the source (destination) port and source (destination) address distribution arises due to the nature of the underlying traffic patterns. However, the correlations across the source and destination distributions stem from the uni-directional nature of flow-level measurements available today.

• The anomalies detected by the port and address distributions overlap significantly. In our dataset, almost all the anomalies detected by these distributions are *alpha flows* [10]. In contrast, host degree distributions and FSD identify anomalous scan, DoS, and P2P activity that are not detected by the port and address distributions.

• Experiments with synthetically generated anomalies show that FSD and the degree distributions detect scanning events that cannot be detected by the port and address distributions. For DDoS-style events, port and degree distributions detect only high-magnitude events that would have appeared as traffic volume anomalies.

These observations have important implications for entropy-based analysis. First, we should select candidate distributions with care. While ports and addresses have been commonly suggested [10] as good candidates for entropy-based

IMC'08, October 20–22, 2008, Vouliagmeni, Greece.

anomaly detection, our measurements question this rationale. Our results also suggest a natural approach: select traffic distributions that inherently complement one another and thus provide different views into the underlying traffic structure. Second, we need to move beyond the traditional uni-directional flow semantics available today (e.g., [14, 16]), since it can artificially skew the properties of the underlying distributions. Thus, it is prudent for administrators to use bi-directional flow collection tools whenever possible. Finally, we discuss how to use the correlations to design a better anomaly detection system. Our preliminary results show that using time-series anomaly detection on the correlation scores can expose new anomalies that do not manifest in the raw time-series.

2. PRELIMINARIES

Datasets: Our primary dataset uses (bi-directional) flow data [2] captured in February 2005 at Carnegie Mellon University. [1] The dataset contains traffic to and from tens of thousands of active IP addresses involving roughly 92 TB of total traffic over 2.5 billion flows. IP addresses in the dataset were anonymized preserving a one-to-one mapping between actual and anonymized IP addresses [19]. Application ports were not anonymized. The traffic feature distributions we study are unchanged by the anonymization.

The dataset is split into five minute non-overlapping epochs consisting of flows that completed within the epoch. Each (bi-directional) flow record consists of source/destination pairs for the IP address, port, packet count, and byte count. It also includes the connection time, protocol used, connection state, and flow direction. However, in some cases the directionality is not evident from the flow record (e.g., UDP flows, long-lived TCP flows that extend beyond the flow timeout). In such cases, we use application port numbers to infer flow direction.[2]

We also corroborate specific parts of our analysis with Netflow [14] data from Internet2, GÈANT, and a (different) university department.

Approach: The entropy of a random variable X is $H(X) = -\sum_{i=1}^{N} p(x_i)\log\left(p(x_i)\right)$, where $x_1, \ldots, x_N$ is the range of values for X, and $p(x_i)$ represents the probability that X takes the value x_i.[3] We compute the normalized entropy (between zero and one) as $\frac{H}{\log(N_0)}$, where N_0 is the number of distinct x_i values present in a given measurement epoch.

We study seven empirical traffic distributions. Five of these are obtained from flow-headers: source address, destination address, source port, destination port, and flow size distribution measured in packets per flow (FSD). Prior work on using flow-header features in entropy-based analysis uses uni-directional flow information (e.g., [14, 16]). Hence, we explicitly convert each bi-directional flow record [2] into two uni-directional flows for computing the distributions over the flow-header features. For each source (destination) address (port) x_i, we calculate the probability

$$p(x_i) = \frac{\text{Number of pkts with } x_i \text{ as src (dst) address (port)}}{\text{Total number of pkts}}$$

The normalization factor is $\log(N_0)$, where N_0 is the number of active source (destination) addresses (ports) observed during the measurement epoch.

The remaining two distributions are based on inter-host communication behavior. We consider the in- and out-degree of each active internal IP address inside the network under consideration (e.g., in our dataset we only consider hosts inside the university): these are the only hosts for which we have a complete view of both incoming and outgoing traffic. For a host X, the out-degree is the number of distinct IP addresses that X contacts, and the in-degree is the number of distinct IP addresses that contact X. The degree distributions are computed using bi-directional flows. For each value of out-degree (in-degree) x_i, we calculate the probability

$$p(x_i) = \frac{\text{Number of hosts with out-degree } x_i}{\text{Total number of hosts}}$$

The normalization factor is $\log(D)$, where D is the number of distinct out-degree (in-degree) values observed during the measurement epoch.

For each measurement epoch, we compute the normalized entropy for the seven distributions. Let Y_{ij} denote the normalized entropy of distribution i (e.g., source address) observed in epoch j, and Y_i denote the timeseries of normalized entropy values for distribution i. Given the Y_is, we compute the pairwise correlation coefficients between every pair of timeseries vectors Y_i and $Y_{i'}$, $\gamma(Y_i, Y_{i'}) = \frac{\sum_j Y_{ij}Y_{i'j} - n\overline{Y_i}\,\overline{Y_{i'}}}{(n-1)\sigma_{Y_i}\sigma_{Y_{i'}}}$, where $\overline{Y_i}$ and $\overline{Y_{i'}}$ are the sample means of Y_i and $Y_{i'}$, σ_{Y_i} and $\sigma_{Y_{i'}}$ are the sample standard deviations of Y_i and $Y_{i'}$, and n is the number of epochs. We also apply timeseries anomaly detection on each Y_i using the wavelet analysis technique proposed by Barford et al [3].[4]

3. MEASUREMENT RESULTS

3.1 Correlations in Entropy Timeseries

Table 1 shows the pairwise correlation scores between the entropy timeseries of different distributions. We find strong correlations (> 0.95) between the address and port distributions. The remaining metrics show low or no correlation. Figure 1 shows the entropy timeseries values over the entire month-long trace. The visual confirmation of the correlations is just as striking as the values themselves. Additionally, we observe that many of the spikes and deviations in the timeseries plots are also highly correlated. We will revisit these anomalies in the subsequent discussions.

To confirm that these results are not an artifact of our dataset, we perform similar analysis using data from other networks and time periods: Internet2, GÈANT, Georgia Tech, and CMU-2008. All the datasets are large; consisting of over a hundred thousand flows per 5-minute bin, and span multiple weeks. The Internet2 and GÈANT traces consist of flow data from each of the vantage points (11 and 22 respectively). Table 2 summarizes the average and standard deviation of the correlation scores among the ports

[1] The router observes all traffic between university hosts and external Internet hosts. It also observes a significant fraction of internal inter-departmental traffic

[2] Rationale: If a host is running a well-known application service, then it is likely to be the server. Since the client initiates a connection in client-server transactions, we assume that the host that does not use the well-known port is the connection initiator.

[3] All logarithms are base 2 and $0 \log 0 = 0$.

[4] We also use a heuristic anomaly detection approach to rule out biases due to the wavelet analysis [15].

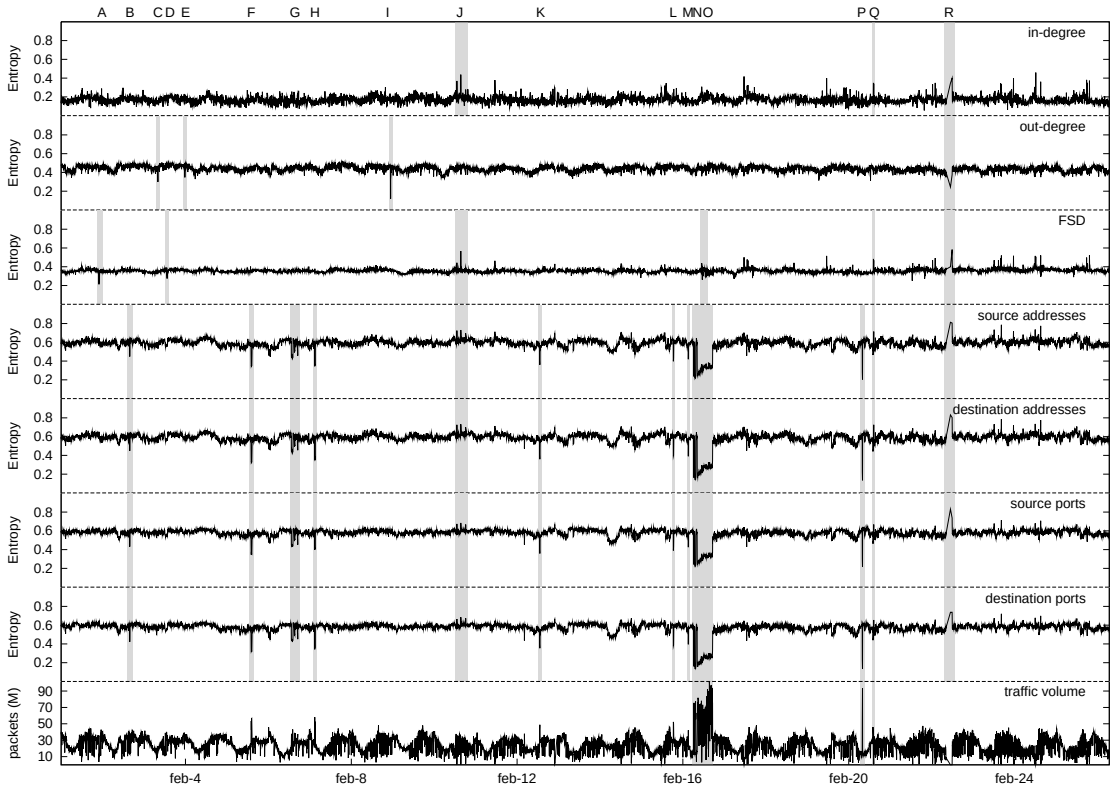


Figure 1: Time series of entropy data for CMU, February 2005, with anomalous event labels.

	Out Deg	Src Addr	Dst Addr	Src Port	Dst Port	FSD
InDeg	0.102	0.100	0.097	0.000	0.007	0.414
OutDeg	-	-0.034	-0.033	-0.054	-0.015	-0.018
SrcAddr	-	-	0.994	0.962	0.956	0.307
DstAddr	-	-	-	0.966	0.969	0.286
SrcPort	-	-	-	-	0.989	0.171
DstPort	-	-	-	-	-	0.181

Table 1: Correlation of entropy timeseries on CMU-2005 dataset.

Trace (#routers)	Date	Avg	SDev	Min	Max
CMU (1)	3/1/08–3/31/08	0.98	0.01	0.96	0.99
GA Tech (1)	2/12/08–3/22/08	0.94	0.020	0.91	0.96
Internet2 (11)	12/1/06–12/14/06	0.84	0.07	0.76	0.98
GÈANT (22)	11/1/05–11/31/05	0.89	0.07	0.81	0.98

Table 2: Correlation from other traces.

and addresses.[5] The strong correlations we observe are not unique to one dataset. Further, the correlations on the CMU dataset are stable across the 2005 and 2008 datasets. For the remainder of the paper, we focus on results using the CMU 2005 dataset.

3.2 Correlations in Anomaly Deviation Scores

Next, we explore if the correlations in the entropy timeseries values also extend to anomalies. We compute the anomaly deviation score of each epoch as the magnitude of normalized local variance computed over a sliding window of size six representing a half-hour interval [3].

[5]Internet2, GEANT, and Georgia Tech only provide unidirectional Netflow [14] style flow records. Thus, we cannot repeat the degree analysis on these.

	Out Deg	Src Addr	Dst Addr	Src Port	Dst Port	FSD
InDeg	0.248	0.199	0.188	0.185	0.156	0.507
OutDeg	-	0.179	0.165	0.143	0.122	0.396
SrcAddr	-	-	0.991	0.971	0.964	0.319
DstAddr	-	-	-	0.970	0.971	0.300
SrcPort	-	-	-	-	0.986	0.256
DstPort	-	-	-	-	-	0.220

Table 3: Correlations of wavelet deviation scores.

Table 3 shows that the port and address distributions are as strongly correlated in deviation scores as they are in terms of the entropy values. Interestingly, the behavioral features become slightly more correlated to the other metrics. For example, the correlation between out-degree and FSD increases by 0.414 from their correlation in entropy to their correlation in deviation scores. We hypothesize the reason for this increase is that the in and out-degree distributions show more stochastic variations than the other distributions. Thus, they tend to be uncorrelated in terms of the timeseries values. However, the wavelet analysis removes the noisy variations and the deviation scores become more correlated.

3.3 Understanding the correlation in port and address distributions

First, we rule out that the correlations arise as an artifact of factors such as entropy normalization, packet vs. byte counts, and time scale of computing correlation values (hour vs. day. vs. week). Once we eliminate these factors, we posit that the correlations are due to (a) a fundamental property of the underlying traffic patterns and/or (b) the unidirectional flow accounting model. For (b), note that this is not specific to our analysis and data. Uni-directional

Anomaly Type	Affected Metrics	Labels
Alpha Flows (Botnet activity)	Addresses, Ports	B, F-H, K-N, P
Scans	FSD	A, D
P2P Supernode Activity	FSD	O
Spoofed DoS	Degree	C, E, I
Measurement Outage	Inconsistent	J, Q, R

Table 4: Labeled traffic anomalies

flow measurements are often the only type of measurements available to network operators today.

To decouple (a) and (b), we consider the CMU-2005 dataset since it has additional bi-directional annotations allowing us to eliminate effects from the uni-directional model. Under a bidirectional flow model, we find that the source (destination) port and source (destination) address distributions are structurally similar due to inherent properties of end-host behavior. However, the correlations between the source and destination entities arise due to the uni-directional nature of the flow measurements – each packet contributes to both source and destination pairs. This causes the uni-directional distributions to be approximately the union of their bi-directional pairs.

3.4 Understanding Anomalies In-Depth

Why do the anomalies detected by the port and address distributions overlap and why do FSD and degree distributions provide unique detection capabilities? To answer this, we use a heuristic approach to identify eighteen major events, indicated by alphabetical labels in Figure 1, and summarized in Table 4. (Anomalies spanning multiple epochs are clustered into a single event.) In the absence of ground truth for our data, we develop a semi-automated anomaly labeling approach that explains the observed anomalies.

The labeling technique consists of the following steps. First, we analyze the *top-k* contributors within each distribution (e.g., top 50 destination address receiving the most number packets) and check if the top-k set changes during the anomaly. The rationale behind this approach is that the top few contributors to the distributions are relatively stable during normal operation but may change significantly during the anomalies (e.g., if a new host/port entry enters the top-k). Next, we identify the flows corresponding to these new entries in the top-k set. Finally, we remove these flows and recompute the entropy and wavelet scores over the remaining flows. If the anomaly subsides (i.e., the new anomaly score computed over the residual data is lower than the anomaly threshold), we attribute the anomaly event to these new top-k entries.

Events J, Q, and R are measurement anomalies, characterized by few or no flow records in our dataset which show no consistent behavior across the different traffic features. In alpha flows (events $F-H$ and $L-N$), a few ports and addresses (both source and destination) dominate the total traffic volume [10], *decreasing* entropy. The events contain a large volume of UDP traffic destined to a single external host on popular application ports (80,53), which seem to be triggered after a small amount of TCP traffic is transferred on port 6667 (IRC botnet control). The alpha flows are detected by *all* the port and address distributions. Further, alpha flows are the only type of anomaly detected by the port and address distributions. This suggests that in our trace, using all four port and address distributions provides no additional detection capabilities compared with using only one of the port or address distributions.

The series of anomalies collectively labeled O are caused by an internal host being recruited as a P2P "supernode" in the Kazaa network [8]. During the event, many hosts connect to this supernode creating a significant number of small flows causing a sharp decrease in the entropy of the FSD. In event A, a single internal host scanned more than 350,000 unique external hosts, using a fixed source port of 666. As there are a large number of small flows, FSD detects the scan. Event D is an outbound scan with a single internal host scanning numerous external hosts on multiple ports. Only FSD detects the scan. In anomalies C, E, and I, a large number of spoofed "hosts" send attack traffic to a single destination on port 6667. The set of source addresses in these flows spans the entire /16 of the university address space using a small range of port numbers. This leads us to believe that an internal host may be sending attack traffic with spoofed source addresses (within the same subnet) to avoid egress filtering.[6]

Events O and H are particularly interesting with respect to the anomaly labeling heuristic. When removing the initial alpha flow event (N), we found that the anomaly O persisted, which ports and addresses alone cannot detect. Event H consisted of two independent alpha flows from which our initial analysis revealed one, and after discovering that an anomaly persisted we discovered the second event. Contrary to conventional wisdom, port and address distributions do not show significant deviations for the scanning anomalies in our data. However, FSD detects such abnormal scanning activity.

4. USING SYNTHETIC ANOMALIES

We use synthetically generated anomaly events to complement our measurement results. Table 5 presents a taxonomy of the five synthetic anomalies we evaluate. For each type of anomaly, we want to identify the traffic distribution(s) that provide the most effective detection capability. To understand the detection sensitivity, we vary the scale of the anomaly using an anomaly-specific control parameter (e.g., number of sources involved in a DDoS or scan attack). We insert the anomaly at 50 random locations in the month-long trace, and report the average to ensure that the results are not biased by time-of-day and day-of-week effects. In the case of the DDoS and bandwidth floods, we are also interested in comparing entropy-based detection to simple volume based detection.

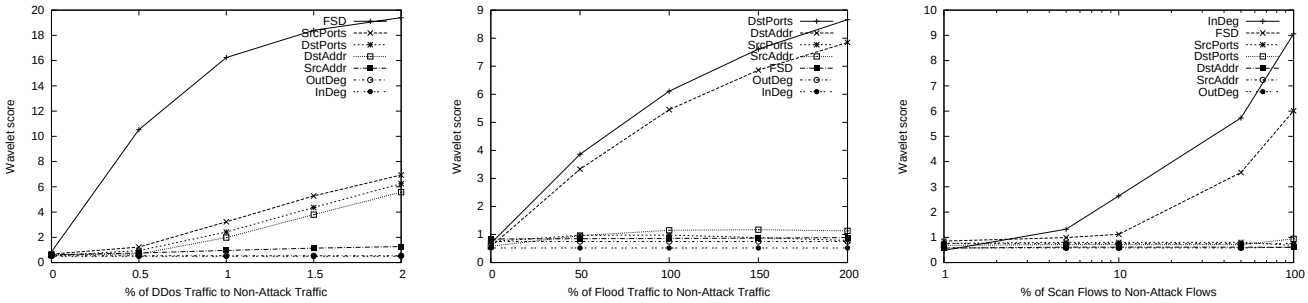
Inbound DDoS Flood: Each DDoS event is characterized by a single destination address receiving a large volume of single-packet flows (to overwhelm the bandwidth and processing capacity of the server and routers). Figure 2(a) shows the anomaly scores as a function of the percentage of total DDoS traffic. Each attack source generates 10 kilobits per second of attack traffic, using a fixed packet size of 57 bytes and a single flow per packet. The attack flows are destined to port 80 on a randomly chosen host inside the university. We have repeated the experiments varying the destination port and the choice of destination address (picking a high-volume, random, and low-volume host) and found similar results.

The change in the FSD can easily detect the anomaly even at a low magnitude since a single flow is used per packet. The destination port and destination address distributions

[6]Since we only have anonymized flow level traces, we could not further validate this hypothesis.

Anomaly Type	SrcAddr	DstAddr	SrcPort	DstPort	FlowSize
Inbound DDoS Flood	Random	Fixed	Random	Fixed	Fixed (10 Kbps), 1 flow per packet
B/W Flood	Random	Fixed	Random	Fixed	Random (300-400 Kbps), 1 flow per host
Single Scanner	Fixed	Random	Random	Fixed	1-3 packets (10% response rate)
Multiple Scanners	Random	Random	Random	Fixed	1-3 packets (10% response rate)
Port Scan	Fixed/Random	Fixed	Random	Sequential	1-3 packets (10% response rate)

Table 5: Taxonomy of synthetic anomalies used in our evaluation



(a) DDoS flood. (b) Bandwidth flood. (c) Multiple scanners.

Figure 2: Synthetic anomaly results.

detect the anomaly (i.e., the score is $\geq \alpha = 2$) only when the anomaly has significantly increased in magnitude. The degree distributions are unaffected by this anomaly.

Bandwidth Flood: In a bandwidth flood, a small number of hosts send large amounts of traffic to a single destination. The key differences with respect to the DDoS attack are that the number of hosts involved is an order of magnitude smaller than the DDoS, and that each attack flow is a single high-volume flow. We vary the number of hosts involved in flooding a single target IP address. The rate of traffic from each host varies uniformly in the range of 300 to 400 Kilobits per second with a fixed packet size of 57 bytes. The target IP is chosen at random within the university with a specific destination port (e.g., port 80 on a webserver). Again, the results were independent of the choice of port and destination host.

Figure 2(b) shows that the behavioral features, FSD, source port, and source address are unaffected. Since each source generates a single flow (and the size of each flow is random), FSD is not effective at detecting this anomaly. As expected, destination ports and addresses exhibit the greatest deviation. However, traffic volume can detect this anomaly just as well given that, at a detection threshold $\alpha = 2$, the total traffic has increased by 25%.

Network Scans: We consider two types of scanning activity: a single host scanning the entire university address space and distributed scanning activity from a set of random source addresses. To model the properties of real scanning activity (e.g., responses probability), we sample 10,000 inbound scan flows to port 445 (associated with many known vulnerabilities [13]) from the traffic trace. We observe that scans receive responses to probes approximately 10% of the time for a flow size of 3 packets, else they are single-packet flows (just a SYN packet). We find that no distribution is able to detect a single scanner. Even with a host scanning 200 hosts per second, which is approximately 6% of total traffic, no distribution is able to detect it. To detect such isolated scan activity, more fine-grained per-host analysis (e.g., flag any host contacting more than X unique destinations in Y seconds [1]) and incorporating other aspects of scanning behavior (e.g., failed connections [6]) are necessary.

In a coordinated scan, multiple hosts (e.g., part of a botnet) scan a particular network (e.g., worm or botnet activity). Each participating host scans at a low rate to avoid detection.[7] We fix the scan rate to 30 hosts per second, and vary the number of hosts generating scanning activity. Figure 2(c) shows that we need to introduce an additional 10% of the total flows before the wavelet score reaches $\alpha = 2$. This implies that even in coordinated scans, entropy-based anomaly detection may not be sufficient.

We also explored several synthetic port scans. The results are similar to the network scans, except that the degree-based metrics are ineffective. A single scanner with a moderate scan rate (30 scans per second) is not detected by any of the entropy metrics. With an increased scan rate (> 1000 scans per second), FSD is the only metric that detects the port scan.

5. IMPLICATIONS

Choice of features: Our analysis suggests that the selection of traffic distributions in entropy-based anomaly detection should be made judiciously, and in particular we should look beyond simple port and address based distributions. The results also suggest a natural approach for choosing traffic features: select traffic distributions that complement one another and provide different views into the underlying traffic structure. For example, the behavioral distributions and the FSD, which are qualitatively different from the port and address distributions, provide distinct and often better anomaly detection capabilities. These complementary distributions can also detect multiple anomalies that occur simultaneously (Sections 3 & 4) .

Computing distributions: Section 3.3 shows that a unidirectional traffic accounting can introduce biases in computing traffic distributions. Obtaining bi-directional measurements may involve additional overhead and instrumentation of current traffic monitoring infrastructure, but recent thrusts for bi-directional flow export [17] may help. This

[7]Worm outbreaks produce similar behavior. With a random scanning worm, the number of incoming scans is $InfectedHosts * ScanRate * \frac{InternalAddressSpace}{TotalIPAddressSpace}$.

may be difficult in large networks when traffic monitors are distributed across the network or when the traffic rates are high enough that sampling is necessary. For enterprise networks, however, the monitor is often a single vantage point co-located with the border router; in this case, bi-directional semantics are easier to obtain and should be preferred.

Leveraging correlations for anomaly detection: The stability of correlations in the entropy values during normal time periods suggests a new anomaly detection technique. We compute the correlations over a finite time window T and detect anomalies over the timeseries of correlation values (computed over a sliding window of size T). In the CMU-2005 dataset, we observe that almost all non-trivial anomalies significantly decrease the entropy correlations between source address (port) and destination address (port) pairs. Additionally, new events are introduced which do not manifest in the wavelet deviation scores. Exploring this observation to strengthen anomaly detection is an interesting avenue for future work.

6. RELATED WORK

The use of entropy and distributions of traffic features has recently received a lot of attention. Feinstein et al. [5] consider the use of entropy of the distribution of source addresses seen at a network ingress point for DDoS detection. Lakhina et al. [10] augment the PCA framework with entropy based metrics and show that this detects anomalies that cannot be identified using volume based analysis alone. These approaches show the promise of entropy-based anomaly detection. Our work studies the selection of traffic feature distributions for entropy based anomaly detection.

Lee and Xiang [12] propose information-theoretic measures for intrusion detection. Entropy has also been to automatically cluster traffic patterns [20]. Wagner et al. [18] use entropy for worm detection by evaluating the compressibility of flow data during attacks.

There is a large body of work related to the accuracy of estimating distributional properties. These include the work on streaming algorithms for estimating the flow size distribution [9] and distribution entropy [11]. Brauckhoff et al. [4] evaluate how packet sampling affects the fidelity of entropy based anomaly detection and show that sampling does not affect the accuracy of detecting the Blaster worm [13].

7. CONCLUSIONS

Entropy-based methods have recently been suggested as good candidates for fine-grained anomaly detection and traffic classification. The goal of our measurement study was understand the analysis and detection capabilities provided by different entropy based metrics.

We find that the port and address distributions are strongly correlated both in their entropy timeseries and detection capabilities. The behavioral metrics (in- and out-degree) and the flow size distribution provide detection abilities that are distinct from other distributions. Using synthetic anomalies, we further confirmed that the port and address distributions have limited utility in anomaly detection: they are ineffective for scanning attacks, and the flood anomalies they detect are large enough to be volume anomalies.

Our results have two main implications. First, we should look beyond port and address distributions for fine-grained anomaly detection. In particular, we should consider distributions that complement each other in their detection capabilities. Second, to avoid the biases arising from unidirectional auditing, it is prudent to use bi-directional flow abstractions for computing traffic distributions.

Acknowledgments

We thank Nick Feamster for providing us access to the additional datasets used in our evaluation, as well as Mukarram Bin Tariq for helping us parse the data. We also thank John Payne for helping us understand the correlations between the entropy timeseries. This work was supported in part by grant CNS-0619525 from the National Science Foundation.

8. REFERENCES

[1] Snort. `http://www.snort.org`.
[2] Argus. `http://qosient.com/argus/`.
[3] Barford, P., Kline, J., Plonka, D., and Ron, A. A signal analysis of network traffic anomalies. In *Proc. of IMW* (2002).
[4] Brauckhoff, D., Tellenbach, B., Wagner, A., Lakhina, A., and May, M. Impact of traffic sampling on anomaly detection metrics. In *Proc. of ACM/USENIX IMC* (2006).
[5] Feinstein, L., Schnackenberg, D., Balupari, R., and Kindred, D. Statistical Approaches to DDoS Attack Detection and Response. In *Proc. of DARPA Information Survivability Conference and Exposition* (2003).
[6] Jung, J., Paxson, V., Berger, A. W., and Balakrishnan, H. Fast Portscan Detection Using Sequential Hypothesis Testing. In *Proc. of the IEEE Symposium on Security and Privacy* (2004).
[7] Karamcheti, V., Geiger, D., Kedem, Z., and Muthukrishnan, S. Detecting malicious network traffic using inverse distributions of packet contents. In *Proc. of ACM SIGCOMM MineNet* (2005).
[8] Kazaa. `www.kazaa.com`.
[9] Kumar, A., Sung, M., Xu, J., and Wang, J. Data streaming algorithms for efficient and accurate estimation of flow distribution. In *Proc. of ACM SIGMETRICS* (2004).
[10] Lakhina, A., Crovella, M., and Diot, C. Mining anomalies using traffic feature distributions. In *Proc. of ACM SIGCOMM* (2005).
[11] Lall, A., Sekar, V., Xu, J., Ogihara, M., and Zhang, H. Data streaming algorithms for estimating entropy of network traffic. In *Proc. of ACM SIGMETRICS* (2006).
[12] Lee, W., and Xiang, D. Information-theoretic measures for anomaly detection. In *Proc. of IEEE Symposium on Security and Privacy* (2001).
[13] Morrison, J. Blaster revisited. ACM Queue vol. 2 no. 4, June 2004.
[14] Cisco Netflow. `http://www.cisco.com/warp/public/732/Tech/nmp/netflow/index.shtml`.
[15] Nychis, G., Sekar, V., Andersen, D. G., Kim, H., and Zhang, H. An Empirical Evaluation of Entropy-Based Traffic Anomaly Detection. Tech. Rep. CMU-CS-08-145, Computer Science Department, Carnegie Mellon University, 2008.
[16] Phaal, P., Panchen, S., and Mckee, N. InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks. RFC 3176, 2001.
[17] Trammell, B., and Boschi, E. Bidirectional Flow Export Using IP Flow Information Export (IPFIX). RFC 5103, 2008.
[18] Wagner, A., and Plattner, B. Entropy Based Worm and Anomaly Detection in Fast IP Networks. In *Proc. IEEE WET ICE* (2005).
[19] Xu, J., Fan, J., Ammar, M. H., and Moon, S. B. Prefix-preserving IP Address Anonymization: Measurement-based Security Evaluation and New Cryptography-based Scheme. In *Proc. of IEEE ICNP* (2002).
[20] Xu, K., Zhang, Z., and Bhattacharyya, S. Profiling internet backbone traffic: Behavior models and applications. In *Proc. of ACM SIGCOMM* (2005).