Evaluation of Different Packet Header Data as Signals for Anomaly Detection

Seong Soo Kim
Department of Electrical Engineering
Texas A&M University
College Station, TX 77843-3128, USA
979-845-9578
skim@ee.tamu.edu

A. L. Narasimha Reddy
Department of Electrical Engineering
Texas A&M University
College Station, TX 77843-3128, USA
979-845-7598
reddy@ee.tamu.edu

ABSTRACT
The frequent and large-scale network attacks have led to an increased need for developing techniques for analyzing network traffic. A number of recent studies have proposed measurement based approaches to network traffic analysis. These techniques treat traffic volume and traffic header data as signals or images in order to make analysis feasible. In this paper, we propose an approach based on classical Neyman-Pearson test employed in signal detection theory to evaluate these different strategies. We use both analytical models and trace-driven experiments, and compare the performance of different strategies. Our evaluations on real traces reveal differences in the effectiveness of different traffic header data as potential signals for traffic analysis in terms of their detection rates and false alarm rates. Our results show that address distributions and number of flows are better signals than traffic volume for anomaly detection. Our results also show that statistical techniques can be sometimes more effective than the NP-test when the attack patterns change over time.

Categories and Subject Descriptors
C.2.0 [**Computer-Communication Networks**]: General---security and protection; C.2.3 [**Computer-Communication Networks**]: Network Operations---network monitoring; C.4 [**Performance of Systems**]: Measurement Techniques, Modeling Techniques, Performance Attributes; G.3 [**Probability and Statistics**]: Time Series Analysis; I.4.8 [**Image Processing and Computer Vision**]: Scene Analysis; I.5.2 [**Pattern Recognition**]: Design Methodology---classifier design and evaluation.

General Terms
Security, Measurement, Performance, Detection theory, Statistics.

Keywords
Network Traffic Analysis, Traffic Engineering, Image Processing, Scene Analysis, Network Anomaly Detection, Estimation and Detection.

1. INTRODUCTION
Malicious network traffic such as worms and other malicious code has become common threat in the Internet. A number of recent studies have pointed to the need for fast detection of such worms for any effective mechanisms for thwarting such traffic [20]. If the spread of such malicious traffic can be detected in real-time, we can alleviate and possibly contain such malicious traffic. Currently, there are many well-known automated self-propagating codes that can be classified as denial of service (DoS), distributed denial of service (DDoS), and distributed reflection denial of service (DRDoS) attacks. Compound attacks consisting of more than one strategy, self-modifying worms, and encryption based worms are likely to increase this threat further in the future [21].

Traditionally, Intrusion detection system (IDS) tools that rely upon operating system logs, process behaviors and firewall logs have been employed to monitor the network traffic. IDS tools monitor network and host traffic to filter the packets that belong to attacks with known behavior patterns. However, the frequent appearance of novel attacks compromises such analysis making true detection of unknown malicious traffic difficult. Measurement-based IDS tools and network traffic analysis have recently started attracting attention as a potential complementary approach [1, 3, 4, 5]. In this paper, we focus our attention on measurement-based approaches to anomaly detection.

Measurement-based tools analyze network traffic to observe statistical properties of traffic. Based on such measurements and some acceptable thresholds on normal network behavior, these tools try to classify traffic as normal or anomalous. Recently, a number of studies have proposed a number of diverse approaches. Some of these studies have treated network traffic as signals, which can

be processed and analyzed to detect anomalies. These studies have considered traffic volume [3, 14, 4, 25], number of flows [5], address and port number distributions [24] as potential signals that can be analyzed in order to detect anomalies in network traffic. The traffic headers carry various pieces of information such as port numbers, protocol numbers which can be further treated as potential signals for analysis.

While all these signals have been shown to be useful for analyzing network traffic, so far, no comprehensive study has been carried out about the relative usefulness of these traffic signals. Which signals are more effective for detecting anomalies? Which signals provide low false alarm rates? Different studies have employed different analysis techniques and different traces making such a comparison difficult. In this paper, we propose to employ classical detection theory based NP-test to study the relative effectiveness of various pieces of information in traffic headers for detecting traffic anomalies.

Our approach employs the Neyman-Pearson (NP) test from the classical detection theory. In our approach, we treat normal traffic as noise and attack traffic as containing the signal along with noise and employ the NP-test to detect attacks in network traffic. NP-test is known to be optimal and hence can be employed to reveal the inherent strengths of the various traffic header signals. We also employ statistical analysis of the traffic data to compare its effectiveness against the NP-test.

The rest of the paper is organized as follows. In section 2, we discuss related work. In section 3, we introduce the various traffic signals that have been proposed for analysis and anomaly detection. In section 4 and 5, we outline our approach for evaluating the effectiveness of these various signals. Our analysis is based on a number of real-world traces. In section 6 and 7, we present the results from our study. Section 8 concludes and provides directions for future work.

This paper makes the following significant contributions: (a) provides a comprehensive evaluation of effectiveness of a number of signals derived from network traffic headers, (b) provides an approach based on classical NP-test for evaluating the effectiveness of network traffic signals, (c) provides data from a number of real-world traces to compare the different traffic signals and (d) shows that statistical techniques can be as effective or sometimes more effective than the NP-test because of changing attack patterns.

2. RELATED WORK

A number of popular measurement tools such as FlowScan, Cisco's FlowAnalyzer, and AutoFocus [1], are used as traffic analyzers. FlowSan is open source software to gather and analyze network flow data taken from NetFlow records of Cisco routers [5]. In the FlowScan, *cflowd* writes raw flow files that wait to be post-processed by *flowscan* for providing against heavy-traffic or flood-based DoS attacks. However, excessive backlog of flow files may cause difficulties in real-time analysis. Using FlowScan, characterization of anomalous network traffic behavior can be described at the flow level [6].

Recently traffic volume metrics, such as byte counts and packet rates, have been analyzed using wavelets to detect anomalies in network traffic [3]. Work in [4] has considered correlation of addresses as a signal for analysis and anomaly detection. While earlier work analyzed traffic as a time series of a single variable, recent work analyzes distributions over different domains of packet header data, particularly the address space and port number space [24]. Sketch-based techniques are shown to perform close to that of per-flow methods for network traffic analysis [8].

Recent studies have shown that Gaussian approximation should work well for aggregated traffic if the level of aggregation in the number of traffic sources and observed time scales is high enough so that individual sources are swallowed according to Central Limit Theorem [7]. Work in [4] has shown the possibility of analysis of WSS (wide-sense stationary) property in network traffic. Based on these results, if we appropriately select the sampling rates, normally distributed and stationary signals can be generated from traffic and hence make statistical analysis of signals feasible.

3. Traffic Signals

We broadly categorize anomaly detection schemes into two groups. The two groups are based on the amount of information maintained or kept per sample. The scalar signals keep track of a single variable (such as traffic volume) as a time series. The vector signals keep track of a vector (1 or 2 dimensions in this paper) of values over a domain of traffic header data (such as addresses, protocol numbers etc.). In order to keep the paper self contained and to provide a basis for our work here, we briefly outline the various traffic signals below before we outline our approach for evaluation.

3.1 Scalar Signals

The scalar signals typically employ the traffic volume such as byte counts, packet counts and the number of flows. Many of the commonly exploited malicious attacks are based on high-bandwidth floods, or other repetitive streams of packets. Whether individual packets are malicious or otherwise, flood-based attacks have been

used for staging DoS attacks. Work in [3, 4] has analyzed traffic volume, measured in byte counts, packet counts and flow counts, using wavelets to detect anomalies in network traffic. For reasons of scalability, we look at the aggregate traffic volumes at the observation point (not per-flow measurements).

3.1.1 Byte Counting

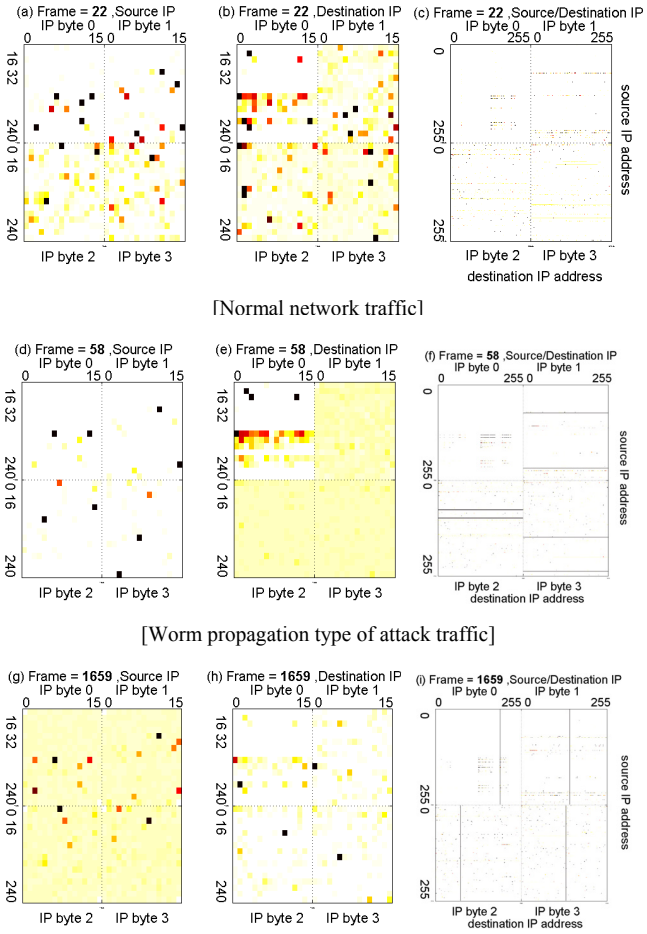
This approach simply counts the number of bytes of traffic seen at the observation point during each sample. The traffic volume in bytes ***b(t)*** at each sample *t* is constructed as a time varying signal that can be analyzed to detect anomalies. Sudden variations in ***b(t)*** or traffic beyond normal statistical bounds can be considered as anomalies. Byte counting has *O(1)* processing cost per packet and *O(1)* storage cost per sample.

3.1.2 Packet Counting

In packet counting, traffic volume again is measured at each sample, but now in terms of packets. The traffic volume in packets ***p(t)*** is the time-varying signal that is analyzed to detect anomalies. Packet counts can be more useful than byte counts when links are carrying traffic to the capacity most of the time. Most of automated self-propagating codes use constant size packets and hence it is possible the packet counts can change as a result of an attack compared to normal traffic. Packet counting has *O(1)* processing cost per packet and *O(1)* storage cost per sample.

3.1.3 Flow counting

This approach counts the number of flows at the observation point. A flow can be classified by the 5-tuple (or by another definition) of (source address, source port, destination address, destination port, protocol number). During each sampling period, the number of such distinct tuples are counted to generate the flow signal ***f(t)***. Flow counting is inherently more difficult than byte counting or packet counting. Hashing and other such techniques would be required to reduce the 5-tuple space to a single flow count number. At the end of each sample, we would need to scan the flow space to count the number of distinct flows seen during the current sample. The costs of such scheme would depend on the hashing techniques employed, the number of tuples chosen to define a flow and the number of flows likely to be seen at the observation point. Hashing can be accomplished in *O(1)* time per packet, individual flow observations can be marked in *O(1)* assuming no hash collisions and the number of flows can be counted in *O(n)* time, where *n* is the size of the flow space. More sophisticated approaches based on bloom filters can be employed to reduce the cost of such approaches to *log(n)* or *loglog(n)* [22, 23].



The (a), (d) and (g) sub-pictures show the intensity of network traffic of the source IP addresses in each frame. The (b), (e) and (h) sub-pictures show that of destination IP addresses. Each quadrant corresponds to each byte in IP address structure. The color of each pixel shows the intensity of traffic at the source or destination, and the descending order of intensity is black, red, orange, yellow and white. The (c), (f) and (i) sub-pictures show the intensity of network traffic of the (source, destination) pair in 2-dimension simultaneously. The x-axis corresponds to the distribution of the destination IP addresses, and y-axis does that of the source addresses. In each quadrant, source and destination addresses consist of 256-by-256 pixels.

Figure 1. Illustrations of address-based image signal

The number of flows could vary from the norm due to DDoS attacks, wide-scale worm propagation etc. It is expected through monitoring changes in the number of flows, it would be feasible to recognize such anomalies. FlowScan analyzes, visualizes and reports Internet traffic flow profiling on flow-centric measurements [5]. Using FlowScan, characteristics of network traffic flow anomalies are illustrated at flow level [6].

3.2 Vector Signals

Vector signals maintain a vector of data for each sample. This requires more space per sample, but allows more sophisticated analysis of traffic header data. We

investigate a new signal based on the protocol numbers. Additionally, we also investigate approaches based on representing network traffic as images. According to employed packet header data and observation domain, we categorize the image-based signals into address-based, flow-based and port-based signals.

3.2.1 Protocol composition

This approach is based on the observation that during the attacks, the protocol employed by the attack traffic should see considerably more traffic than during normal traffic. For example, the recent SQL Slammer worm infected over 90% of the vulnerable hosts employing UDP protocol. In this approach, the amount of ICMP traffic ***i(t)***, TCP traffic ***t(t)***, UDP traffic ***u(t)*** and ETC. traffic ***e(t)*** is monitored as a fraction of the total traffic volume at each sampling interval. Because the proportion of each protocol in traffic is closely interrelated to each other, the increase of a proportion of one protocol makes the proportions of other protocols to decrease. Hence, an abrupt increase or decrease of the proportion of traffic of a protocol can indicate anomalies. Protocol composition has *O(1)* cost per packet and *O(n)* storage cost per sample, where *n* is the number of protocols monitored.

3.2.2 Image-based Signals

In this approach, traffic distribution in a domain is used as a signal that can be analyzed. First, the traffic volume, such as normalized packet counts and the number of flows, is measured along the packet header domain, such as IP addresses and port numbers. Each resultant traffic datum is converted to corresponding pixel intensity in image representation of traffic in the chosen domain. For example, traffic volume can be counted based on destination port numbers. Since the IP port number field is 16 bits, we would obtain 2^{16} = 64K values for each sample indicating the traffic distribution in the port number domain. For reducing the storage and computation complexity, the traffic header domain can be processed in byte-segments which separate out each byte of the IP address (or the port number) as shown in Fig. 1. The image-based signals then originate from the distribution of pixel intensity in each byte of the chosen domain. Based on the kinds of traffic data and the header domain, we categorize the image-based signals into address-based, flow-based and port-based signals. Address-based signal employs traffic volume distribution over address domain (either source address alone, or destination address alone, or a 2-dimensional source and destination address domain). Flow-based signal employs flow number distribution over address domain(s). Port-based signal employs traffic volume distribution over port number domain. (Note: We plan to consider the fourth possibility of employing the flow number distribution over port number domain in the future.)

Image-based signals require two samples of packet header data 2**P*, where *P* is the size of the sample data. We also maintain summary information (DCT coefficients etc.) over a larger number of samples *S*, for statistical evaluation of the current data sample. So, the total space requirement is *O(P+S)*. In our example of source address domain analysis, *P = 4*256 (4 bytes of IP address * 256 values for each byte) = 1024*. DCT based image analysis requires *O(P+S)* processing.

Much of the work on image-based signals draws from the large body of work in image processing and video analysis [9,10,11,12]. These techniques enable the detection of abrupt transitions in images or enable the detection of traffic anomalies.

4. NP-test

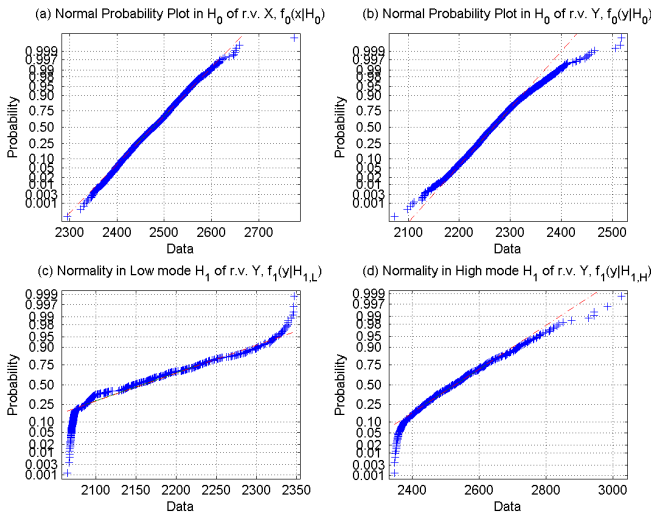
In order to compare the inherent strengths of various signals, we employ the classical and well-established detection theorem, Neyman-Pearson (NP) detection. Briefly, here the normal traffic can be seen as noise and attack traffic can be seen to contain a signal (along with noise) that is of interest that needs to be detected. In this section, we introduce the foundations of the considerably promising detection theory principles into the anomaly detection space in network traffic.

NP-test is optimal and works with any distribution of the underlying hypotheses. As explained below, NP-test employs apriori classified datasets for modeling the distributions of the noise and the signal. For anomaly detection purposes, NP-test would require samples of both normal traffic and attack traffic.

4.1 PDF of H_0 and H_1

In the binary hypothesis testing problem, each of two outputs corresponds to one of two statistical hypotheses [15], the null mode (H_0) and alternative mode (H_1), and an observed datum in the observation space maps into one of the hypotheses. The former null hypothesis represents normal network traffic mode with only noise (N), and the latter alternative hypothesis represents any attack modes with noise and signal (N+S). The probabilistic transition mechanism generates observed data according to two prior conditional probability densities, $p(X=x|H_0)$ and $p(X=x|H_1)$, where X is a random variable denoting the observation. We can define the sample space of one-lateral signals, such as scalar signals, as a random variable X. Similarly the two-lateral signals, such as image-based signals, can be defined on two random variables X (source domain) and Y (destination domain).

The NP-test requires these density functions to be known. To implement this theorem, the total sample



If the data comes from a normal distribution, the plot will appear linear. Other probability functions will introduce curvature in the plot.

Figure 2. Normality of address-based signals in real-time mode.

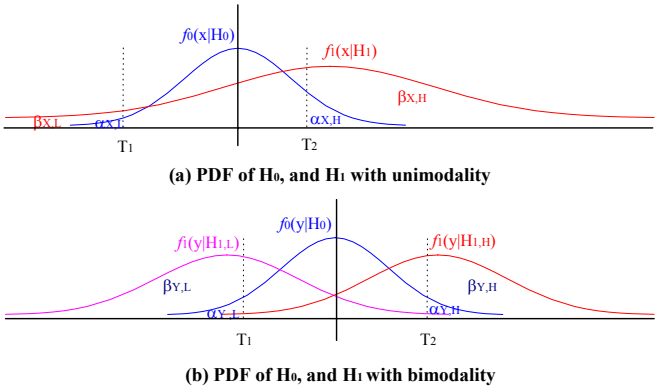
space S on the real traces is divided into two parts, S_0 and S_1. Observations that fall into S_0 elicit the H_0 hypothesis, and observations that fall into S_1 elicit the H_1 hypothesis.

For accurately detecting the anomalous behavior, it requires a solid model of normal behavior. We look at some statistical properties of aforesaid feasible signals in the normal mode. Based on the probability distribution, we assume that the short-term network traffic S_0 exhibits approximately normal distribution. For example, two random variables X and Y in the address-based signal are distributed with approximately Gaussian distribution as shown in Fig. 2(a) and 2(b). We verify that observation space S_0 has a normal distribution at 5% significance level through the Lilliefors test, namely H_0: $X \sim N(\mu_{XN}, \sigma_{XN}^2)$ in scalar signals and source domain of image-based signals, and H_0: $Y \sim N(\mu_{YN}, \sigma_{YN}^2)$ in destination domain of image-based signals. The probability density functions (PDF) of H_0 can be expressed as follows.

$$f_0(x|H_0)=\frac{1}{\sigma_{XN}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_{XN}}{\sigma_{XN}}\right)^2\right] \quad (1\text{-}1)$$

$$f_0(y|H_0)=\frac{1}{\sigma_{YN}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_{YN}}{\sigma_{YN}}\right)^2\right] \quad (1\text{-}2)$$

Similarly, to model the distribution of abnormal traffic S_1, we excerpt only trace with attacks as samples and investigate the statistical measures. In the case of traffic volume-based signals, the distribution of traffic under H_1 could be considered as approximately normal distribution with large variance. Source domain in image-based signal has also unimodality as $f_1(x|H_1)$ in Fig. 3(a). In the case of



(a) PDF of H_0, and H_1 with unimodality

(b) PDF of H_0, and H_1 with bimodality

Figure 3. Illustration of PDF and true / false positive rates

destination domain, however, the PDF shows shape close to a bimodal distribution as $f_1(y|H_{1,L})$ and $f_1(y|H_{1,H})$ in Fig. 3(b). These two separated modes are located in the tail of the normal distribution of H_0. Each of the modes can be locally modeled to have a rough normal distributed component. For instance, Fig. 2(c) and 2(d) illustrate two normal probabilities of abnormal traffic in address-based signal Y. The histogram indicates that the data might be appropriately fitted with a mixture of two normal distributions with the different locations and standard deviations [16, 17].

$$f_1(y|H_1)=\varepsilon\phi_1+(1-\varepsilon)\phi_2$$
$$=\varepsilon f_1(y|H_{1,L})+(1-\varepsilon)f_1(y|H_{1,H}) \quad (2)$$

, where ε is mixing proportion

ϕ_1, ϕ_2 are normal PDFs with location and scale parameters $\mu_1,\sigma_1,\mu_2,\sigma_2$.

The mixing proportion (between 0 and 1) can be fitted using either least squares or maximum likelihood. We set the contamination factor from likelihood of the histogram. Through analysis of sample space, we embody the distribution of H_1, namely H_1: $X \sim N(\mu_{XN} + \mu_{XS}, \sigma_{XN}^2 + \sigma_{XS}^2)$ in scalar signals and source domain of image-based signals, and H_1: $Y_L \sim N(\mu_{YN} + \mu_{YLS}, \sigma_{YN}^2 + \sigma_{YLS}^2)$ and $Y_H \sim N(\mu_{YN} + \mu_{YHS}, \sigma_{YN}^2 + \sigma_{YHS}^2)$ in destination domain. The PDF under H_1 can be expressed as follows.

$$f_1(x|H_1)=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_{XN}^2+\sigma_{XS}^2}}\exp\left[-\frac{1}{2}\frac{(x-(\mu_{XN}+\mu_{XS}))^2}{\sigma_{XN}^2+\sigma_{XS}^2}\right] \quad (3\text{-}1)$$

$$f_1(y|H_1)=\varepsilon*\frac{1}{\sqrt{2\pi}\sqrt{\sigma_{YN}^2+\sigma_{Y_LS}^2}}\exp\left[-\frac{1}{2}\frac{(y-(\mu_{YN}+\mu_{Y_LS}))^2}{\sigma_{YN}^2+\sigma_{Y_LS}^2}\right] \quad (3\text{-}2)$$
$$+(1-\varepsilon)*\frac{1}{\sqrt{2\pi}\sqrt{\sigma_{YN}^2+\sigma_{Y_HS}^2}}\exp\left[-\frac{1}{2}\frac{(y-(\mu_{YN}+\mu_{Y_HS}))^2}{\sigma_{YN}^2+\sigma_{Y_HS}^2}\right]$$

4.2 Bayes' Likelihood Ratio Test

The Bayesian criterion method assumes that the two outputs are governed by apriori probabilities, π_0 and π_1, and that a cost is assigned to each of the four outcomes. These costs are denoted by C_{00}, C_{10}, C_{11}, and C_{01}, where the first subscript indicates the hypothesis accepted and the second indicates the unknown truth. These outcomes respectively map to true negative, false positive, true positive and false negative.

Bayesian criterion leads to likelihood ratio test [18], where a hypothesis is accepted when it is sufficiently likely relative to the other hypothesis. The optimal test is a threshold test of the likelihood ratio. The notion of using the magnitude of the ratio of two PDFs as the basis of a best test will help to provide an intuitively appealing method of constructing a test of a null hypothesis against an alternative hypothesis. The test is defined in scalar signals X and source domain X of image-based signals as follows.

$$\Lambda(x)=\frac{f_1(x|H_1)}{f_0(x|H_0)}=\frac{\frac{1}{\sqrt{2\pi}\sqrt{{\sigma_{XN}}^2+{\sigma_{XS}}^2}}\exp\left[-\frac{1}{2}\frac{\left(x-(\mu_{XN}+\mu_{XS})\right)^2}{{\sigma_{XN}}^2+{\sigma_{XS}}^2}\right]}{\frac{1}{\sigma_{XN}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_{XN}}{\sigma_{XN}}\right)^2\right]}$$

$$\begin{cases}\text{if } \Lambda(x)\geq\eta, \text{anounce } \mathrm{H}_1\\ \text{if } \Lambda(x)<\eta, \text{anounce } \mathrm{H}_0\end{cases} \quad (4\text{-}1)$$

And detector is defined in destination domain Y of image-based signals as follows.

$$\Lambda(y)=\frac{f_1(y|H_1)}{f_0(y|H_0)}=\frac{\varepsilon^*\frac{1}{\sqrt{2\pi}\sqrt{{\sigma_{YN}}^2+{\sigma_{Y_LS}}^2}}\exp\left[-\frac{1}{2}\frac{\left(y-(\mu_{YN}+\mu_{Y_LS})\right)^2}{{\sigma_{YN}}^2+{\sigma_{Y_LS}}^2}\right]}{\frac{1}{\sigma_{YN}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_{YN}}{\sigma_{YN}}\right)^2\right]}$$

$$+\frac{(1-\varepsilon)^*\frac{1}{\sqrt{2\pi}\sqrt{{\sigma_{YN}}^2+{\sigma_{Y_HS}}^2}}\exp\left[-\frac{1}{2}\frac{\left(y-(\mu_{YN}+\mu_{Y_HS})\right)^2}{{\sigma_{YN}}^2+{\sigma_{Y_HS}}^2}\right]}{\frac{1}{\sigma_{YN}\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_{YN}}{\sigma_{YN}}\right)^2\right]}$$

$$\begin{cases}\text{if } \Lambda(y)\geq\eta, \text{anounce } \mathrm{H}_1\\ \text{if } \Lambda(y)<\eta, \text{anounce } \mathrm{H}_0\end{cases} \quad (4\text{-}2)$$

This value is a random variable, and is tested against the threshold η as

$$\eta=\frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})}=\frac{\pi_0}{\pi_1}, \text{ when } \alpha+\beta=1 \quad (5)$$

Threshold η can be defined by the ratio of *P(H_0 is true)* and *P(H_1 is true)*, which are to be determined from a prior knowledge. If the likelihood ratio is greater than η, the detection output is H_1, otherwise the output is H_0

In many cases, it may be difficult to determine the costs or a prior distribution $\Pi = (\pi_0, \pi_1)$, and we have serious difficulties in setting proper threshold. The Neyman-Pearson (NP) test bypasses these factors by introducing the conditional probabilities as (6) and (7). For a practical fidelity criterion, given a constrained significance level α (i.e., false alarm rate) we can derive the threshold η of the test which correspondingly renders the maximum detection rate β. Appendix A-1 shows the relationship between thresholds and measurement criteria during real-time analysis of flow-based signal.

To evaluate our approach against the NP-test, we calculate the false alarm rate and the detection rate based on a given threshold. We set appropriate thresholds of the NP-test so as to correspond to statistical threshold of 3σ. Given the threshold, we solve the Bayesian detector in (4-1) and (4-2), and derive critical regions (Z_1) of either boundary (T_1 and T_2) in Fig. 3. These critical regions are close to those of 3σ-based method due to normality.

4.3 Expected True and False positive rates

We can define the false alarm rate α (type I error) as the overall probability that H_0 is actually true and likelihood ratio test detects H_1 as (6-1) and (6-2) from blue-colored PDFs of Fig. 3.

$$\alpha_X=\int_{Z_1} f_0(x|H_0)\,dx=\int_{-\infty}^{T_1} f_0(x|H_0)\,dx+\int_{T_2}^{\infty} f_0(x|H_0)\,dx \quad (6\text{-}1)$$

$$\alpha_Y=\int_{Z_1} f_0(y|H_0)\,dy=\int_{-\infty}^{T_1} f_0(y|H_0)\,dy+\int_{T_2}^{\infty} f_0(y|H_0)\,dy \quad (6\text{-}2)$$

, where Z_1 is critical region : $[-\infty,T_1]+[T_2,\infty]$

And the detection rate β in scalar signals and source domain of image-based signals with unimodality is defined as the probability that we successfully detect the anomalies, i.e., H_1 is true and the likelihood ratio test detects H_1 as (7-1) from red-colored PDF of Fig. 3(a). Consequently false negative rate (type II error) is calculated as 1-β. Similarly, the true positive rates in destination domain of image-based signals with bimodality can be defined as (7-2) from red/pink-colored PDFs of Fig. 3(b).

$$\beta_X=\int_{Z_1} f_1(x|H_1)\,dx=\int_{-\infty}^{T_1} f_1(x|H_1)\,dx+\int_{T_2}^{\infty} f_1(x|H_1)\,dx \quad (7\text{-}1)$$

$$\beta_Y=\int_{Z_1} f_1(y|H_1)\,dy=\varepsilon^*\left[\int_{-\infty}^{T_1} f_1(y|H_{1,L})\,dy+\int_{T_2}^{\infty} f_1(y|H_{1,L})\,dy\right]+(1-\varepsilon)^*\left[\int_{-\infty}^{T_1} f_1(y|H_{1,H})\,dy+\int_{T_2}^{\infty} f_1(y|H_{1,H})\,dy\right] \quad (7\text{-}2)$$

The objective of the NP-test is to make α as small as possible and β as large as possible. To accomplish this objective, α is constrained by a given tolerable lower

bound, and β is maximized using Lagrange multipliers. From derived density function and given thresholds, we can induce the expected true positive rates and false positive rates of each feasible traffic signal.

4.4 Application of traffic signals in NP-test

We use real-time image-based signal of flow distribution in destination address domain in Table 5 for explanation of how NP-test is to be applied. Through analysis of image-based signals, the distribution of H_0 has a normal distribution at 5% significance level, namely $Y \sim N(910.9, 146.5^2)$ in destination address. And, we simplify the distribution of H_1, namely $Y_L \sim N(405.9, 45.0^2)$ and $Y_H \sim N(1518.4, 462.2^2)$. The mixing ratio of low mode and high mode is 0.27 and 0.73.

Under given threshold and PDFs which is defined as (1-2) and (3-2), we solve the NP-test as (4-2) and derive critical regions of either boundary.

For destination address variable Y

$$\Lambda(y)=\frac{0.27\frac{1}{45.0\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-405.9}{45.0}\right)^2\right]+0.73\frac{1}{462.2\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-1518.4}{462.2}\right)^2\right]}{\frac{1}{146.5\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-910.9}{146.5}\right)^2\right]}=14.08$$

through numerical analysis

$$y_{1,2}=910\pm425$$

$$\begin{cases} Z_0: 485<y<1335 \\ Z_1: y\le 485, 1335\le y \end{cases} \quad (8)$$

These critical regions are close to those of 3σ-based method, 471<y<1350.

We can compute the false alarm rate α (type I error) as the interval probability distribution from (6-2) as,

$$\alpha_Y \approx \int_{-\infty}^{485}\frac{1}{146.5\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-910.9}{146.5}\right)^2\right]dy+\int_{1335}^{\infty}\frac{1}{146.5\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-910.9}{146.5}\right)^2\right]dy$$
$$\approx 0.0037 \quad (9)$$

Similarly, the detection rate β is calculated as a mixture of two interval probabilities from (7-2) as,

$$\beta_Y \approx 0.27\left\{\int_{-\infty}^{485}\frac{1}{45.0\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-405.9}{45.0}\right)^2\right]dy+\int_{1335}^{\infty}\frac{1}{45.0\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-405.9}{45.0}\right)^2\right]dy\right\}+$$
$$0.73\left\{\int_{-\infty}^{485}\frac{1}{462.2\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-1518.4}{462.2}\right)^2\right]dy+\int_{1335}^{\infty}\frac{1}{462.2\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-1518.4}{462.2}\right)^2\right]dy\right\}$$
$$\approx 0.746 \quad (10)$$

In flow-based signal of the Table 5, in case of destination address domain in real-time, the false alarm rate is about 0.37% and the detection rate is 74.6%. With expected detection rates and false alarm rates, we can evaluate the power of feasible traffic signals.

5. Evaluation Methodology

5.1 Traces

In order to evaluate the different signals mentioned above, we employ three real traffic traces. We name these traces "*Access Link*", "*ISP*" and "*Campus*" based on where these traces are collected from without divulging the source of traces (to preserve author anonymity). The "*Access Link"* trace contained 2 weeks of network traffic data from Oct. 12, 2003 to Oct. 26, 2003, and contained actual worm attacks. Currently *Access Link* member institutions are over 230 organizations, which include 50 government research institutes, 72 universities, 15 industrial research laboratories, etc. *Access Link* trace is a collection of NetFlow trace files collected by the 155Mbps international ATM link. In the trace employed, there are 5 major attacks as described in Table 1 and a few instantaneous probe attacks. It generates 4345 samples in case of 2-minute sampling period. Among the observations, the suspected activities reach to 782 times, which are judged by traffic engineering. Additionally we examine the signals on traces from the *ISP*, which contains real network attacks in the pcap header format. Third, we examine the signals on a live network in *Campus*.

5.2 Measurement Criteria

To quantitatively evaluate different packet header data signals for detecting anomalies, we employ two kinds of measurement criteria, which allow comparative and normative studies. The former is to evaluate the performance of various signals based on the detection rates, false alarm rates and likelihood ratios. The latter is to judge the various signals effectiveness through NP-test based measure which allows the potential of these signals-in-themselves to be compared.

We evaluate the signals through two analyses: first analysis based on statistical properties of the signals and appropriate thresholds (3σ or higher) and the second analysis based on the classical Neyman-Pearson detection methodology.

5.2.1 Type I and II Errors

Measurement-based anomaly detection techniques have

TABLE 1
THE DESCRIPTIONS OF FIVE ATTACKS IN *ACCESS LINK* TRACES

	1	2	3	4	5
Duration	5.3 h	4.5 h	4.1 hours	12.3 h	3.6 h
IP	semi-random	random[a]	random	semi-random	random
Protocol	TCP	UDP	TCP/UDP	TCP/UDP/ICMP	UDP
Port	#80	#1434	random/#1434	#80 / #1434 / #0	#1434
Size	48B	404B	random/ 404B	48B / 404B/ 28B	404B

a. SQL Slammer

to contend with two types of errors.

The true positive (sensitivity or detection β) is the probability that a statistical test will be positive for a true statistic. On the other hand, a type I error (false positive error α) occurs if a difference is declared when the null hypothesis is true. In other words, a false attack alarm is declared when the traffic is normal.

$$\alpha = p(\text{anounce } H_1 \mid H_0 \text{ is true}) = p(\text{detect anomaly} \mid \text{traffic is normal}) \quad (11\text{-}1)$$

$$\beta = p(\text{anounce } H_1 \mid H_1 \text{ is true}) = p(\text{detect anomaly} \mid \text{traffic is anomaly}) \quad (11\text{-}2)$$

The true negative (specificity 1-α) is the probability that a statistical test will be negative for a negative statistic. On the other hand, a type II error (false negative error 1-β) occurs if no difference is declared when the null hypothesis is false. In other words, a false negative is declared that the traffic is normal even when the traffic suffers from attacks.

5.2.2 Likelihood Ratio

To test non-nested complementary hypotheses, the LR (likelihood ratio) and NLR (negative likelihood ratio) are used as follows [26].

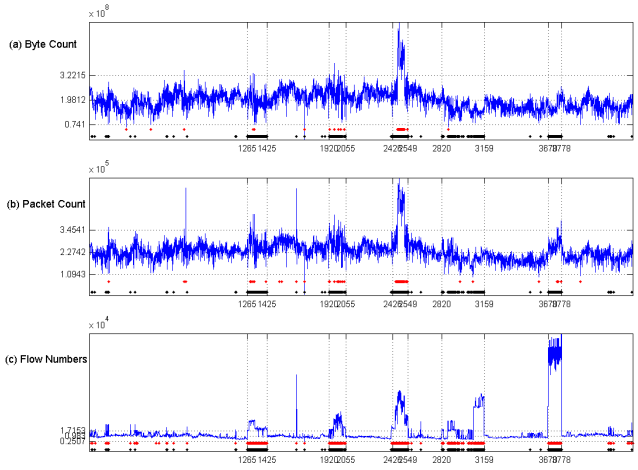
$$\text{LR} = \frac{\text{true positive rate}}{\text{false positive rate}} = \frac{\text{sensitivity}}{\text{1-specificity}} = \frac{\beta}{\alpha} \quad (12)$$

$$\text{NLR} = \frac{\text{false negative rate}}{\text{true negative rate}} = \frac{\text{1-sensitivity}}{\text{specificity}} = \frac{1-\beta}{1-\alpha}$$

Because type I and type II errors are dependent on each other (e.g., as β increases, α increases), we carry out the evaluation experiments upon the same set of real traces for all the different traffic signals. LR and NLR help to estimate the efficient trade-off between the power of detection and false alarm. Ideally, LR is infinity and NLR is zero.

5.3 Statistical analysis based on 3σ

NP-test's requirement of parametric knowledge of distributions of the normal and attack traffic may make the test difficult as new attack patterns emerge. In order to overcome this difficulty, we also employ statistical analysis of the traffic data. Statistical analysis of traffic data requires only a model of normal traffic and hence possibly can distinguish new forms of attacks. We developed a theoretical basis for deriving thresholds for analyzing traffic signals and anomaly detection. For 3σ-based statistical analysis, we set 2 kinds of thresholds, a high threshold T_H and a low threshold T_L. When we respectively set the T_H and T_L thresholds to $\pm 3.0\sigma$ of aforementioned traffic signal distributions in ambient traffic, attacks can be detected with an error rate of 0.3% (if the signal is normally distributed) which can be expected as target false alarm rate. We can judge the current traffic status by calculating the standard intensity deviation of signals in each sampling instant as (13-1) and (13-2).



The red dots located on the top are marked when traffic volume-based signals declare anomalies. The black dots located on the bottom show actual anomalous traffic. The vertical dotted lines mean 5 major attacks. The horizontal dotted lines respectively show the T_H, mean and T_L based on 3σ method of ambient signals.

Figure 4. Trace-driven detection results in scalar signals

$$\textit{traffic status} \begin{cases} \textit{normal}, & \text{if } T_L < \sigma < T_H \\ \textit{attack}, & \text{if } \sigma \le T_L \text{ or } T_H \le \sigma \end{cases} \quad (13\text{-}1)$$

$$P(\mu - 3.0\sigma < X \le \mu + 3.0\sigma) \approx 99.7\% \quad (13\text{-}2)$$

6. EVALUATION OF SIGNALS

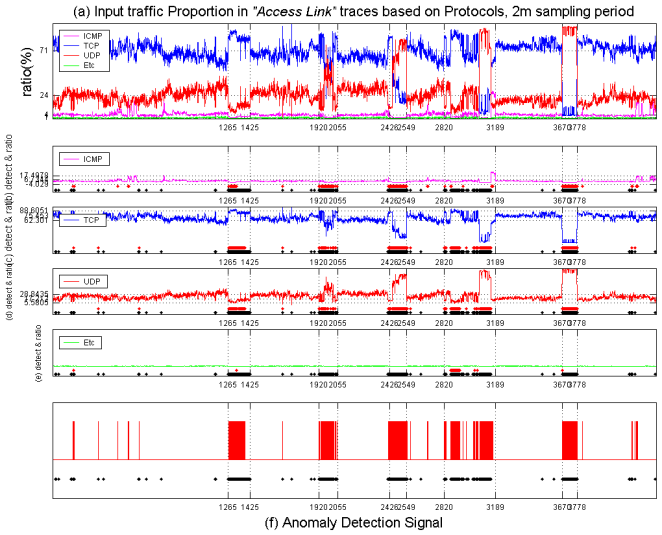
6.1 SCALAR SIGNALS

6.1.1 Byte counting, packet counting

Fig. 4 and Table 2 shows the results of employing scalar signals on the *Access Link* trace for anomaly detection.

Fig. 4(a), 4(b) and the 2nd and 3rd column of the Table 2 demonstrate the detection strength of the traffic volume-based signals. In *Access Link* traces, there are 5 major different styled attacks marked by dotted vertical lines and a few instantaneous probe attacks. As the pictures show, except the 3rd attack, the remaining 4 attacks did not set off any distinguishable variance in traffic volume. SQL Slammer worm located in the 5th attack, for instance, spawned a large number of traffic connections with a single packet of relatively small size of 404 bytes. Compared to large elephant flows, these flows have insignificant prominence in byte and packet counts.

The results show that byte count and packet count signals with statistical thresholds achieved detection rates of 11.0% and 18.3% respectively. This is compared to



The (a) sub-picture shows the proportions of all protocols. The (b) shows the variation of ICMP protocol with time, (c) shows that of TCP, (d) shows that of UDP and (e) shows that of remnant protocols. The above red dots located on the bottom in each sub-picture are marked when each protocol declares anomalies. The below black dots located on the bottom show real anomalies. The horizontal dotted lines respectively show the average proportion of each protocol. The figure (f) shows the generated composite attack detection signal in red which is combined from the attack detection of each protocol. The black dots located on the bottom show actual anomalous traffic.

Figure 5. Detection results based on protocol composition.

feasible detection rates of 35.3% and 32.4% with the NP-test based analysis of the same signals. The byte count signal is slightly better; however, the difference between the two measurements is marginal. These results show that the traffic volume signals may not be adequate for providing reliable signals for anomaly detection.

6.1.2 Flow counting

Flow counting shows significantly better performance in detecting anomalous traffic as shown in Fig. 4(c) and the 4th column of the Table 2. Flow counting with statistical thresholds achieved a detection rate of 95.1% at a false alarm rate of 0.73%. Flow counting is clearly superior to both byte counting and packet counting signals. Flow counting with NP-test based analysis provided a detection rate of 91.7% with an accompanying false alarm rate of 0.24%. NP-test based analysis achieved a smaller false alarm rate with an accompanying loss in detection rate compared to the statistical approach.

We study the flow counting signal further in section 7.1 because of its promise based on the results in Table 2.

Table 2. Results of scalar signals[1]

Signals	T.P. β [1]	F.P. α [2]	*NP β* [3]	*NP α* [4]	LR [5]	NLR [6]
Byte count	11.0% 86/782	0.11% 4/3563	 *35.3%*	 *0.15%*	98.0/ *241.8*	0.89/ *0.65*
Packet count	18.3% 143/782	0.25% 9/3563	 *32.4%*	 *0.16%*	72.4/ *206.8*	0.82/ *0.68*
Flow number	95.1% 744/782	0.73% 26/3563	 *91.7%*	 *0.24%*	130.4/ *384.6*	0.05/ *0.08*

1. True Positive rate by 3σ-based statistical analysis
2. False Positive rate by 3σ-based statistical analysis
3. *expected true positive rate by Neyman-Pearson test*
4. *expected false positive rate by Neyman-Pearson test*
5. Likelihood Ratio in measurement by 3σ / *LR in NP-test*
6. Negative Likelihood Ratio by 3σ / *NLR in NP-test*

6.2 VECTOR SIGNALS

6.2.1 Protocol Composition

We employ 2 kinds of thresholds, a high threshold T_H that indicates that the fraction of volume of one of the network protocols increases abnormally and a low threshold T_L indicating that the fraction of the traffic volume of the network protocol decreases inordinately. When each protocol proportion in current input traffic is larger (or lower) than the 3σ of normal distribution for individual protocol, the detector declares anomalies. Fig. 5(f) shows the generated composite attack detection signal which declares an anomaly when 2 out of 4 individual protocol signals declare an anomaly.

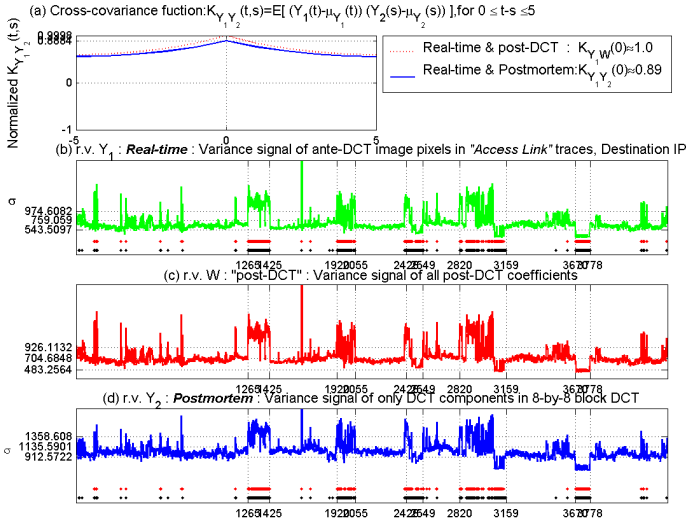
Table 3. Results of protocol composition signals

Signals	T.P. β	F.P. α	*NP β*	*NP α*	LR	NLR
ICMP	72.9% 570/782	1.94% 69/3563	 *72.4%*	 *0.37%*	37.6/ *195.7*	0.28/ *0.27*
TCP	81.0% 633/782	0.42% 15/3563	 *83.7%*	 *0.31%*	192.3/ *270.0*	0.19/ *0.16*
UDP	77.5% 606/782	0.39% 14/3563	 *78.5%*	 *0.26%*	197.2/ *302.0*	0.23/ *0.22*
ETC.	31.7% 248/782	0.00% 0/3563	 *76.8%*	 *0.81%*	∞ / *94.8*	0.68/ *0.23*
Composite signal	80.8% 632/782	0.28% 10/3563	–[1]	–	288.0	0.19

1. Multi-component signal is hard to model in NP-test

Fig. 5 and Table 3 show the measurement results for protocol composition. Protocol composition could achieve a detection rate of 80.8% at a false alarm rate of 0.28% as shown in Table 3. This shows that protocol composition could be a useful signal. Protocol composition signal based detector has lower complexity than the flow counting signal.

[1] Statistical analysis is in non-italic type whereas NP-test is in italic type.



The (b) subpicture shows the standard deviation of pixels in the image in real-time analysis. The (c) and (d) show the standard deviation of all of the DCT coefficients and only 1 most significant DCT coefficient respectively during postmortem analysis. The horizontal dotted lines show the T_H, mean and T_L thresholds based on 3σ method. The bottom red dots in (b) and (d) illustrate the anomaly detection results. The black dots located on the bottom show real anomalies.The (a) subpicture shows the cross-correlation coefficients between these signals.

Figure 6. The trace-driven evaluation results from address-based image signal in destination address for detecting attacks

6.3 Image-Based Signals

6.3.1 Analysis Methods

Our approach to detecting anomalies envisions two kinds of detection mechanism, real-time and postmortem. Real-time analysis may rely on less sophisticated analysis because of the resource demands and imminence of attacks.

6.3.1.1 Real-time analysis

We employ the Discrete Cosine Transform (DCT) for scene change analysis in postmortem. In real-time, we employ the variance of pixel intensities in the image for analysis and anomaly detection. Using the variations of these image signals for deriving thresholds, we can obtain an approximation of the energy distribution of the normalized packet counts within observation domain as follows.

$$\sigma=\left[\frac{1}{N}\sum_{k=1}^{N}(x_k-\bar{x})^2\right]^{\frac{1}{2}} \qquad (14)$$

, where $\begin{cases} x_k \text{ are pixel intensities, N=1024 in real-time} \\ x_k \text{ are DCT coefficients, N=16 in postmortem} \end{cases}$

and $\bar{x}=\frac{1}{N}\sum_{k=1}^{N}x_k$

The detection signal is calculated instantaneously upon sampling instants for real-time analysis. Fig. 6(b) shows the detection results from *Access Link* trace-driven evaluation for 8 days.

6.3.1.2 Postmortem analysis

For post-attack forensic analysis and traffic engineering, the captured images need to be stored. Instead of storing the entire image of each sample, a few DCT coefficients for each sample could be stored. The relationship of two stationary random processes can be estimated using cross-covariance function which is the cross-correlation of mean-removed sequence as follows.

$$K_{XY}(t,s)=E\left[\left(X(t)-\overline{X(t)}\right)\left(Y(s)-\overline{Y(s)}\right)\right] \qquad (15)$$

, where $\overline{X(t)}, \overline{Y(s)}$ are the mean values and E[.] is the expected value operator

If all of the DCT coefficients were used in the detection, this method would be approximately equivalent to the variance of pixels in real-time (as in Parseval's Theorem) as shown in $K_{Y1W}(0)\approx 1.0$ in Fig 6(a), and the variance signal in Fig. 6(b) and 6(c). However, by using only the *n* most significant DCT coefficients, we filter the image and are able to focus on the broader characteristics of each image type. We can vary *n* to see how many coefficients we ought to compare to get optimal results in variance. For simplicity, we can use only the most significant DCT coefficient (*n*=1) in an 8-by-8 block. Using the variance of only DCT component in DCT blocks, we can obtain an approximation of the energy distribution in postmortem analysis as shown in Fig. 6(d).

Instead of performing the computationally intensive task of reconstruction, it is possible to analyze the image by analyzing the DCT coefficients directly. However, by taking only few DCT coefficients, we could potentially perform worse than the reconstruction scheme if the traffic image is not represented well by the retained set. As shown in $K_{Y1\ Y2}(0) = 0.89$ in Fig 6(a) and the approximated variance signal in Fig. 6(d), this technique would classify nearly as well as the reconstruction algorithm.

6.3.2 Packet Distribution in Address Domain

Usually the packet counts at the source addresses of the outbound aggregate traffic can illustrate the distributed properties of the network traffic usage. Meanwhile, the analysis at the destination address of the outgoing traffic can show the concentration of the flow target.

Table 4. Results of Address-based signals

Time	D.	TP β	FP α	*NP β*	*NP α*	LR	NLR
Real-time	SA [1]	81.5% 637/782	0.06% 2/3563	*76.3%*	*0.15%*	1451.2/ *508.7*	0.19/ *0.24*

	DA [2]	87.1% 681/782	0.42% 15/3563	*88.4%*	*0.15%*	206.9/ *589.3*	0.13/ *0.12*
	(SA, DA)[3]	94.2% 737/782	0.48% 17/3563	–	–	197.5	0.06
Post mortem	SA	88.6% 693/782	0.06% 2/3563	*92.0%*	*0.05%*	1578.7/ *1840.0*	0.11/ *0.08*
	DA	80.2% 627/782	0.14% 5/3563	*84.1%*	*0.14%*	571.4/ *600.7*	0.20/ *0.16*
	(SA, DA)	95.9% 750/782	0.20% 7/3563	–	–	488.2	0.04

1. SA stands for Source Address.
2. Destination Address.
3. Source Address and Destination Address in combination.

The results of the address-based signals are shown in Fig. 6 and Table 4. The Fig. 6(b) illustrates that the true positive rate is 87.1 % (681 detected out of 782) and the false positive rate is 0.42% (15 falsely detected out of 3563) for real-time detection based on destination address. And the Fig. 6(d) shows that detection rate is 80.2 % (627 detected) and the false alarm rate is 0.14% (5 false detections) for postmortem analysis, again based on destination address.

NP-test generally shows a little higher performance than statistical analysis with higher LRs and lower NLRs excepting source address in real-time.

The results indicate that the different signals exhibit different strengths in anomaly detection. The destination address based signal performed better in real time and the source address based signal performed better in postmortem when compared with each other. In both of real-time and postmortem analyses, the (source, destination) based signal performed significantly better than the single dimensional signals. However, this signal has a higher storage and processing cost.

6.3.3 Flow Distribution in Address Domain

An analysis of the distribution of flows gives an idea what peer to peer transmissions consist of and how they are distributed over the address domain. This flow-based signal deals with not only the variation of the number of flows, but also with the changes in the distribution of flows.

An analysis of the flow-based image could be effective for revealing flood types of attacks. When a flow is defined as the triple of source address / destination address / destination port, the flood-based attacks spread flows over the destination IP addresses (or ports) in random or dictionary mode style attacks. The distribution of the number of flows in address space would then be expected to be much different from its normal and historical distribution.

Table 5. Results of Flow-based signals

Time	D.	TP β	FP α	*NP β*	*NP α*	LR	NLR
Real-time	SA	90.3% 706/782	0.22% 8/3563	*90.0%*	*0.14%*	402.1/ *663.8*	0.10/ *0.10*
	DA	56.5% 442/782	0.25% 9/3563	*74.6%*	*0.37%*	223.8/ *201.1*	0.44/ *0.25*
	(SA, DA)	92.8% 726/782	0.42% 15/3563	–	–	220.5	0.07
Post mortem	SA	91.6% 716/782	0.20% 7/3563	*91.8%*	*0.14%*	466.0/ *676.6*	0.08/ *0.08*
	DA	52.0% 407/782	0.17% 6/3563	*70.7%*	*0.40%*	305.9/ *178.7*	0.48/ *0.29*
	(SA, DA)	94.8% 741/782	0.20% 7/3563	–	–	482.3	0.05

Results of analysis of flow-based images are shown in Table 5. As shown in the Table 5, the source address based images/signals generally exhibit higher confidence than the destination address based images/signal for detecting traffic anomalies due to bimodality of destination addresses. If source and destination address signals are jointly adopted, we can expect higher confidence in detection rate with a consequent deterioration of the false alarm rate.

The results from Tables 2 and Table 5 can be compared to understand the relative strengths of scalar flow counting signal and the flow-based image signal. It is observed that flow-based images could reduce the false alarm rates. However, it is observed that flow-based images did not improve the detection rates when compared to the scalar flow counting signal. From these results, the significant additional storage and processing cost entailed in flow-based images may not be warranted unless the reduction of the false alarm rate is paramount.

In our experiments, the destination flow-based signals failed to identify the 3rd attack in the *Access link* trace. This attack consists of a concurrent host scan aimed at specific destinations (high threshold), and the SQL Slammer worm which targeted random machines (low threshold). These two simultaneous conflicting attacks complicate the detection by offsetting the address distribution characteristics of each other. As a result, it shows that composite attacks may require multiple signals for analysis. The multidimensional signal is motivated from the grounds.

6.3.4 Packet Distribution in Port Domain

Besides address domain, we could analyze and visualize the packet header information in port number domain. An analysis of the port number-based image can reveal portscan types of attacks. When a machine is the target of a portscan, the distribution of the exploited port numbers would deviate from its normal distribution.

Table 6. Results of Port distribution signals

Time	D.	TP β	FP α	*NP β*	*NP α*	LR	NLR
Real-time	SP [1]	83.4% 652/782	0.14% 5/3563	*94.9%*	*0.07%*	594.1/ *1428.8*	0.17/ *0.05*
	DP [2]	96.2% 752/782	0.17% 6/3563	*90.5%*	*0.14%*	571.1/ *630.4*	0.04/ *0.09*
	(SP, DP) [3]	96.8% 757/782	0.25% 9/3563	–	–	383.2	0.03
Post mortem	SP	93.9% 734/782	0.11% 4/3563	*94.2%*	*0.08%*	836.1/ *1183.9*	0.06/ *0.06*
	DP	95.7% 748/782	0.14% 5/3563	*87.1%*	*0.21%*	681.6/ *406.1*	0.04/ *0.13*
	(SP, DP)	96.0% 751/782	0.17% 6/3563	–	–	570.3	0.04

1. SP stands for Source Port
2. Destination Port
3. Source Port and Destination Port in combination

The results in Table 6 indicate that port-based signal could be a powerful signal for anomaly detection achieving detection rates of up to 96% with very low false alarm rates. The statistical analyses have occasionally outperformed the NP-test based analysis. For example, the destination-port based postmortem signal achieves a detection rate of 95.7% at a false alarm rate of 0.14% compared to the 87.1% detection rate and 0.21% false alarm rate of the NP-test based analysis. The difference could be verified in terms of LR (681.6 vs. 406.1) and NLR (0.04 vs. 0.13).

The simultaneous improvement in both the measures is a result of the nature of attacks which probe accessible ports in random or dictionary fashion for infiltration. It also demonstrates the complexity of correct NP-test modeling due to multimodality.

6.3.5 Multidimensional signal

With three distinct image-based signals, we can analyze the traffic properties of each IP address and port number from multiple and diverse viewpoints. We develop a multi-component image based analysis of traffic signal. With three distinct traffic signals, that are address-based, flow-based and port-based signals, can we improve the rate of detection of the anomalous traffic? We study this issue for both real-time and postmortem analysis. The results are shown in Table 7.

Table 7. Results of Multidimensional signals

Signals	T.P. β	F.P. α	LR	NLR
Real-time (S,D) [1]	97.1% 759/782	0.62% 22/3563	157.2	0.03
Postmortem (S,D)	97.4% 762/782	0.34% 12/3563	289.3	0.03

1. Source Domains and Destination Domains in combination

The results from Table 7 suggest that it is possible to improve the detection rates considerably by considering multidimensional signals with an accompanying higher rate of false alarms compared to the individual components of the signal. Using various aspects of packet header data simultaneously facilitates analysis of different strategies of attacks and the effect of composite anomalous traffic. Even though sophisticated attacks could go undetected when only one-component signal is investigated, it could become possible to detect the anomalies employing other signals.

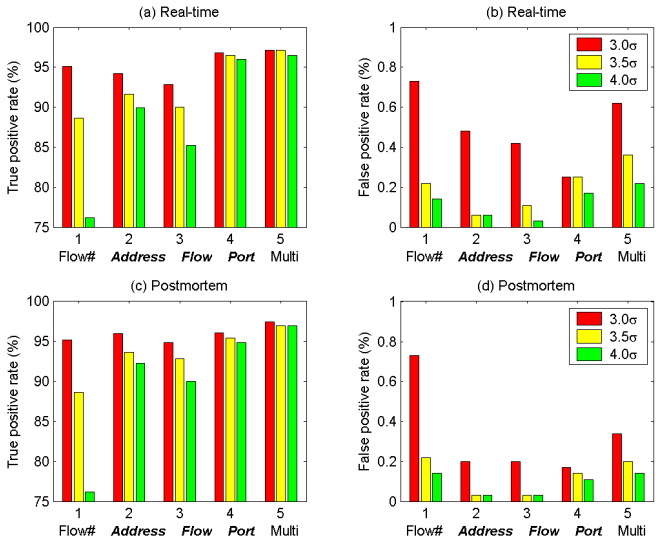


Figure 7. The Relationship between measurements and thresholds.

7. ANALYTICAL RESULTS

7.1 Sensitivity of Signals to Thresholds

In order to evaluate the effectiveness of employing different thresholds, we compare the detection results of schemes employing the image-based analysis with a scalar signal, especially the number of flows. The anomaly detection measurements in combination of source and destination domains are shown in Fig 7. At medium confidence levels (3σ), the four kinds of image-based analyses (group 2 to group 5) do not offer significant advantage in detection rates over a scalar signal of flow counting. However, the image-based signals offer significant advantage in false alarm rates. When higher confidence levels (3.5σ ~ 4.0σ) are considered (for decreasing the false alarm rates further), the image-based signals provide significantly better detection results than the flow counting approach. This clearly shows that the vector signal offers more significant improvement in the detection of anomalies than scalar signal.

7.2 General discussion of Results

The statistical analysis and NP-test show that the performance of the destination address domain slightly

degrades due to the bimodality of distribution over the destination address space. These analyses show that the source address based images/signals generally exhibit higher confidence than the destination address based images/signals for detecting traffic anomalies especially in false alarm rates and likelihood ratios.

In most experiments, the results of statistical 3σ bounds match those of the NP detector. And these two analyses illustrate consistent evaluation on the power of various signals. Based on these two compatible analyses, we can judge which signals are more effective in detecting traffic anomalies. Our evaluations indicate that the vector signals, which track variations of distributions of the underlying traffic header domains, provide more reliable signals than the scalar signals of traffic volume. Our evaluations also indicate that flow counting is superior to volume counting in scalar signals. Our evaluations show that flow counting can approach the detection rates of vector signals but suffers from higher false alarm rates.

Between the two approaches employed, 3σ approach does not require the analysis of the distribution of H_1 and can be more easily implemented. On the other hand, the NP-test requires PDFs of the H_0 and H_1 to be known and be parametrizable. If the NP-test uses insufficient observation samples S_1 for analyzing anomalous traffic H_1, it could result in a biased modeling, for example inaccurate histogram of the multimodality. While the NP-test is optimal and useful for understanding the inherent strengths of various signals as used here, statistical approaches may be more effective for anomaly detection in practice (due to unknown attacks).

7.3 Consideration using other traces

Results from *ISP* and *Campus* traces are shown in the Table 8 through Table 11 in Appendix A-2 and A-3. Overall, the performances of traffic volume-based signals and image-based signals show consistency regardless of traffic types exploited.

8. CONCLUSION

In this paper, we have evaluated a number of signals proposed for detecting traffic anomalies. We evaluated the signals on three different traces using two different methods. We proposed classical detection theory based on NP-test for detecting anomalies in traffic. We also employed statistical techniques for unknown attack detection. Both the statistical techniques and the NP-test based evaluations provide similar conclusions. Our evaluations indicate that the vector signals, which track variations of distributions of the underlying traffic header domains, provide more reliable signals than the scalar signals of traffic volume. Our evaluations also demonstrated that statistical techniques can be simpler and as effective as the NP-test based analysis.

9. REFERENCES

[1] C. Estan, S. Savage and G. Varghese, "Automatically Inferring Patterns of Resource Consumption in Network Traffic", in Proc. of ACM SIGCOMM 2003, Karlsruhe, Germany, August 2003.

[2] A. Kuzmanovic and E. Knightly, "Low-Rate TCP-Targeted Denial of Service Attacks", in Proc. of ACM SIGCOMM 2003, Karlsruhe, Germany, August 2003.

[3] P. Barford, J. Kline, D. Plonka and A. Ron, "A Signal Analysis of Network Traffic Anomalies", in Proc. of ACM SIGCOMM Internet Measurement Workshop (IMW), Marseille, France, November 2002.

[4] S. Kim, A. L. N. Reddy and M. Vannucci, "Detecting traffic anomalies through aggregate analysis of packet header data", in Proc. of Networking 2004, LNCS vol. 3042, pp.1047-1059, Athens, Greece, May 2004.

[5] Dave Plonka, "FlowScan: A Network Traffic Flow Reporting and Visualization Tool", in Proc. of the USENIX 14th System Administration Conference, New Orleans, LA, December 2000.

[6] P. Barford and D. Plonka, "Characteristics of Network Traffic Flow Anomalies", in Proc. of ACM SIGCOMM Internet Measurement Workshop (IMW), October, 2001.

[7] Jorma Kilpi and Ilkka Norros, "Testing the Gaussian approximation of aggregate traffic", in Proceedings of ACM SIGCOMM Internet Measurement Workshop (IMW), Marseille, France, November 2002.

[8] B. Krishnamurthy, S. Sen, Y. Zhang and Y. Chen, "Sketch-based Change Detection: methods, Evaluation, and Applications", in Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC) 2003, Miami, USA, October 2003.

[9] Dan Lelescu and Dan Schonfeld, "Statistical Sequential Analysis for Real-time Video Scene Change Detection on Compressed Multimedia Bitstream", IEEE Transactions on Multimedia, vol. 5, issue 1, pp 106-117, 2003.

[10] H. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic partitioning of Full-motion Video", Multimedia Systems, vol. 1, no. 1, pp 10-28, 1993.

[11] R. Lienhart, C. Kuhmunch, and W. Effelsberg, "On the Detection and Recognition of Television Commercials", in Proc. Of the International Confernce on Multimedia Computing and Systems, pp 509-516, Ottawa, Canada, 1997.

[12] K. Shen and E. J. Delp, "A fast Algorithm for Video Parsing Using MPEG Compressed Sequences", in IEEE Conference on Image Processing, pp 252-25, 1995.

[13] Gyaourova, A., C. Kamath, and S.-C. Cheung, "Block matching for object tracking", LLNL Technical report, October 2003. UCRL-TR-200271.

[14] A. Hussein, J. Heidemann, and C. Papadopoulus, "A framework for classifying denial of service attacks", ACM SIGCOMM, Aug. 2003.

[15] H. Vincent Poor, An Introduction to Signal Detection and Estimation, Springer Press, 2nd Edition, pp. 11, 1994

[16] NIST/SEMATECH e-Handbook of Statistical Methods. Available:http://www.itl.nist.gov/div898/handbook/eda/section3/histogr5.htm

[17] Emanuel Parzen, "On Estimation of a Probability Density Function and Mode", The Annals Mathematical Statistics, Vol. 33, No. 3, pp 1065-1076, September 1962

[18] Robert V. Hogg and Allen T. Craig, Introduction to mathematical statistics, Macmillan Company, 2nd Edition, pp. 285, 1965.

[19] Martin Roesch, "Snort-Lightweight Intrusion Detection for Networks", in Proc. of LISA '99, 13th Systems Administration Conference, Seattle, Washington, USA, November 1999.

[20] David Moore, Colleen Shannon, Geoffrey M. Voelker, and Stefan Savage, "Internet Quarantine: Requirements for Containing Self-Propagating Code", in INFOCOM 2003 conference, April, 2003

[21] Stuart Staniford, Vern Paxson, and Nicholas Weaver, "How to 0wn the Internet in Your Spare Time", in the Proceedings of the 11th USENIX Security Symposium (Security '02), San Francisco, CA, USA, August 5-9, 2002.

[22] Christian Estan and George Varghese, "New Directions in Traffic Measurement and Accounting", in ACM SIGCOMM 2002, Pittsburgh, PA, USA, August 2002.

[23] Marianne Durand and Philippe Flajolet, "Loglog Counting of Large Cardinalities", in the "Engineering and Applications Track" of the *11th Annual European Symposium on Algorithms* (ESA03)., LNCS vol. 2832, pp.605-617, September 2003.

[24] Reference is omitted for preserving author's anonymity.

[25] A. Lakhina, M. Crovella and C. Diot "Diagnosing network-wide traffic anomalies", in ACM SIGCOMM, Sept. 2004.

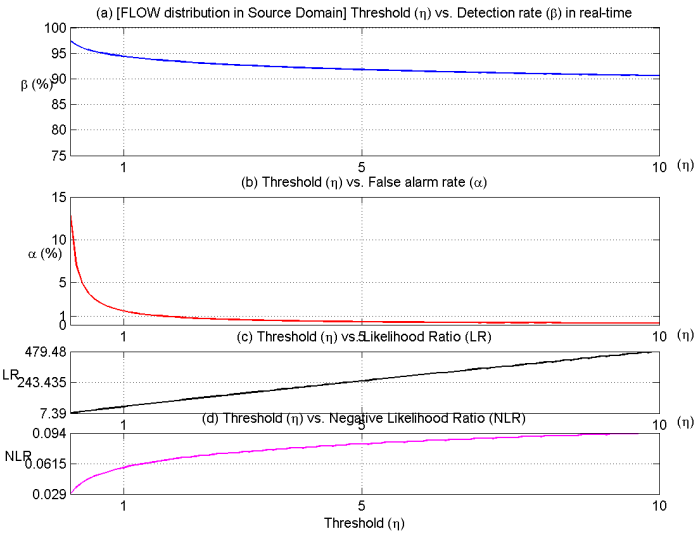
[26] MathWorld The web's most extensive mathematics resource, Available:http://mathworld.wolfram.com/LikelihoodRatio.html.

[APPENDIX]

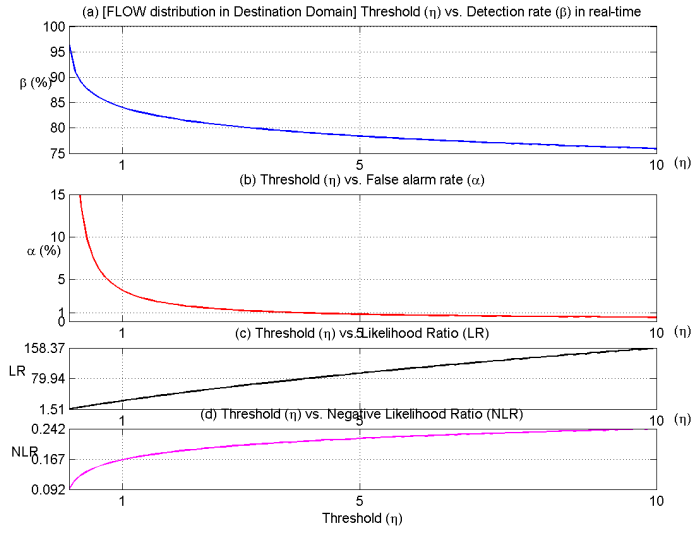
A-1 Effects of Thresholds (η) in NP-test

Bayesian criterion requires knowing a prior distribution π_0, π_1, where a threshold is determined as the ratio of π_0/π_1 for choosing decision region so as to minimize total probability of error. The optimal test is the threshold test of the likelihood ratio. Due to inexact knowledge of priors, however, we adopt an alternative fidelity criterion, NP criterion, in which threshold is not in terms of priors. Given a constrained significance level α, we can derive the threshold η of the NP-test for maximizing the detection rate β. For example, when false alarm rate is constrained to 1.0%, the threshold is derived as about 2.0 and corresponding detection rate reaches 94.0% in source domain as shown in the Fig. 8. We can carry an acceptable trade-off between detection rates and false alarm rates by varying the threshold.

Fig 8 shows the relationship between various performance criteria and thresholds in flow-based image signal in real-time. As shown in Fig. 8, as the threshold increases, the detection rate degrades and the false alarm rate improves. Similarly, the increased threshold make likelihood ratio better and make negative likelihood ratio



[Source Domain]



[Destination Domain]

Figure 8. The Relationship between thresholds and measurements criteria.

worse. Generally the measurement criteria of source domain are superior to those of destination as shown in LR and NLR. The β and α of destination domain change more drastically than those of source domain due to bimodality.

A-2 Campus Traces

Table 8. Port-based Measurement Results of *Campus* traces

Time	D.	TP β	FP α	*NP β*	*NP α*	LR	NLR
Real-time	SP	93.5% 58 / 62	0.15% 2 / 1313	*98.4%*	*0.28%*	614.1/ *352.9*	0.06/ *0.02*
	DP	91.9% 57 / 62	0.08% 1 / 1313	*95.7%*	*0.18%*	1207.1/ *542.0*	0.08/ *0.04*
	(SP, DP)	95.2% 59 / 62	0.15% 2 / 1313	–	–	634.7	0.05
Post mortem	SP	93.5% 58 / 62	0.08% 1 / 1313	*97.4%*	*0.08%*	1228.3/ 1188.9	0.06/ *0.03*
	DP	91.9% 57 / 62	0.08% 1 / 1313	*94.0%*	*0.12%*	1207.1/ *803.4*	0.08/ *0.06*
	(SP, DP)	93.5% 58 / 62	0.08% 1 / 1313	–	–	1228.3	0.06

Table 9. Results of Multidimensional signals of *Campus* traces

Signals	T.P. β	F.P. α	LR	NLR
Real-time (S,D)	96.8% 60 / 62	0.15% 2 / 1313	635.3	0.03
Postmortem (S,D)	98.4% 61 / 62	0.08% 1 / 1313	1291.8	0.02

A-3 ISP Traces

Table 10. Address-based Measurement Results of *ISP* traces

Time	D.	TP β	FP α	*NP β*	*NP α*	LR	NLR
Real-time	SA	84.6% 11 / 13	3.85% 1 / 26	*65.6%*	*0.14%*	22.0/ *468.6*	0.16/ *0.34*
	DA	92.3% 12 / 13	0.00% 0 / 26	*78.5%*	*0.14%*	∞ / *560.7*	0.08/ *0.22*
	(SA, DA)	100.0% 13 / 13	3.85% 1 / 26	–	–	26.0	0.00
Post mortem	SA	84.6% 11 / 13	3.85% 1 / 26	*64.4%*	*0.15%*	22.0/ 429.3	0.16/ *0.36*
	DA	84.6% 11 / 13	0.00% 0 / 26	*75.8%*	*0.14%*	∞ / *541.4*	0.01/ *0.24*
	(SA, DA)	84.6% 11 / 13	3.85% 1 / 26	–	–	22.0	0.16

Table 11. Results of Multidimensional signals of *ISP* traces

Signals	T.P. β	F.P. α	LR	NLR
Real-time (S,D)	100.0% 13 / 13	3.85% 1 / 26	26.0	0.00
Postmortem (S,D)	92.3% 12 / 13	3.85% 1 / 26	24.0	0.08