Detection of Low-Rate Attacks in Computer Networks

Gautam Thatte[1], Urbashi Mitra[1] and John Heidemann[2]
University of Southern California
[1]Ming Hsieh Department of Electrical Engineering, 3740 McClintock Ave, Los Angeles, CA 90089
[2]Information Sciences Institute, 4676 Admiralty Way, Marina Del Rey, CA 90292
Email: {thatte,ubli}@usc.edu, johnh@isi.edu

***Abstract*—This paper develops two parametric methods to detect low-rate denial-of-service attacks and other similar near-periodic traffic, without the need for flow separation. The first method, the periodic attack detector, is based on a previous approach that exploits the near-periodic nature of attack traffic in aggregate traffic by modeling the peak frequency in the traffic spectrum. The new method adopts simple statistical models for attack and background traffic in the time-domain. Both approaches use sequential probability ratio tests (SPRTs), allowing control over false alarm rate while examining the trade-off between detection time and attack strength. We evaluate these methods with real and synthetic traces, observing that the new Poisson-based scheme uniformly detects attacks more rapidly, often in less than 200ms, and with lower complexity than the periodic attack detector. Current entropy-based detection methods provide an equivalent time to detection but require flow-separation since they utilize source/destination IP addresses. We evaluate sensitivity to attack strength (compared to the rate of background traffic) with synthetic traces, finding that the new approach can detect attacks that represent only 10% of the total traffic bitrate in fractions of a second.**

I. Introduction

Early detection and mitigation of denial of service (DoS) attacks, link congestion and other traffic anomalies are vital to efficient network management. While a DoS attack is obvious at the victim, who is overwhelmed with traffic, it is much more difficult to detect attacks near their source, where network operators can take remedial actions. A common concealment method is to transmit attacks from multiple hosts ("zombie" machines) with each zombie sending at a low rate; thus detection of low-rate traffic at or near zombie sources is an important problem. In this paper we introduce new parametric detection methods that allow one to tune detection sensitivity to latency, making detection of low-rate zombies near the source feasible.

Our approaches are parametric—our detection method models key, gross features of the traffic to enable informed decisions after changes in traffic. We propose the parametric Modeled Attack Detector (MAD) based on time-series data and modify the method of He et al [7] to a sequential version, the Periodic Attack Detector (PAD), using spectrally-based techniques [4], [7], [8].

MAD models traffic without attacks as Poisson traffic, and it assumes that a constant rate attack results in a fixed increase in the number of packet arrivals per unit time. We *do not believe* that that model represents *general* Internet traffic accurately. However, we show that, in spite of this known mismatch, it is sufficient to effectively capture changes in the level of traffic which are associated with a DoS attack. Our goal is to see whether these simple and only approximate statistical models can yield detection methods of high performance by capturing sufficient, salient features of the traffic. Although we cannot quantify the error due to this model inaccuracy, our proposed models are able to detect attacks in real Internet traces as shown in Section V. Since our goal here is to develop new detection methods, not develop general traffic models, we believe this preliminary evidence suggests that a perfect model fit is not required to detect attacks in our case.

PAD is based on spectral methods which are motivated by the fact that constant rate attacks result in a prominent spectral peak that is computed as

$$\text{Dominant frequency} = \frac{\text{Link Bandwidth}}{\text{Packet Size}}.$$

We refer readers to the work by He et al [7] for an exposition on spectral detection methods.

The contributions of this paper are therefore to develop the new MAD detection scheme and sequentialize the scheme in [7]; these methods operate on aggregate traffic streams and we quantify their effectivness on both controlled synthetic traffic and real traces captured in the wild. We show that MAD can detect attacks as small as 10% of total traffic in less than a second. This work suggests it is feasible to detect and prune attacks at

network edges based on aggregate traffic, and not just near attack victims.

II. Related Work

Other authors have considered change-point algorithms for a variety of Internet detection problems. Non-parametric methods in are able to detect a TCP SYN attack within a few seconds, as [15] and [18] show. The challenges with the methods of [15] and [18] are the need for an empirically designed threshold, and knowledge of the SYN flag within the TCP header via flow separation, respectively. Exponential signal models were the basis for worm detection as developed in [2].

The techniques in [6], [12] and [17] use the entropy content of certain flow-separated traffic parameters, *e.g.* source and destination IP addresses, to detect an attack. Other methods, such as in [9] and [13], use parameters in the TCP field to detect an attack.

The above methods directly employed packet arrival time-series data; a contrasting approach is to compute the power spectral density of the time-series. It has been argued that spectrum-based approaches are adept at detecting features with near-periodic signatures, such as bottlenecks in the link layer, the TCP windowing mechanism and DoS attacks [7], traffic anomalies [1], and even for attack fingerprinting [8]. The sequential probability ratio test (SPRT), a time-adaptive detection technique, has been used to distinguish between reduction-of-quality (RoQ) flows and legitimate TCP flows in a distributed setting [4] and fast portscan detection [10].

The important features of our parametric detection schemes include:

- The fact that the analysis is based on packet interarrival times, not packet contents. Unlike the algorithms in [9] and [13], our methods are robust to changes in the transport-layer headers, *e.g.* time-to-live (TTL) in the IP header.
- In contrast to [4], [6], [12], [17] and [18], we operate on aggregate traffic *without* flow separation, enabling analysis of encrypted traffic in a passive monitoring framework.
- We estimate attack parameters in real-time and do not assume them *a priori*; our method requires less tuning and empirical parameter design than the work in [15].
- Our approach can accurately detect *low-rate* attacks, *i.e.* cases where attack traffic is 10% or less of the total bitrate.

Both the PAD and the MAD can detect attacks on the order of milliseconds when the attack bitrate is around 10% of the total traffic bitrate. We empirically find that the MAD is at least twice as fast as the PAD, and also test our detection methods on real Internet traffic.

III. Overview of Sequential Detection Methods

Hypothesis testing exploits prior knowledge of statistical descriptions of data in order to choose amongst a candidate set of populations. In our problem setup, we have two hypotheses:

$$\begin{aligned} & H_1: && \text{Presence of an attack in traffic,} \\ \text{and} \quad & H_0: && \text{No attack.} \end{aligned}$$

Assume that we are given independent and identically distributed (i.i.d.) observations $\{x_k, k = 1, 2, \ldots\}$ which are drawn from one of the two probability distributions; the conditional probability density when H_i is true is denoted $p(x|H_i)$ for $i = 0, 1$. Focusing on methods that detect quickly, we design SPRTs which take observations one at a time until a confirmed decision can be made. We refer readers to the seminal work by Wald [16] for details and derivations.

Given the two hypotheses, there are four possible outcomes for the detector; we focus on a particular pair of outcomes. A *false alarm* (FA) or *false positive* is declared when the algorithm selects H_1 when H_0 is in fact true. Similarly, choosing H_0 even though H_1 is true is termed a *miss* (M) or *false negative*. We use the probabilities of these two outcomes,

$$\alpha = P_{\text{FA}} = P[H_1|H_0] \ , \ \gamma = P_{\text{M}} = P[H_0|H_1], \quad (1)$$

to specify the performance criterion of the sequential detection test.

Having defined the detection performance criterion, we consider the statistical measure used to implement the SPRT, which is termed the *likelihood ratio*. Since the likelihood ratio is employed to decide between the two hypotheses, it is defined in terms of the associated conditonal densities. Given i.i.d. observations $\mathbf{x}_N = \{x_1, \ldots, x_N\}$, the likelihood ratio is defined as

$$L_N(\mathbf{x}) = \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)} = \frac{p(x_N|H_1)}{p(x_N|H_0)} L_{N-1}(\mathbf{x}), \quad (2)$$

where the second equality illustrates that the likelihood ratio can be updated as each new observation is made available.

Given a new observation, the likelihood ratio is compared to two thresholds A and B, which correspond to

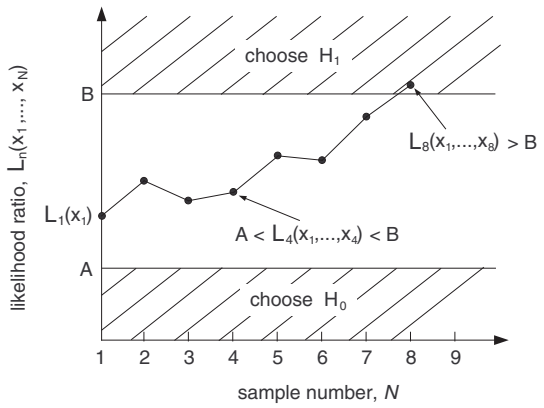

Fig. 1. Depiction of the sequential probability ratio test (SPRT).

choosing H_0 or H_1, respectively. Figure 1 depicts a realization of the SPRT wherein if $A < L_N(x_1, \ldots, x_N) < B$, the experiment continues, and an additional observation x_{N+1} is taken. But if $L_N(x_1, \ldots, x_N) \geq B$ or $L_N(x_1, \ldots, x_N) \leq A$, the test terminates and we choose hypothesis H_1 if the former, or hypothesis H_0 if the latter, is true. The boundaries, A and B, of the SPRT would ideally be selected to minimize the probability of error for all possible values of N. Determining the exact value is generally intractable, thus we use Wald's approximations [16] to approximate

$$B \cong (1-\gamma)/\alpha \quad \text{and} \quad A \cong \gamma/(1-\alpha), \tag{3}$$

which are a function of the required detection performance parameters from Equation (1). We observe that the approximate values of A and B are independent of $p(x|H_i)$.

Since the number of samples required for the sequential hypothesis test is a random variable, the efficacy of the test can be evaluated by computing the average number of samples required to make a decision by a particular test. This average sample number (ASN) function is simulated experimentally, and is used as the performance metric to compare the two methods developed. In the following section, we describe the parametric detection methods and derive the associated SPRTs.

IV. Developing the Detection Algorithms

The MAD operates on the sampled time-series of network traffic, whereas the PAD uses traffic spectra obtained from processing non-overlapping segments of the sampled time-series. The purpose of segmentation in the latter method is to ensure that traffic spectra are computed and compared for constant time intervals as described by He et al [7]. Both methods assume that the parameters related to the background traffic, *i.e.* the null hypothesis H_0, are initially known. The background parameters are updated in real-time whenever the H_0 hypothesis is chosen, and the attack parameters are continually updated with each available observation. In this section, we outline the development of the SPRT for these two methods, particularized to the specific densities used in the PAD and MAD, based on the generic framework from the previous section.

A. The Modeled Attack Detector (MAD)

The MAD models the attack stream as a constant rate source with *deterministic, unknown* rate $r_A > 0$. The no-attack case is modeled as Poisson arrivals, and the attack case as a constant rate attack in Poisson background traffic, which corresponds to the shifted Poisson density[1]. The relevant densities are given by

$$p(x|H_0) = \frac{\lambda_B^x e^{-\lambda_B}}{x!} \quad \text{and} \quad p(x|H_1) = \frac{\lambda_B^{x-r_A} e^{-\lambda_B}}{(x-r_A)!} \tag{4}$$

for an attack with rate r_A and background Poisson traffic with rate λ_B, where x represents the number of packet arrivals in a unit time.

The likelihood ratio assumes that parameters of the underlying probability density are known. Since the attack rate is not known, we employ the generalized likelihood ratio test (GLRT) [3]. The GLRT uses each observation for two tasks: to compute the maximum likelihood estimate (MLE) of the rate parameter $\widehat{r}_A$ and to decide between H_1 and H_0. The MLE of the attack rate, from N observations, is

$$\underbrace{\widehat{r}_A}_{\substack{\text{estimate of}\\ \text{attack rate}}} = \underbrace{-\lambda_B}_{\substack{\text{known background}\\ \text{traffic rate}}} \underbrace{+\frac{1}{N}\left[\sum_{i=1}^{N} x_i\right]}_{\substack{\text{estimated rate}\\ \text{of total traffic}}}, \tag{5}$$

wherein we assume the background rate λ_B is known or can be estimated previously. We can further derive the corresponding SPRT as

$$\prod_{i=1}^{N} \frac{\lambda_B^{\lambda_B - \overline{x}_i} x_i!}{\Gamma(x_i + \lambda_B - \overline{x}_i + 1)} \;\geq\; B \tag{6}$$

for the test to determine the presence of an attack, where $\overline{x}_N$ is the mean of observations $\{x_k, k = 1, \ldots, N\}$.

[1]The shifted Poisson density with parameters (λ, c) is similar to the Poisson density, but requires that at least c arrivals occurs in any unit of time, and thus the number of arrivals when there is an attack is always strictly greater than zero.

As the mean and background rate are generally real numbers, we round the quantity $\lambda_B - \overline{x}_i$ to the nearest integer in practice. This is the generic SPRT equation $L_N(x_1 \ldots, x_N) > B$, particularized to the Poisson and shifted Poisson densities for the MAD, and incorporating the MLE of the attack rate $\widehat{r}_A$.

B. The Periodic Attack Detector (PAD)

This frequency-domain method uses the traffic spectrum to decide between the two hypotheses. The PAD is the sequential version of the scheme developed by He et al [7] with the attack parameters estimated in real-time, rather than known *a priori*. A detailed methodology and processing details for obtaining a traffic spectrum from a trace of packet arrivals is provided in [7].

As mentioned in Section I above, spectral methods for the detection of periodic attacks are motivated by the presence of a dominant frequency. This dominant frequency, or base frequency, is the statistic in the spectral-domain that we use to detect between H_1, the presence of an attack, and H_0, no attack (the case of just background traffic). Even if packets are of different sizes, a few common packet sizes (in particular those related to a DoS attack) tend to dominate the traffic [5], [14]. We model the log of the maximum amplitude in the traffic spectrum for both hypotheses with a Gaussian distribution [7]. This specifies the detection problem as distinguishing between two Gaussian distributions with parameters (μ_1, σ_1^2) and (μ_0, σ_0^2). Given the Gaussian distribution $\mathcal{N}(y; \mu, \sigma^2)$, y represents the log of the maximum amplitude of the traffic spectrum. The parameters for the null hypothesis H_0 are pre-determined using training data, and perioddically updated. On the other hand, similar to the MAD, the parameters for hypothesis H_1 are estimated using each observation at the same time as PAD is deciding between the two hypotheses.

The SPRT to determine the presence of an attack, given N observations, is:

$$a_2 \sum_{k=1}^{N} y_k^2 + 2a_1 \sum_{k=1}^{N} y_k + N a_0 \geq \tau \qquad (7)$$

where

$$\begin{aligned} a_2 &= \widehat{\sigma}_1^2 - \sigma_0^2 \ , \quad a_1 = \widehat{\mu}_1 \sigma_0^2 - \mu_0 \widehat{\sigma}_1^2 \ , \\ a_0 &= \mu_0^2 \widehat{\sigma}_1^2 - \widehat{\mu}_1^2 \sigma_0^2 \ , \\ \text{and} \quad \tau &= 2\sigma_0^2 \widehat{\sigma}_1^2 \left[\ln B - N \ln \frac{\sigma_0}{\widehat{\sigma}_1} \right], \end{aligned}$$

and y_k is the log of the maximum amplitude of traffic spectrum for k-th segment. The MLEs of the mean and variance parameters for hypothesis H_1 are computed as

$$\widehat{\mu}_1 = \frac{1}{N} \sum_{k=1}^{N} y_k \quad \text{and} \quad \widehat{\sigma}_1^2 = \frac{1}{N} \sum_{k=1}^{N} (y_k - \widehat{\mu}_1)^2 \ . \quad (8)$$

The SPRTs derived for the MAD and PAD are used for both synthetic traces and real traffic in the following section. The simulations yield the experimental ASN that is used to compare the two detection schemes, which we discuss in the next section. As an aside, we outline an argument for the lower complexity of the MAD in the following subsection.

C. Computational Complexity

We shall see that MAD detects more quickly than PAD; herein we show that it is also more computationally efficient. The cost of PAD is a function of sample rate and the FFT window size,

$$\text{PAD} \sim O\left(\text{ASN}_{PAD} + \text{ASN}_{PAD} \cdot N_{FFT/p} \log N_{FFT/p}\right),$$

while the cost of MAD is proportional only to the sample rate,

$$\text{MAD} \sim O\left(\text{ASN}_{MAD}\right).$$

Thus, MAD is algorithmically more efficient than PAD ($O(n)$ compared to $O(n \log n)$) in the limit.

Furthermore, the sample rate of MAD is approximately a factor of 100 smaller than PAD. Normalizing to the cost of each sample, MAD costs 10^3 samples/second, while PAD costs 2×10^5 samples/second plus an FFT for each of the 1000 segements/second, with each FFT costing about as much as 1600 samples.

V. Numerical Results

We use both synthetic traces and real traffic to validate and characterize the detection methods developed in this paper. The synthetic traces allow us to methodically examine the sensitivity of our algorithm: the trade-off between time-to-detection and the rate of the attack. We first note that the sampling rate for MAD is 10^3 Hz, *i.e.* each observation x_i represents the number of packet arrivals in the interval $\left[\frac{i}{10^3}, \frac{i+1}{10^3}\right)$ seconds. On the other hand, PAD segments the trace into 1 ms segments, samples each segment at 2×10^5 Hz, and computes the traffic spectrum via the discrete Fourier transform. Each observation y_k represents the log of the maximum amplitude of the traffic spectrum of the k-th segment.

A. Synthetic Traces

Both the MAD and PAD have been characterized using synthetic traces[2] generated from real background

[2]The five-minute attack traces are available as PREDICT ID USC-LANDER UniformAttack-Traces-Generated20070821-20041202.

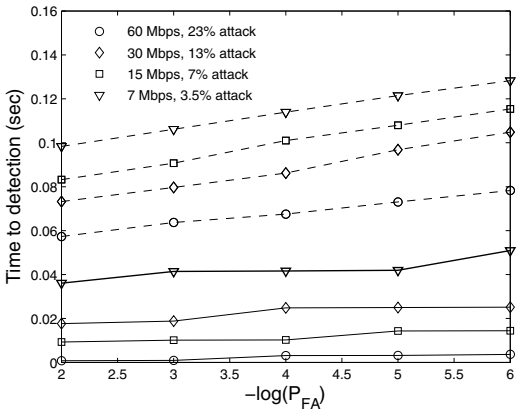

Fig. 2. Comparing time to detection for the MAD (solid lines) and PAD (dashed lines) as a function of false positive rate.

traffic and artificial attacks with rates ranging from 60 Mbps to 7 Mbps using the stream merger application [11]. The attack rate is expressed as a percentage of the total traffic, *i.e.*

$$\% \text{ attack } = \frac{\text{attack rate}}{\text{attack rate} + \text{background rate}},$$

e.g. a 60 Mbps attack in 196 Mbps background traffic is denoted as a 23% attack.

To test the capabilities of the detection algorithms, we fix the probability of miss to be 10^{-3} and vary the probability of false alarm from 10^{-2} to 10^{-6}. The experimental ASN curves in Figure 2 were generated by simulating the SPRTs of the two methods, Equations (6) and (7), and averaging the results over 25 runs.

The experimental ASN for both the MAD (solid lines in Figure 2) and PAD (dashed lines in Figure 2) decrease *sublinearly* as the attack rate increases; specifically, they require 50% and 80% as many samples when the attack rate doubles. We also see that the average number of required samples increases as the probability of false alarm decreases, which is equivalent to the performance constraints becoming more demanding. Both methods require an average of 40% more samples as the probability of false alarm changes from 10^{-2} to 10^{-6}.

A direct comparison between the detection methods shows that the MAD uniformly detects attacks more quickly than the PAD. Furthermore, we find that the PAD requires relatively more time as the attack rate increases, *e.g.* the PAD takes three times as long to detect a 7 Mbps attack, but eight times as long to detect a 30 Mbps attack in 196 Mbps background traffic. The PAD relies on the near-periodic nature of the trace to achieve rapid and efficient detection, which is not always a realistic assumption. We are currently investigating whether the collisions resulting from the merger of attack and background streams as attack rate increases introduces jitter in the packet arrivals, potentially affecting the periodicity of the resulting trace.

The synthetic trace simulations provide us with guidelines on the effectiveness and capability of our detection methods. They suggest that low-rate attacks, even as low as 3.5%, can be detected on the sub-second timescale with no *a priori* knowledge of the attack parameters. We do, however, need to know the background traffic rate and we are currently examining methods to estimate this quantity as well. Furthermore, we can use the experimental ASNs as rough estimates when detecting real attacks. The detectability of low-rates attacks and the average time to detection obtained from the synthetic trace simulations will necessarily be optimistic due to the controlled setup, and this must be noted when real attacks are detected.

B. Real Attacks

MAD uniformly detects attacks more quickly than the PAD; we use the former method to detect the presence of an attack in real Internet traces[3]. The second subplots in Figures 3 and 4 plot the evolution of the SPRT until the likelihood ratio crosses the threshold B, at which point the attack is detected. The technique proposed by Feinstein et al [6] was also simulated (in the third subplot) by computing the entropy of the destination IP address using a window of 5000 packets. A comparison of the two methods yields equivalent detection times, but the method in [6] requires flow-separated traffic since it uses the IP address as part of its statistic.

The time to detection of the real attacks using MAD confirms that the average time to detection of the synthetic attacks is indeed very optimistic. For similar 25% attacks, the time required to detect a real attack is approximately 10 times greater than the detection delay associated with the synthetic trace.

VI. Discussion and Future Work

We have developed the MAD and PAD, which can detect low-rate attacks on sub-second timescales. The MAD operates on the sampled time-series and can detect attacks more quickly than the PAD, which relies on the near-periodic nature of the traffic. The lack of strong periodicities adversely affects the performance of the

[3]The traces used to validate the performance of the MAD are available as PREDICT ID USC-LANDER DoS_traces-20020629 and PREDICT ID USC-LANDER TBD

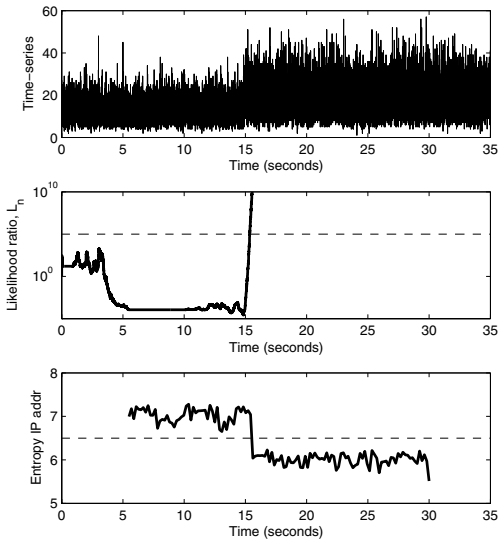

Fig. 3. Detection of a 24% reflector attack (time-series in subplot 1) by both the MAD (subplot 2) in 150 milliseconds, and the method in [6] (subplot 3) in 400 milliseconds.

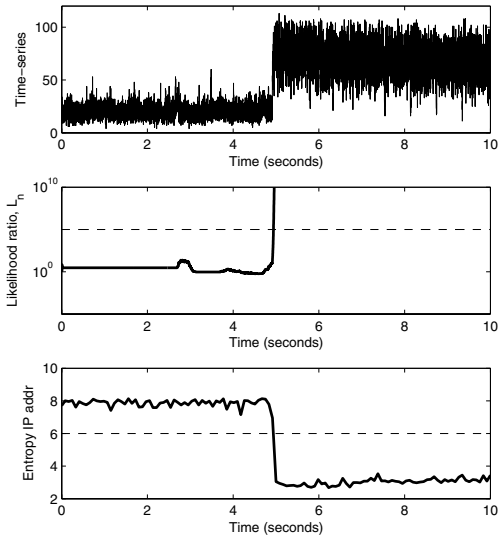

Fig. 4. Detection of a 46% IP-proto 255 attack (time-series in subplot 1) by both the MAD (subplot 2) in 27 milliseconds, and the method in [6] (subplot 3) in 55 milliseconds.

latter method. The MAD is also lower complexity than PAD, both theoretically and practically.

When tested on real Internet traffice, the MAD detected a 24% attack in under 0.2 seconds. We have further characterized the capabilities of the two methods using synthetic traces. We find that the results of testing on real Internet traffic are within the experimental limits of those predicted by the experimental ASNs from synthetic traces. Our current work focuses on incorporating more features of the traffic spectrum into the PAD to lessen its dependence on the trace periodicities, designing other parametric variations of the MAD to operate at different timescales, and further validating the model using real traffic.

ACKNOWLEDGMENTS

Research has been funded by DHS NBCHC040137 and NSF CNS-0626696. Traces used in this work were provided by the USC/LANDER project.

REFERENCES

[1] P. Barford, J. Kline, D. Plonka and A. Ron, "A Signal Analysis of Network Traffic Anomalies," in *SIGCOMM Internet Measurement Workshop*, France, Nov 2002.

[2] A.A. Cardenas, J.S. Baras and V. Ramezani, "Distributed Change Detection for Worms, DDoS and other Network Attacks," in *ACC '04*, Boston, MA, 2004.

[3] G. Casella and R. Berger, *Statistical Inference*, Duxbury, 2002.

[4] Y. Chen and K. Hwang, "Spectral Analysis of TCP Flows for Defense against Reduction-of-Quality Attacks," in *ICC 2007*, Glasgow, Scotland, June 2007.

[5] K. Claffy, G. Miller and K. Thompson, "The nature of the beast: Recent traffic measurements from an internet backbone," in *INET '98*, Geneva, CH, July 1998.

[6] L. Feinstein, D. Schnackenberg, R. Balupari and D. Kindred, "Statistical Approaches to DDoS Attack Detection and Response," Proc. DARPA Information Survivability Conf. and Exposition, vol. 1, IEEE CS Press, pp. 303 - 314, 2003.

[7] X. He, C. Papadopoulos, J. Heidemann, U. Mitra, U. Riaz and A. Hussain, "Spectral Analysis of Bottleneck Traffic," Technical Report USC-CSD-TR-05-854, May 2005.

[8] A. Hussain, J. Heidemann and C. Papadopoulos, "Identification of Repeated Denial of Service Attacks," in *INFOCOM 2006*, Barcelona, Spain, April 2006.

[9] C. Jin, H. Wang and K. Shin, "Hop-Count Filtering: An Effective Defense Against Spoofed DoS Traffic," in *ACM CCS*, Oct 2003.

[10] J. Jung, V. Paxson, A.W. Berger and H. Balakrishnan, "Fast Portscan Detection Using Sequential Hypothesis Testing," *IEEE Symposium on Security and Privacy*, Oakland, CA, May 2004.

[11] P. Kamath, K.-C. Lan, J. Heidemann, J. Bannister and J. Touch, "Generation of High Bandwidth Network Traffic Traces," in *MASCOTS*, Fort Worth, TX, 2002.

[12] A. Lakhina, M. Crovella and C. Diot, "Mining anomalies using traffic feature distributions," ACM SIGCOMM Computer Communication Review, 35(4), pp. 217 - 228, 2005.

[13] J.C.C. Rodriguez, A.P. Briones and J.A. Nolazco, "Dynamic DDoS Mitigation based on TTL field using fuzzy logic," in *CONIELECOMP '07*, Mexico, Feb 2007.

[14] R. Sinha, C. Papadopoulos and J. Heidemann, "Fingerprinting internet paths using packet pair dispersion." Technical Report USC/CS-TR-2006-876, 2006.

[15] A. Tartakovsky, B. Rozovskii, R. Blazek, and H. Kim, "A Novel Approach to Detection of Intrusions in Computer Networks Via Adaptive Sequential and Batch-Sequential Change-Point Detection Methods," *IEEE Trans. Sig. Proc.*, 54(9), 2006.

[16] A. Wald, *Sequential Analysis*, New York: John Wiley, 1947.

[17] A. Wagner and B. Plattner, "Entropy Based Worm and Anomaly Detection in Fast IP Networks," 14th IEEE WETICE, 2005.

[18] H. Wang, D. Zhang, and K. G. Shin, "Detecting SYN Flooding Attacks," in Proceedings of IEEE INFOCOM 2002.