

Detecting Malware Outbreaks using a Statistical Model of Blackhole Traffic

Sohraab Soltani[†], Syed Ali Khayam[§] and Hayder Radha^{†*}

[†]Department of Computer Science & Engineering, [‡]Department of Electrical & Computer Engineering
Michigan State University
East Lansing, MI 48824, USA

[§]NUST Institute of Information Technology (NIIT), National University of Sciences & Technology (NUST), Rawalpindi, Pakistan
{soltanis, khayamsy, radha}@msu.edu

Abstract—Internet blackholes have emerged as very effective tools for monitoring changes in the Internet’s traffic behavior. Prior studies have shown that traffic observed at a blackhole contains valuable information about emerging malware. While blackhole traffic has been effectively used for attack forensics, a systematic method of leveraging this traffic for online Internet-scale anomaly detection is not available. In this paper, we propose a novel technique to detect malware outbreaks using deviations in a robust statistical model of a blackhole’s traffic. First, we introduce a novel and accurate *Piecewise Poisson process Model* (PPM) of traffic observed at an Internet Motion Sensor (IMS) blackhole which provides a statistical quantification of the intensity or *rate* of incoming traffic at a blackhole, which can in turn be used to detect malware outbreaks. After establishing the accuracy of the proposed PPM model, we develop a regression model that can characterize variations in the PPM’s traffic rates. Once an accurate model of traffic rates is in place, malware outbreaks can be detected using deviations from the model’s likely statistical patterns. After removing simple deterministic patterns, we observe that a blackhole’s traffic rate residuals have a skewed and heavy-tailed behavior. Consequently, we employ a *stable distribution* that models variations in traffic rate residuals with very high accuracy. Finally, we propose an online detection mechanism that utilizes deviations from the rate residual distribution of blackhole traffic data to detect malware outbreaks. Experimental results using the IMS data for approximately one year show that the proposed mechanism accurately detects malware outbreaks in a timely manner.

I. INTRODUCTION

A blackhole is a sensor that continuously monitors and logs the Internet’s traffic [1]– [3]. A blackhole sensor monitors a block of unused (or dark) IP address space. Since unused address blocks are not assigned to valid hosts, incoming traffic on these unused addresses contains abnormal (but not necessarily malicious) activities like misconfigurations, benign scanning or probing, and overtly malicious behaviors like backscatter from spoofed source addresses, scan traffic from malware probing etc. While blackhole traffic contains information about new malware attacks which can and should be used for Internet-scale anomaly detection, so far blackhole data has mostly been used for attack forensics [4]– [8]. Despite

the interesting insights provided in [1]– [3], blackholes are not being used to detect malware outbreaks because a robust and online detector that can correlate and leverage the received blackhole traffic for anomaly detection does not exist. This paper presents an unsupervised online detection technique that can be deployed on a blackhole to detect malware outbreaks. The proposed technique detects outbreaks by flagging deviations from a sound statistical model of blackhole traffic.

We analyze the traffic patterns observed at a blackhole with the objective of developing a suitable stochastic model of blackhole traffic. Subsequently, we aim at utilizing the blackhole traffic model to detect malware outbreaks. Benefits of an accurate stochastic model of blackhole traffic are manifold. First, the model provides important insights into the fundamental characteristics of the traffic observed at a blackhole. These insights can be used to evaluate and improve malware detection and defense techniques. Modeling insights can also be used to improve malware monitoring techniques, for instance by identifying effective blackhole locations. Second, an accurate model allows simulations of blackhole traffic (e.g., by simulating the behavior of a worm outbreak) without having an actual blackhole in place. Lastly, and perhaps most importantly, an accurate model characterizes the *typical* behavior of a blackhole’s traffic. Typical behavior is characterized by harmless traffic observed at a blackhole. While most of the blackhole traffic is a consequence of illegitimate/anomalous activities, not all of this traffic is necessarily malicious in nature. For instance, a significant amount of traffic observed at a blackhole is due to benign scanning, probing, misconfiguration, or some already patched prior malware [1]. An accurate model should be able to differentiate such typical activities from truly malicious traffic. Deviations from this typical behavior can also be used to detect notable malicious activity such as a worm outbreak.

As a first step toward modeling blackhole traffic, we show that the intensity or *rate* of incoming traffic at a blackhole can be characterized as a time-varying Poisson process with a corresponding time-dependent Poisson arrival-rate parameter. Thus, we propose a novel *Piecewise Poisson process Model* (PPM) that can accurately capture the behavior of blackhole

*This work was supported in part by NSF Award CNS-0721550, NSF Award CCF 0728996, NSF Award CCF-0515253, and NSF Award CNS-0430436.

traffic. After establishing the high accuracy of the PPM, we model the PPM's traffic rates using a statistical regression model. This regression model reveals important and interesting statistical properties of blackhole traffic. We note that after removing simple deterministic patterns, the remaining traffic rate residuals exhibit heavy-tailed and skewed behavior due to the presence of overtly malicious traffic. We show that the well-known *stable distribution* [9] can accurately model the traffic rate residuals.

The detector proposed in this paper employs the residual rate's stable distribution to find an accurate threshold for the detection of malicious patterns. The proposed detector can adapt the detection threshold in an online and unsupervised manner to provide accurate and real-time detection of malware outbreaks. We test the proposed detection technique using traffic observed at an IMS blackhole for a period of approximately one year. We show that the proposed technique detects most of the significant worm outbreaks [10] in a very timely fashion. Moreover, the proposed detector can accurately identify the attacked ports used in a malware outbreak.

This paper is organized as follows. Section II briefly describes the blackhole data collected by the IMS project. In Section III, we present the Piecewise Poisson process Model (PPM) and explain how the PPM can be used to simulate the blackhole traffic. Section IV describes the regression model for the traffic rates and models the traffic residuals using the stable distribution. The detection technique and evaluations are provided in Section V. We conclude the paper in Section VI.

II. INTERNET MOTION SENSOR DATA

In this work, we use data collected at a blackhole that is part of the Internet Motion Sensor (IMS) project [1]. The IMS consists of distributed blackhole sensors, each monitoring certain unused IP addresses. These sensors are located at major service providers, large enterprises, academic networks, and broadband providers. Specifically, there are 28 monitored IP blocks at 18 physical locations with subnet sizes varying from /25 to /8. The data analyzed in this paper are monitored activities recorded on /24 IP blocks on a daily basis from 21st of March 2004 until 28th of February 2005, a total recording duration of 345 days. IMS sensors monitor the activities on all ports (i.e., 0 to 65535). An IMS sensor records three types of traffic data:

- *IP-traffic*: this traffic type contains the number of unique/distinct sources (as characterized by source IP addresses) that scanned a particular port on a particular day.
- *MD-traffic*: on receiving a packet, the blackhole computes the Message Digest (MD) algorithm-5 signature of the payload and compares it with previously received payloads. A packet is logged if the MD5 signature is new.
- *PC-traffic*: the count of the total number of packets received on every port.

III. A PIECEWISE POISSON MODEL OF BLACKHOLE TRAFFIC

As explained above, the IMS sensors data used in this paper are based on aggregate traffic statistics, which were collected using a rather simple tallying (counting) of the parameter under consideration (e.g., IP-traffic count). Hence, one option for modeling this recording operation is by employing a counting random process. In this section, we introduce a novel counting process, namely the *Piecewise Poisson Process Model* (PPM), to accurately model a blackhole's traffic. PPM model is a collection of homogenous Poisson processes with varying rates (as explained below). Under the proposed PPM model, each recording period's traffic is modeled using a different Poisson process; thus for w samples of training data, $w \times 3$ different homogeneous Poisson processes with different rates are developed for IP-, MD- and PC-traffic data. In particular, at one hand, we are making the assumption that the traffic exhibits *independent* and *stationary increments* within each recording period in the sense that the arrival (counting) process within a given recording period can be divided into smaller periods of non-overlapping time segments that are independent and stationary (i.e., having the same Poisson rate). At the same time, we improve the accuracy of the model by capturing the changes in the process rate (based on the available data) from one recording period to another using a different Poisson process for each period. There are several analytical subtleties that one needs to handle carefully when developing the proposed PPM model; these will be discussed in the following sections.

A. Model Description

We describe the model development steps only for the IP-traffic counting process; the steps for MD- and PC-traffic modeling are identical. Henceforth, while we only discuss modeling of IP-traffic, it is implied that three PPM models are being generated, one for each traffic type. All three traffic models are then used collectively for malware detection in later sections.

Let $\{n = 1, 2, \dots, w\}$ represent discrete indices of w equal-sized, non-overlapping and ordered epochs defined on a continuous-time axis $t \geq 0$. The discrete indices n are referred hereafter as recording epochs or recording periods. Values of the continuous-time variable t on the boundaries of epoch n are denoted as t_{n-1} and t_n .

We define $N_n(t)$, for $t_{n-1} < t \leq t_n$, as the total number of incoming traffic events during epoch n . For each epoch n , we assume that $N_n(t)$ is a homogenous Poisson process on the epoch $(t_{n-1}, t_n]$ with a rate r_n , where r_n measures the intensity of traffic arrivals during the epoch $(t_{n-1}, t_n]$. Thus, for a given recording epoch n , the probability of recording k IP-traffic entries on a blackhole are given by the probability:

$$\Pr [N_n(t) = k] = \frac{(r_n t)^k}{k!} \exp(-r_n t), t_{n-1} < t \leq t_n.$$

Since we define a different Poisson process to model each epoch, we refer to this model as the Piecewise Poisson

process Model (PPM). Because any Poisson process is completely defined by its rate, the series of PPM rates, $\{r_n, n = 1, 2, \dots, w\}$, play a critical role in characterizing the proposed model. Moreover, abrupt changes in the rates can be used to detect notably malicious activity. Since the PPM rate sequence is pivotal in the present context, we now focus on accurately estimating the traffic arrival rates r_n in each epoch n .

B. Estimation of PPM Rates

The backhole data does not contain actual arrival times of each IP-traffic packet. Thus we do not have complete information to estimate the rates $\{r_n, n = 1, 2, \dots, w\}$ using simple empirical methods. In the absence of actual arrival time information, one simplistic and rather naïve approach for estimating the rate r_n is by directly using the recorded count in the dataset under consideration. For example, if the IP-traffic count recorded for a recording epoch n is some number X , then we can approximate the arrival rate during that epoch as $r_n = X$. However, this could generate highly abrupt transitions in the rate of an underlying time-continuous (Poisson) process.

To estimate the rates accurately, we define a new non-homogeneous Poisson process $S(t)$, for $0 < t \leq t_w$, with a time-varying rate function $r(t)$. The process $S(t)$ is a counting process that counts the total number of traffic events observed up to time instance t . In other words, $S(t_n)$ is the total number of IP-traffic counts from epoch 0 to epoch n . It can be seen that the non-homogeneous $S(t)$ process is in essence a cumulative function that aggregates the piecewise and homogeneous Poisson processes, $N_n(t)$. Thus $S(t)$ can be expressed in terms of $N_n(t)$ as:

$$S(t_n) = \sum_{i=1}^{t_i \leq t_{n-1}} N_i(t) + N_n(t) = S(t_{n-1}) + N_n(t), 1 < n \leq w. \quad (1)$$

Due to the above relation, the rate vector $\{r_n, n = 1, 2, \dots, w\}$ of $N_n(t)$ is a uniformly sampled version of $S(t)$'s rate function $r(t)$. That is, the time-varying rate function $r(t)$ at any time $t > 0$ can be expressed as:

$$r(t) = \sum_{n=1}^w r_n I_n(t), \quad 0 < t \leq t_w, \quad (2)$$

where $I_n(t) = \begin{cases} 1, & t_{n-1} < t \leq t_n \\ 0, & \text{otherwise.} \end{cases}$

Here the non-homogeneous Poisson (NPP) process $S(t)$ serves as a bridge between a traditional homogeneous Poisson process with a constant rate, say λ , (over all epochs) and the proposed PPM process with rates $\{r_n, n = 1, 2, \dots, w\}$ that vary after discrete epochs. For the NPP process $S(t)$, we define the expected value function $G(t) = E[S(t)]$ as the average number of arrivals over the epoch $(0, t]$. Therefore,

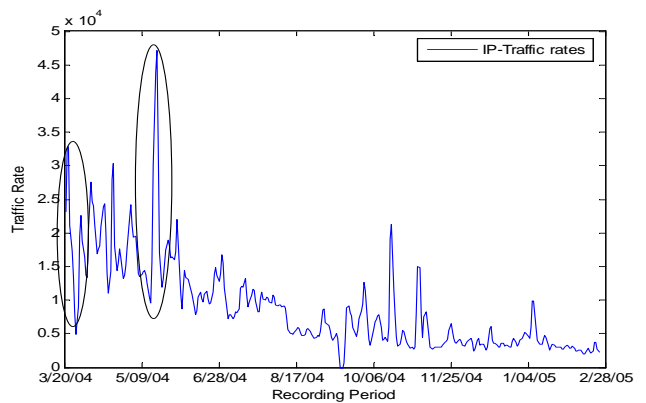


Fig. 1. The estimated IP-traffic rates.

NPP has the following generalized Poisson distribution:

$$\Pr[S(t) = k] = \frac{(G(t))^k}{k!} \exp(-G(t)).$$

Note that a traditional Poisson process is a special case of the NPP model when $G(t) = \lambda t$. Further note that the standard homogeneous rate λ simply represents the derivative of λt . Consequently, we compute the (time-dependent) rate $r(t)$ at time $t > 0$ of the NPP model using: $r(t) = dG(t)/dt$.

Now we can focus on estimating the rates $\{r_n, n = 1, 2, \dots, w\}$ of arrivals during the epoch $(0, t_n]$ of the PPM process $N_n(t)$ by utilizing the continuous-time rate function $r(t) = dG(t)/dt, 0 < t \leq t_n$. We revert back to the original problem that we only have access to discrete epoch-aggregated values in the present dataset. Consequently, we only have the discrete-time $\{G_n = E[S(t_n)], n = 1, 2, \dots, w\}$ counterpart of the expected value function $G(t) = E[S(t)]$, where G_n is directly approximated from the set of measurements (samples) that are being recorded in the IMS data as follows:

$G_n =$ number of traffic counts up to time $t_n, n = 1, 2, \dots, w$.

While the problem of discrete measurements prevails, due to the non-homogeneous process $S(t)$ this problem is transformed to a more tractable problem of accurately measuring the expected value of $G(t)$ and subsequently $r(t)$. The expected value function $G(t)$ can be simply obtained by fitting a regression function on G_n . It is worth mentioning that since $N_n(t) = S(t_n) - S(t_{n-1})$ on the epoch $(t_{n-1}, t_n]$ is a homogenous Poisson process with rate r_n , its corresponding expectation has a linear slope r_n . As a result, the expected value function $G(t) = E[S(t)]$ becomes a polygon with varying slopes $\{r_n, n = 1, 2, \dots, w\}$ on different epochs $\{(t_{n-1}, t_n], n = 1, 2, \dots\}$. To accurately fit a regression function on G_n that preserves the polygonal property of $G(t)$, we utilize a nonparametric regression approach and in particular the spline smoothing method [11].

Once the regressed spline function of $G(t)$ is approximated, the time-varying rate $r(t)$ of the continuous-time NPP process $S(t)$ is computed using $r(t) = dG(t)/dt$. The estimated sequence of rates $\{\hat{r}_n, n = 1, 2, \dots, w\}$ of the PPM process can then be directly computed by uniformly sampling $r(t)$ for each epoch. Figure 1 shows the estimated IP-Traffic rates

for the IMS data. The variations in traffic rates over the recording period suggest that the incoming traffic observed at the blackhole exhibits erratic behavior over time. For example, the traffic rates for the first 10 recordings (March 20th–29th 2004) and from epochs 55 to 70 (May 13th–28th 2004) are highlighted in Figure 1. We observe that the rates fluctuate significantly during these periods. We believe that a radical increase or decrease in the traffic rates can render robust indications of significant anomalous activity. In Section V, we describe how one can employ the traffic rate variations at a blackhole to detect substantially anomalous activity, such as a malware outbreak.

IV. REGRESSION MODEL OF PPM’S TRAFFIC RATES

Blackhole data is collected on a block of unused or dark address space. Since these address blocks are not assigned to valid hosts, we categorize incoming blackhole traffic into two clusters: (i) *abnormal* (but harmless) traffic due to misconfigurations, benign scanning/probing, or scan traffic from legacy malware which for the most part has been eradicated by patching; (ii) *malicious* traffic due to backscatter from a spoofed source addresses or scans by currently-active malware. In general, the abnormal activities have a low traffic load and repeatedly appear at a blackhole. On the other hand, malicious activities occur relatively infrequently, but they carry high traffic loads. In other words, abnormal traffic describes the *typical blackhole traffic*, while malicious traffic is the *anomalous blackhole traffic*.

The PPM models the traffic rates for every recording epoch. If the traffic contains only typical abnormal traffic then the corresponding traffic rates should exhibit slow-varying patterns throughout the recording period. However, the presence of anomalous blackhole traffic makes the traffic rates to abruptly deviate from the smooth and slow-varying trends. We observed that these abrupt patterns introduce heavy-tailed behavior in the rate distribution. To accurately capture this heavy tail pattern, we propose the following probabilistic regression model for rates:

$$r_n = f_n + \varepsilon_n, \quad n \geq 1, \quad (3)$$

where $\{f_n\}$ is a deterministic regression function and $\varepsilon = \{\varepsilon_n, n \geq 1\}$ are traffic rate residuals that are assumed to be independent, identically distributed (iid) random variables. Once an appropriate random variable is identified for the rate residuals, the statistics of that random variable can be used to determine an accurate threshold to detect anomalous blackhole patterns.

The presence of anomalous traffic has a direct impact on the distribution of the rate residuals, ε . Specifically, the magnitudes of the rate residuals associated with the anomalous blackhole traffic are much larger than the magnitudes of typical rate residuals. Consequently, the variance of the rate residual probability density tends to infinity, and the density function becomes heavy-tailed and skewed. Due to the heavy-tailed behavior, normal and exponential distributions are not suitable to model the residuals. To capture the heavy-tailed behavior of

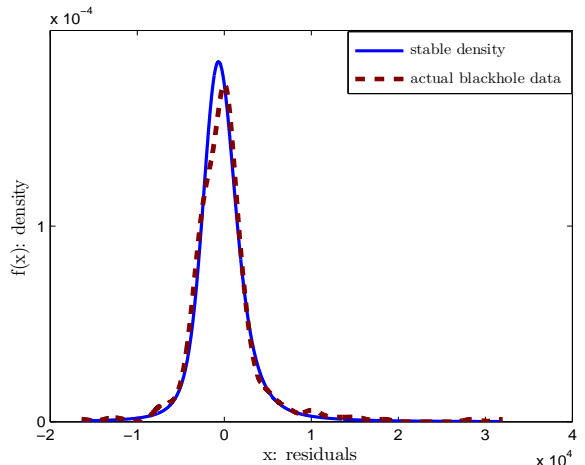


Fig. 2. Probability density of traffic rate residuals: Comparison of the stable density approximation with the actual density observed in the IMS data.

traffic rate residuals, we propose to use a stable distribution [9] to model ε .

We now explain the randomized regression structure for the IP-traffic rates; the structure for MD- and PC-traffic is similar. Using equation (3), let the rate residuals be written as

$$\hat{\varepsilon}_n = \hat{r}_n - f(t_n), \quad n = 1, 2, \dots, w \quad (4)$$

where \hat{r}_n is the estimated rate of the PPM process in epoch n , $f(t_n)$ is a deterministic regression function for the traffic in epoch n , and w is the total number of time epochs under consideration. The deterministic $f(t_n)$ function represents simple non-random trends in the blackhole data. For the present dataset, we observed that an exponential function of the form

$$f(t_n) = \sum_i a_i \exp(b_i t_n), \quad a_i, b_i \in \mathcal{R}, b_i < 0 \quad (5)$$

can adequately represent the deterministic trends of the IMS data. After subtracting these deterministic trends from the rates, the heavy-tailed behavior of the remaining traffic residual, $\hat{\varepsilon} = \{\hat{\varepsilon}_n, n = 1, 2, \dots, w\}$, is modeled using a stable distribution.

A general stable distribution is defined using four parameters [9]: a characteristic exponent $\alpha \in (0, 2]$, a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\gamma > 0$ and a location parameter $\delta \in \mathcal{R}$. We compute the following maximum likelihood estimates of these parameters using the STABLE program [12].

Figure 2 demonstrates the approximated stable density and the density of the residuals observed in the actual data. Figure 2 clearly shows that the stable distribution captures the heavy tail of the actual residual density very accurately.

V. THE PPM/RATE-RESIDUAL-BASED MALWARE DETECTION TECHNIQUE

The focus of this section is to filter typical blackhole traffic and to detect anomalous blackhole traffic. To that end, we

TABLE I
ATTACKS DETECTED BY THE PROPOSED DETECTOR

Traffic Type	Worm Name	Attacked Ports	Detection Date of Proposed Detector	Symantec Discovery Date
MD,IP	W32.Gaobot.SY	80 135 445	21-Mar-04	26-Mar-04
MD,IP	W32.HLLW.Gaobot.RS	80 135 445	21-Mar-04	22-Mar-04
IP,MD	W32.HLLW.Donk.L	4444	21-Mar-04	23-Mar-04
MD	W32.Gaobot.ZX	80 135 445	8-Apr-04	12-Apr-04
PC	W32.Blaster.T.Worm	135 4444	18-Apr-04	21-Apr-04
PC	Hacktool.LsassSba	137 138 139 445	25-Apr-04	27-Apr-04
IP	W32.Sasser.Worm	445 5554 9996	30-Apr-04	30-Apr-04
IP	W32.Welchia.K	80 135 445 3127	2-May-04	5-May-04
IP	W32.Bobax.C	445 5000	17-May-04	18-May-04
IP	W32.Gaobot.RB	80 135 445	25-May-04	26-May-04
IP,MD	W32.Gaobot.AQS	80 135 445	1-Jun-04	7-Jun-04
MD	W32.Gaobot.CEZ	80 135 445 8000	20-Jan-05	25-Jan-05

develop an unsupervised online detection technique that uses the piecewise Poisson model and the rate's residual distribution to detect the anomalous blackhole traffic.

A. Detection Algorithm

Consider a training dataset which contains the traffic monitored during a training period. The training dataset contains unlabeled typical and anomalous blackhole traffic. Clearly, mining anomalous traffic from this unlabeled dataset requires an unsupervised change detection algorithm. We develop such an algorithm using the PPM model and the stable distribution. The steps for the proposed online detection technique are as follows:

- 1) Given a training dataset, apply the PPM to estimate the continuous-time rate function of equation (2). Sample this function to generate the estimated rates of the PPM.
- 2) Using the estimated rates, compute parameters of the deterministic regression function of equation (5).
- 3) Compute the traffic rate residuals by subtracting the deterministic regression function from the rates [see equation (4)].
- 4) For online detection, fit a stable distribution on the rate residuals and set the threshold to be the x_τ quantile of the distribution. As mentioned in Section IV, the main objective of residual modeling is to ascertain an accurate detection threshold. Residual values above that threshold can then be flagged as anomalous. Since anomalous/malicious traffic carries high density rates, we set this threshold to be the x_τ quantile of the stable distribution with $\tau \in [0.95, 0.99]$.
- 5) Label the traffic for an epoch as anomalous if the residuals exceed the threshold.

B. Detection Results and Performance Comparison

To substantiate the efficiency of our detection technique, we compared the exact dates of anomalous epochs with their corresponding high-activity ports. We then searched the Symantec website [www.symantec.com] to find the worms that were reported at or around the anomalous epochs and which used the detected high-activity ports to propagate. We detect an attack only if all the ports that are exploited by the attack are

flagged as high-activity ports on the corresponding anomalous epoch. Due to brevity, Table I shows a sample of the detected attacks and their corresponding ports with a port threshold of $v = 0.95$.

VI. CONCLUSIONS

In this paper, we analyzed the blackhole traffic recorded by the Internet Motion Sensor project. A Piecewise Poisson process Model (PPM) was introduced for blackhole data. Using the PPM, we captured the traffic rates for every recording interval. A regression model of PPM's traffic rates was used to demonstrate that during a malware outbreak traffic rates vary dramatically and exhibit heavy-tailed behavior. We found that the stable distribution can accurately model the traffic rates. Using the statistical PPM and traffic rate models, we proposed a detection mechanism to detect malware outbreaks. We showed that the proposed detector detects significant threats in the analyzed data in a fast and accurate manner.

REFERENCES

- [1] M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson, "Internet motion sensor: A distributed blackhole monitoring system," *ISOC Network and Distributed System Security Symposium (NDSS)*, 2005.
- [2] D. Moore, "Network telescopes: Observing small or distant security events," *USENIX Security Symposium*, 2002.
- [3] E. Cooke, M. Bailey, Z. M. Mao, D. Watson, and F. Jahanian, "Toward understanding distributed blackhole placement," *ACM CCS Workshop on Rapid Malcode (WORM)*, 2004.
- [4] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the Slammer worm," *IEEE Security & Privacy*, 1(4):33–39, 2003.
- [5] D. Moore, G. M. Voelker, and S. Savage, "Inferring Internet denial-of-service activity," *USENIX Security Symposium*, 2001.
- [6] C. Shannon and D. Moore, "The spread of the Witty worm," *IEEE Security & Privacy*, 2(4):46–50, 2004.
- [7] M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson, "The Blaster Worm: Then and Now," *IEEE Security & Privacy*, (3)4:26–31, 2005.
- [8] A. Kumar, V. Paxson and N. Weaver, "Exploiting Underlying Structure for Detailed Reconstruction of an Internet Scale Event," *ACM Internet Measurement Conference (IMC)*, 2005.
- [9] M. S. Taqqu, G. Samorodnitsky, *Stable Non-Gaussian Random Processes: stochastic models with infinite variance*, CRC Press, 1994.
- [10] Symantec Internet Security Threat Reports, September 2004, March 2005, September 2005.
- [11] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, 2nd ed., 1999.
- [12] J. Nolan, STABLE program for Windows, <http://academic2.american.edu/~jpnolan/stable/stable.exe>