DDoS Attack Detection and Wavelets

Lan Li and Gyungho Lee
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, Illinois U.S.A
{lli, ghlee}@ece.uic.edu

Abstract— **This paper presents a systematic method for DDoS attack detection. DDoS attack can be considered system anomaly or misuse from which abnormal behavior is imposed on network traffic. Attack detection can be performed via abnormal behavior identification. Network traffic characterization with behavior modeling could be a good guidance of attack detection. Aggregated traffic has been found to be strong bursty across a wide range of time scales. Wavelet analysis is able to capture complex temporal correlation across multiple time scales with very low computational complexity. We utilize energy distribution based on wavelet analysis to detect DDoS attack traffic. Energy distribution over time would have limited variation if the traffic keeps its behavior over time (i.e. attack-free situation); while an introduction of attack traffic in the network would elicit significant energy distribution deviation in short time period. Our experimental results with typical Internet traffic trace show that energy distribution variance changes markedly causing a "spike" when traffic behaviors affected by DDoS attack. In contrast, normal traffic exhibits a remarkably stationary energy distribution. In addition, this spike in energy distribution variance can be captured in early stage of attack, far ahead of congestion build-up, making it an effective attack detection.**

I. INTRODUCTION

Distributed denial of service (DDoS) attack has been one of the major attention grabbing security attacks as it explicitly threatens the stability of the Internet. Computer Economics [4] estimated that the total economic impact of Code Red was $2.6 billion, and Sircam cost another $1.3 billion. A recent attack via SQL Slammer caused an estimated $1 billion in damage during the first five days as it rapidly spread around the globe [3]. Unlike denial of service attacks relying on specific network protocol or system weakness, the DDoS attack simply exploits the huge resource asymmetry between the Internet and the victim. A sufficient number of zombies generate huge "useless" traffic volume towards the victim. Through this "many to one" attack dimension, the DDoS attack is able to block the access to the "thoroughfare" reaching the victim, effectively taking the victim off the Internet so that any victim' level of defense becomes irrelevant. In addition, the DDoS attack's strategies of hierarchical attack and IP spoofing make attackers difficult to trace. Although great efforts has been involved in attack detection and prevention, there is still a lack of effective and efficient solutions to intercept ongoing attack in a timely fashion, i.e. short enough to prevent traffic build up from DDOS attack.

Several methods have been proposed for attack detection and prevention, such as pattern-based filtering, and queue management associated with flow state (e.g. LRU-RED) [6]. However, no common characteristics of DDoS packets can be used as general signatures of detection and filtering. Attackers can shape the volume of attack streams and vary all packet fields to avoid exposing their own identity. In addition, even if the detector (or filter) is able to identify the pattern of the attacks, massive amount of traffic may paralyze it and make it ineffective. That is the reason why most of current techniques are still unable to withstand large-scale attacks.

DDoS attack can be considered system anomaly or misuse by which abnormal behavior is imposed on network traffic. Attack detection becomes traffic behavior change identification. Traditional anomaly and misuse detections, however, are confined in detecting the deviation from preset reference (e.g. normal traffic pattern) or identifying traffic with known attack signature. The pattern and signature in use are still on packet or flow level, instead of traffic behavior level in which we believe traffic nature is presented. Network traffic characterization could be a good guidance of attack detection, as long as the traffic behavior can be explicitly captured. Recent researches have shown that time series of aggregated traffic is scale invariant or bursty across a wide range of time scales [2][5][8][9]. Since time scales can be naturally represented by wavelets [19] and wavelet representation also matches the properties of the bursty network traffic, wavelet-based scaling analysis has been applied to characterize the Internet traffic [18][19]. Analytic study in [18] shows that variances of wavelet coefficients are determined by the nature of traffic itself. All these propel us to develop energy distribution analysis based on wavelets, for traffic behavior characterization to detect DDoS attack. Following the wavelet method in [12] [25], energy distribution in traffic is defined through the variances of wavelet coefficients on the time series of network traffic.

We applied our traffic behavior characterization with energy distribution to DDoS detection. Our experimental results with Internet traffic trace show that energy distribution variance changes markedly as traffic behavior changes due to DDoS Attack, while normal traffic exhibits a remarkably stationary energy distribution. Furthermore, such change can be captured in a timely manner, i.e. short enough to prevent traffic build up from DDOS attack.

The rest of the paper is organized as follows. We first briefly introduce related work in Section II. We then propose our wavelet based energy distribution analysis in Section III. Normal traffic trace without evident behavior anomaly (including real and simulation trace) has been investigated for its energy distribution. In Section IV, through simulation, attack traffic (a typical cause of traffic behavior change) is studied with our energy distribution analysis. Finally, we conclude the paper in Section V.

0-7803-7945-4/03/$17.00 (C) 2003 IEEE

II. RELATED WORK

Several detection methods have been proposed against DDoS attack. Obviously, detecting a DDoS attack is relatively easy at victim network, since attack traffic near the victim is unusually overwhelming. Attack can be captured based on identifying unusually high traffic with certain classification (e.g. packet type). However, the responsiveness of this approach is fairly poor due to the downstream location. Moreover, if an upstream link has been jammed by attack packets, there is not much to do on victim side.

In contrast, attack packets with spoofed source address can be effectively detected at attack source side [1]. Network Ingress Filtering (NIF) [7] is based on this. Routers with NIF drop packets with illegitimate source IP addresses. However, this approach cannot capture attack packets generated by reflectors (here, source addresses are valid) [1]. In addition, effectiveness of this approach significantly depends on the coverage of NIF. Ensuring all ISP networks to install NIF is evidently not practical. Instead of source network, route-based packet filtering (RPF), proposed by Park and Lee, implements enhanced NIF in intermediate network [10]. RPF validates the route taken by the packets based on the inscribed source and destination addresses, and the BGP routing information. If the route includes an illegitimate path, the packet is considered an attack packet. RPF has more practicability and less coverage requirement than NIF. However, there are several problems that prevent wide deployment of these approaches, such as BGP modification, router overhead, and the lack of inter-domain cooperation. Moreover, similar to the NIF, the RPF approach cannot filter attack packets with valid source addresses (e.g. reflected packets).

Most of the methods introduced so far are based on appearance of DDoS attack, such as spoofed source IP address, bandwidth distribution, attack packet pattern etc. However, attacker can hide the appearance of attack traffic via packet reshaping. DDoS attack can be considered system anomaly or misuse from which abnormal behavior is imposed on network traffic. Some statistical approaches have been proposed for anomaly detection based on behavior profiling, such as neural networks [11], Markov models [22], and signal analysis [23]. Behavior profiles for subjects are initially generated. As the system continues running, the anomaly detection can be performed via the variance of the present profile from the original one. In network environments, traffic characterization mechanisms possessing the ability of behavior modeling can also be applied to attack detection against inscribed anomaly. In this paper, we propose the energy distribution analysis, a characterization mechanism of traffic behavior, to implement attack detection. This mechanism can detect traffic behavior change based on its inherent characteristic. In the following sections, we will present our proposed method in terms of basics of the technique employed and verification of its effectiveness via simulation.

III. ENERGY DISTRIBUTION ANALYSIS BASED ON WAVELETS

Measurements and analytical studies have shown that network traffic exhibits self-similarity or long-range dependence. With inherent scaling property, wavelet is well-suited for analyzing self-similar process [25][19]. Our proposed energy distribution analysis justifiably develops on the top of the wavelet technique proposed by Abry and Veitch [12]. It is also based on a conjecture that the Internet traffic is long-range dependent or self-similar. Although more complex, perhaps multifractal-like, scaling behaviors under sub-second scales have been reported in recent researches [25][26], we still consider self-similar scaling over large time scales (more than 100 ms) by which we believe traffic behavior change can be presented. We measured self-similarity of Internet traffic trace (ITA trace [14]) and found that ITA traces have high Hurst parameter [1] values (> .7) under large time scales for different detecting points and observation time windows. This is consistent with results found by Paxson and Floyd [5] using the same trace.

A. Wavelet analysis and energy distribution

1) Wavelet analysis

Wavelet analysis defines a collection of nested subspace V_j corresponding to a collection of scalable and shiftable functions $\Phi_{j,i}(t)$. Time series $x(t)$ is projected into each of the subspaces V_j:

$$\hat{x}_j(t) = (proj_{V_j} x)(t) = \sum_i a_x^j(i)\Phi_{j,i}(t), \quad a_x^j(t) = x(2^j t) \tag{1}$$

$(proj_{V_j} x)(t)$ is coarser than $(proj_{V_{j-1}} x)(t)$, so the key of wavelet analysis is to examine the loss of information (information difference). We define detail signals [12][17]:

$$Details_j(t) = \hat{x}_{j-1}(t) - \hat{x}_j(t) = (proj_{V_{j-1}} x)(t) - (proj_{V_j} x)(t) \tag{2}$$

$Details_j(t)$ can also be obtained from projecting x(t) onto a collection of subspaces W_j (called wavelet subspace). $\Psi_{j,i}(t)$ are wavelet functions used by projecting operation in wavelet space:

$$Details_j(t) = (proj_{W_j} x)(t) = \sum_i d_x(j,i)\Psi_{j,i}(t) \tag{3}$$

where $d_x(j,i)$ is wavelet coefficient. $d_x(j,i)$ can be considered independent and identical distribution variable with zero mean [12][19]. $|d_x(j,i)|^2$, as variance of $d_x(j,i)$, measures the amount of energy distributed at time instant $2^j i$ ($2^{-j}V_0$ in frequency domain) [12]. Using the average of $|d_x(j,i)|^2$, one can estimate the spectrum of x:

$$\hat{\Gamma}_x(2^{-j}v_0) = \frac{1}{n_j}\sum_i |d_x(j,i)|^2 \tag{4}$$

, where n_j is the available number of wavelet coefficients at j. $\hat{\Gamma}_x(2^{-j}v_0)$ is then measuring the energy that lies in subband with central frequency of $2^{-j}v_0$. We use E_j to represent energy in subband with central frequency of $2^{-j}v_0$.

In our study, we utilize a time series {x(t)}, in which x(t) is defined as the byte counts in a fixed time interval. We set a time interval of 10 milliseconds in our study as in Abry-Veitch wavelet analysis [24]. Our study also shows formed time series with 10 milliseconds interval is able to represent self-similarity of sampled trace, while adequate data can be collected in available time window(s). Other two parameters, sliding window W and time step increment T, are also utilized in our

[1] Hurst parameter (H) presents degree of long-range dependence

study. Every time of T, traffic is sampled with size of W.

2) Energy distribution

Wavelet analysis actually uncouples the scaled traffic. $|d_x(j,i)|^2$ tells us how much difference (dissimilarity) between the two neighboring scaled traffic patterns. On the other hand, having all $|d_x(j,i)|^2$ we can reconstruct the signal series $x(t)$. We notice that the wavelet spectrum provides complete information of the correlation structure of given processes without any loss. In other words, energy distribution (spectrum) has great potential to characterize traffic behavior.

If we observe the traffic at two consecutive points (we use a sliding sampling window of size W with an incremental time step of T), we have energy distribution E_j^2 at second point and E_j^1 at the first point. The variation between E_j^2 and E_j^1 may show the characteristic /behavior change in observed traffic. Because of significant autocorrelation in large time scale (i.e. long-range dependence), the variation of E_j is very limited if the traffic has no characteristics/behavior change. We measured energy distribution and its variation in Internet traffic (using ITA trace [14]) and found that is the case (see Figure 1(a)(b), showing an example from our experimental results). Under given sampling window (21 minutes), energy distribution variation [(2)] of every trace is quite small (<0.15). Since ITA traces were captured at separate time slots in a day, it implies that daily traffic change has little impact on energy distribution property. Throughout this verification procedure, the energy distribution has stayed relatively constant.

Although real trace is preferred in network traffic study, it has limitations, such as short length and fixed network context. Simulated trace is then considered in our study. We can check if our traffic characterization with energy distribution works in simulated trace. Through NS simulator [13], we setup a dumbbell-topology (similar topology has been used in [18] [19]) and typical web workload (similar to SURGE developed at Boston University [20]). Dumbbell topology (see Figure. 2(a)) consists of 40 web server pools, 420 clients and 7 intermediate nodes. One bottleneck exists in the link between the servers and clients. During the simulation, we sweep the number of web sessions from 500 to 3000 and obtain packets trace at bottleneck link. With the same sliding sampling window and incremental time step applied to ITA trace, we have energy distribution variations of simulation trace (shows in Figure 2. (b)). Since we extend simulation time to 180 minutes and only extract the middle section of trace for analysis, trace can be considered stable. Object trace presents nice similarity (Hurst Parameters > 0.8) and stationary (we can also say traffic keeps its characteristics/behavior), so the variation of E_j shows very little variation (<0.01). This results match with our findings with the real Internet trace, ITA trace, shown in Figure 1 (a) and (b).

B. Energy distribution analysis

Since energy distribution of Internet traffic changes little, we conjecture that any anomaly in traffic, like attack traffic, will cause a sudden change in energy distribution for a short time window. Based on this we develop a threshold based traffic signature as follows:

Suppose two time series $x(t)$ and $x(t+\tau)$ are monitored successively and let's define $Eg^t{}_j = \frac{1}{n_j}\sum_k |d^t{}_x(j,k)|^2$ and $Eg^{t+\tau}{}_j = \frac{1}{n_j}\sum_k |d^{t+\tau}{}_x(j,k)|^2$ as the energy function of $x(t)$ and $x(t+\tau)$, respectively. Then, the difference of energy distribution in two time series is

$$\Delta Eg_j = \log Eg^t{}_j - \log Eg^{t+\tau}{}_j = \log \frac{Eg^t{}_j}{Eg^{t+\tau}{}_j} \quad (5)$$

We consider the variance of ΔEg_j, i.e. energy distribution variation in two time series, to be the traffic signature. Thus we define the normal traffic as time series

$$x(t) \in \{x(\tau) \mid \mathrm{var}(\Delta Eg_j) < \delta, \tau > T\} \quad (6)$$

,where δ is a threshold and T is a time step increment for sliding sampling window. For a given value of δ, the traffic behavior is deemed to be normal as long as the traffic signature var(ΔEg_j) is not larger than δ. Since δ and T should reflect traffic behavior, they may be adaptively adjusted. Also, sampling window size W may be sensitive to traffic behavior. In our later simulation, DDoS attack traffic is employed as a cause of traffic behavior change, resulting in noticeable change in energy distribution variation. However, note that other anomalies, i.e. deviation from "normal" traffic, can be captured in energy distribution variation

C. Method Limitations and Discussion

1) Trace size

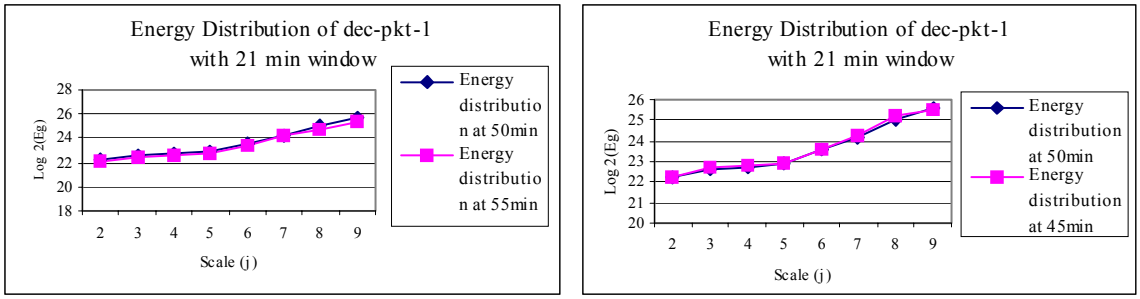
According to the trace investigation in Subsection A, our method requires a certain size of sliding sampling window. Since wavelet analysis demands number of input data must be power of 2, window size needs to follow this rule. After we tried a series of window sizes, 43min, and 21min [(3)] were selected for our experiments. Smaller window size may not provide enough samples to build up traffic self-similarity, while too large window size may cause unnecessary computation during the analysis and weaken the energy distribution variation. As a guidance of network control, our method may apply to traffic on-the-fly. Compared with other studies [18][19], however, 43 minutes trace seems to be quite long for real time analysis, especially for high bandwidth link (may have longer data length). Fortunately, an on-line version of the Abry-Veitch wavelet analysis has been proposed [24]. With the filter-banks, it can effectively process sampling data without redundant computation. It also has low memory requirement and scales naturally to arbitrarily high data rates for real time analysis.

2) Boundary effect

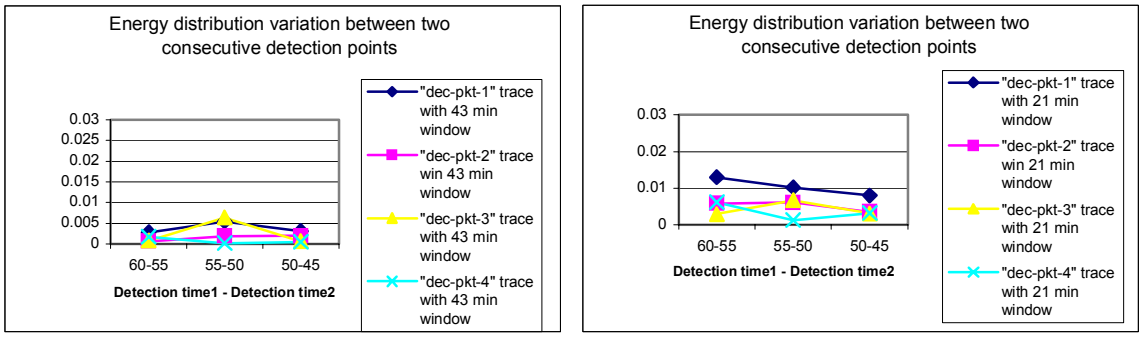
Boundary effect can exist in wavelet analysis. Given input data, how to select proper range (scale j) of wavelet coefficients was the problem mentioned in [12][17][24]. Since our method is based on Abry-Veitch wavelet analysis, we also need carefully choose wavelet coefficient to mitigate boundary effect. Roughan [24] suggested the upper bound of scale ($j_{\max}'$)

[(2)] The definition of energy distribution variation is described in Subsection B

[(3)] window sizes are 2621.44 and 1310.72 seconds (a power of 2 times 10 ms).

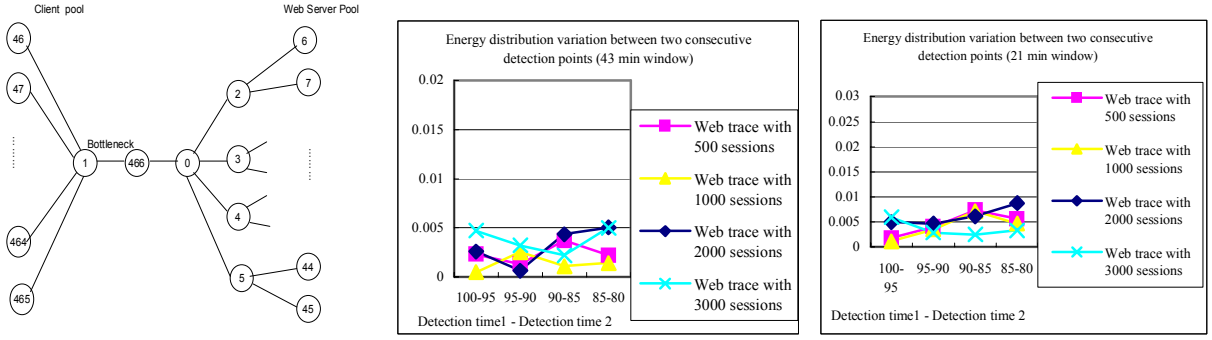


(a) Energy distribution in ITA trace dec-pkt-1



(b). Energy distribution variations in ITA traces

Figure 1. Energy Distribution in Traffic Trace (ITA Real Trace)



(a) Dumbbell-topology

(b). Energy distribution variations in NS-2 simulation trace

Figure 2. Energy Distribution in Traffic Trace (Simulated Trace)

should be less than $\log_2 n$, the largest scale in sampling data, where n is the length of sampling data. Due to initialization errors in wavelet decomposition, the lower bound of scale ($j_{\min}'$) is larger than one. However, there is no rule that can tell what is the best range of scale for given sampling data. In our practice, we select a range based on visual inspection of log-scale diagrams for a given network environment. We set ($j_{\min}'$, $j_{\max}'$) to (2,9) for ITA trace, while (2,10) for simulated trace.

3) Load effect

In order to check load effect on energy distribution analysis, we would have obtained complete picture of energy distribution analysis over all load levels (based on average link utilization).

However, due to the difficulty of network measurements under all load levels, we perform our investigation on simulated traffic (see Figure 2 (b)). In the simulation described in Section 3.1, we sweep the number of sessions from 500 to 3000 to build workloads with average link utilization varying from 19% to 95%. Note that our method is applicable to traffic with moderate and high load. We found a significant deviation in very light load (only 200 sessions and 6% link utilization). Figure 3 shows the comparison between energy distribution variation of light load and that of moderate load. Lower traffic base holds more volatile energy distribution, because even a few occasion of modest change is more distinctively reflected to energy distribution than in moderate or heavy load. Therefore, our energy distribution analysis is limited to traffic with moderate or high load. Since DDoS attack detection does not kick in with low load, this does not cause a practical limit of our approach.

IV. ATTACK DETECTION SIMULATION

Attack traffic is capable of making a sudden change, distorting normal behavior. Stationary energy distribution could be broken, since the sudden behavior change distorts temporal correlation over multiple scales.

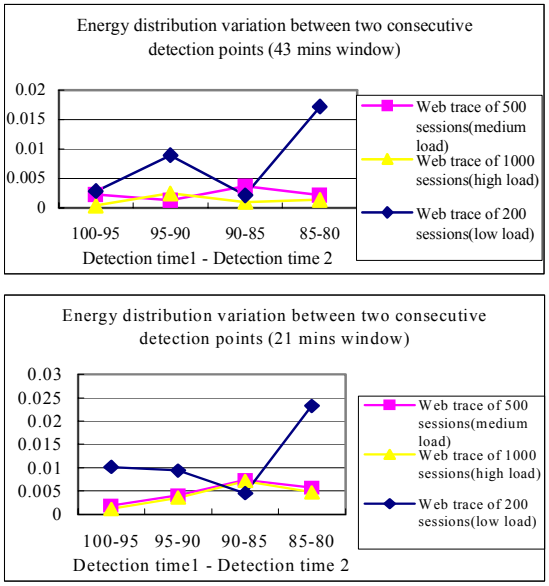


Figure 3. The deviation of energy distribution in light traffic

According to equation (6), our detection method can be considered a threshold-based method. Energy distribution variation caused by anomaly behavior (e.g. attack) could violate the threshold δ . DDOS attack having significant variation of energy distribution could then be detected. In order to catch attack as early as possible, sequential sampling with sliding window W is employed. Every time step increment T, traffic is sampled with size of W. For online detection, energy distribution analysis should be completed in time of T.

A. Simulation Environment

In order to simulate attack flows imposed to background traffic with self-similarity, we constructed a large-scale network simulation test-bed through NS simulator [13].

All the network end nodes in the simulation are assumed to be both IP traffic generators and receivers. General nodes and hot spot nodes (victim) would be simulated. In order to avoid a deterministic impact of statistic based traffic generator (such as Parato and Exponential), application based traffic sources are selected in our simulation, e.g. web, ftp, and CBR. In the simulated network, a number of network nodes are selected to be attacker nodes. Different from the normal traffic generator, IP traffic from all attackers has the same destination, the victim node. According to non-responsive feature, CBR traffic source is chosen for simulating UDP flooding attack. The attack scenarios simulated is based on attack observations done by [15][16].

We have two scenarios: 0.05 (scenario 1) and 0.075 (scenario 2) attack coverage, which is defined as the ratio of attack nodes to whole nodes. In each scenario, cases with attack and without attack are both simulated and the DDoS attack is launched at 3100s with exponential acceleration having knee point at 3500s (Network topology and other parameters including attack configuration for the experiments are described in Figure 4).

B. Simulation result

As the first step, the self-similarity of traffic is extracted by estimating Hurst parameter. As in the case of ITA trace, the simulated traffic also exhibits a quite high self-similarity (with and without attack, all scenario cases produced Hurst parameter values in the range of 0.7 and 0.8). We then applied our method to compute the energy distribution variation of different traces.

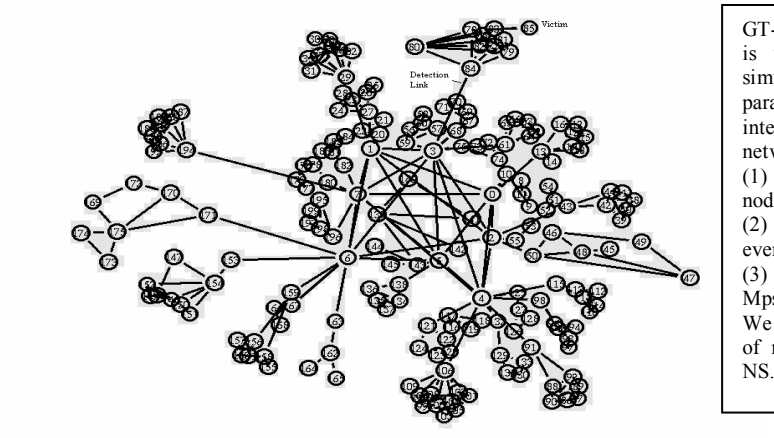
In the simulation results we have, large difference between consecutive detection points is observed in the traces with attack (see Figure 5). As a comparison, the traces without attack have a very limited energy distribution variation, which is very similar to what we obtained from ITA trace. With a threshold of 0.01, our scheme was able to catch all attack cases; four of four cases. The catch points match with attack launch timing shown in Figure 4(c). With the two attack-coverage scenarios and the two sliding window sizes, they are all around 3400s. One important note here is that energy distribution variation analysis is able to catch attacks early in the attack launch, far ahead of congestion build-up due to the attacks.

In contrast, we also show the variation of traffic rate [4] in Figure 6. It is clear that DDoS attack elicits a significant rate change to the traffic. However, rate variation in the early stage of attack (before 3800s in scenario 1) may not be detected by rate watching schemes (e.g. rate threshold), since it is as limited as normal burst. When a significant variation happens (around 4300s), congestion has already built up. With respect to the detection point showing in Figure 5, energy distribution analysis with sliding window of 21 minutes can detect attack at 3400 second, far ahead of congestion.

C. Discussion

We successfully utilize energy distribution to detect DDOS attack in simulation. The deviation of energy distribution variation caused by attack traffic is significant enough to be detected through a threshold $\delta = 0.01$. System parameters, δ ,T and W are chosen tentatively. With too large W and/or T, energy distribution variation may be buried under self-similarity, while too small W and/or T will make a less meaningful stochastic sample. Threshold δ and window size T need in-depth investigation with diverse traffic environment. Developing a proper value for δ , T , and W in various contexts is a key component of our future research. Another issue is sampled target (traffic parameter). Time series $x(t)$ could be sampled for any traffic parameter. Instead of using inter-arrival time, some other traffic parameters could be considered, such as connection amount, packet address distribution, etc. They may represent traffic behavior in different

[4] We only present the result of scenario 1 because of space limitation. Also, we have better detection in scenario 2.

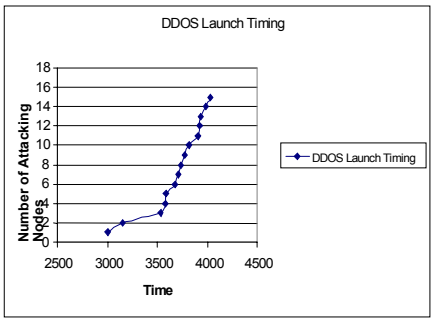


GT-ITM Topology Generator provided by NS-2 [13] is used to generate a three-level network for simulation. Following topology generation parameters are specified: (1) ratio of end nodes and intermediate nodes, (2) connection density of the network nodes and (3) link bandwidth assignment.
(1) Three-level hierarchy: domain, cluster, and nodes.
(2) 10 domains; 4 clusters every domain; 5 nodes every cluster
(3) 10 Mps link for domain; 5 Mps for clusters; 2 Mps for nodes
We scale down link bandwidth and traffic, because of memory usage problem in large simulations of NS.

(a) Network simulation topology

Total nodes	200
Number of background flows	800 (4 flows every nodes in average)
Attack coverage	0.05/0.075 (10/15 nodes, out of 200 total nodes, to be attackers or zombies)
Attack launch curve	Exponential distribution (see Figure 4 (c))
Simulation time:	6000 seconds
Attack period	From 3000s to 5000s
Attack launching period	3100s~4000s
Victim node	85
Detecting path	3->84->85 (node # 3 is the gate way collecting data in the simulation)

(b) Attack configuration



(c) DDoS launch timing

Figure 4. Simulation configurations

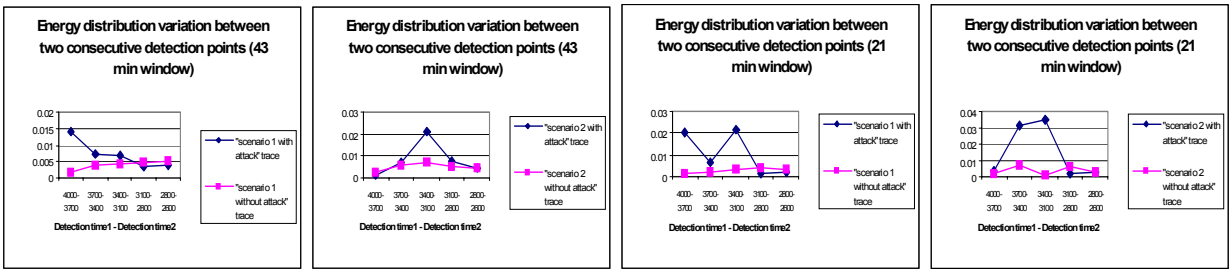


Figure 5. Energy distribution variations in simulated trace

aspects. For diverse applications beyond DDoS detection, we may find more suitable traffic parameter through which efficacy of our method can be improved.

Although this method is still a reactive one, catching attacks behind starting point, it can detect early in the attack launch (cooperate with real-time wavelet analysis), far ahead of congestion build-up. Compared to resource control scheme (such as rate limitation) associated with SRD characteristics of traffic (such as mean, and variance), our method may have better responsiveness and accuracy. Misuse detection schemes, based on preformed patterns, may recognize known "bad" behavior. However, recognizing “known” pattern may cause rather serious overhead due to packet decomposition in high level (such as IP address or TCP/UDP port checking). In addition, how to deal with unknown behavior in proactive way is a very complicated issue with no known acceptable solution. Therefore, our method may outperform existing schemes in attack detection. Also, in cooperation with distributed detection mechanism, we can envision better performance.

V. CONCLUSION

This work is motivated by the fact that abnormal traffic behavior imposed by DDoS attack can be detected via energy distribution based on wavelet analysis. We have shown potential of energy distribution analysis for characterizing network traffic behavior. Wavelet analysis is able to capture complex temporal correlation across multiple time scales with very low computational complexity. Wavelet analysis provides energy distribution data for complete information of traffic behavior. With investigation of both real and simulated traffic trace, we have shown that energy distribution keeps relatively stationary if the traffic has no characteristics/behavior change. Energy distribution analysis based on wavelet analysis then has been developed.

We have applied the energy distribution analysis to detect DDOS attack as a case study to verify the

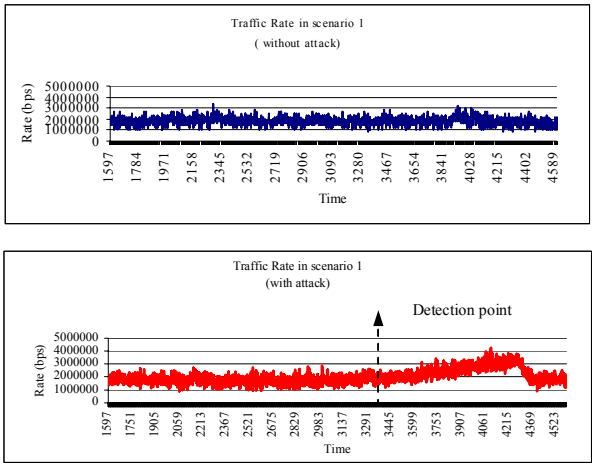


Figure 6. Traffic rate in scenario 1

detection capacity of our method. Parameters of detection method, time step increment T, threshold δ , and sampling window size W, have been studied. Tentative parameter values drawn from our experimental experience have been utilized in simulation. Our results show that energy distribution variance changes markedly, when attack traffic is injected, while normal traffic exhibits a remarkably stationary energy distribution. Our experimental results support that energy distribution analysis can characterize behavior of network traffic under dynamic condition and outperform other existing schemes.

In our study, only Inter-arrival time has been used to construct analyzed date, because it has been widely used in modeling self-similar traffic. However, some other traffic parameters could be considered, such as connection amount, packet address distribution, etc. They represent network traffic behavior in different perspectives. Extending our method to those parameters may improve characterization/detection performance. There seems a great potential for energy distribution to help making better decision for network control and management. We will also study what could be proper time step increment T, and sampling window size W in different network environment and a method through which parameters can be adaptively adjusted.

ACKNOWLEDGEMENT

This work is supported in part by the NSF grant CCR-0242222 and CCR-0209078.

REFERENCES

[1] Rocky K. C. Chang, "Defending against flooding-based distributed denial-of-service attack: a tutorial," IEEE Comm. Magazine, Vol.40 No.10, Oct. 2002, pp. 42-51.

[2] M.E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes," IEEE/ACM Trans. on Networking, Vol.5 No.6, Dec. 1997, pp.835--846.

[3] M. LaMonica, "Microsoft releases anti-Slammer tools," http://zdnet.com.com/2100-1105-983603.html, Feb. 6, 2003.

[4] CERT, "Overview of attack trends," http://www.cert.org/archive/pdf/attack_trends.pdf, Apr. 8, 2002.

[5] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling," IEEE/ACM Trans. on Networking, Vol. 3 No. 3, June 1995, pp.226-244.

[6] S. Sarvotham, R. Riedi, and R. Baranuik, "Connection-Level Analysis and Modeling of Network Traffic," in Proc. of the ACM SIGCOMM IMW, Nov. 2001.

[7] P. Ferguson and D. Senie, "Network ingress filtering: defeating denial of service attacks which employ IP address spoofing," Internet Draft, Jan. 1998.

[8] W. Leland, M. Taqqu, W. Willinger and D. Wilson, "On the self-similar nature of Ethernet traffic," In Proc. ACM SIGCOMM, Vol.23 No.4, Aug. 1993, pp.183-193.

[9] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (Extended Version)," IEEE/ACM Trans. on Networking, Vol. 2 No.1, Feb. 1994, pp.1-15.

[10] K. Park and H. Lee, "On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack," in Proc. of IEEE INFOCOM'01, Apr. 2001, pp. 338–347.

[11] K. Fox, R. Henning, J. Reed, and R. Simonian, "A Neural network approach towards intrusion detection," Technical Report, Harris Corporation, July 1990.

[12] P. Abry and D.Veitch, "Wavelet analysis of long range dependent traffic," IEEE Trans. on Infor. Theory, Vol. 44, No. 1, Jan. 1998, pp.2-15.

[13] The Network Simulator-ns-2. Available from "http://www.isi.edu/nsnam/ns/"

[14] The Internet traffic archive. "http://ita.ee.lbl.gov/", Apr. 2000.

[15] CERT Advisory CA-2001-23, Continued Threat of the "Code Red" Worm, Jan. 17, 2002.

[16] D. Moore, C. Shannon, and J. Brown, "Code-Red: a case study on the spread and victims of an Internet worm," In Proc. Internet Measurement Workshop, 2002.

[17] P. Abry, D. Veitch, and P. Flandrin, "Long-range dependence: revisiting aggregation with wavelets," Journal of Time Series Analysis, Vol. 19, May 1998. pp. 253-266.

[18] X. Tian, J. Wu, C. Ji, "A unified framework for understanding network traffic using independent wavelet models," In Proc. IEEE INFOCOM, June 2002.

[19] S. Ma and C. Ji, "Modeling heterogeneous network traffic in wavelet domain," IEEE/ACM Trans. Networking, Vol.9, No. 5, October 2001, pp. 634-649.

[20] P. Barford and M.E. Crovella, "Generating representative Web workloads for network and server performance evaluation," in ACM SigMetrics, 1998, pp.151-160.

[21] R. Ritke, X. Hong, and M. Gerla, "Contradictory relationship between Hurst parameter and queuing performance (extended version)," Telecommunication Systems, Vol 16, Feb. 2001, pp.159-175.

[22] N. Ye, "A Markov chain model of temporal behavior for anomaly detection," in Workshop on Information Assurance and Security, June 2000.

[23] P. Barford, J. Kline, D. Plonka and A.Ron, "A signal analysis of network traffic anomalies," Internet Measurement Workshop, Nov. 2002.

[24] M. Roghan, D. Veitch, and P. Abry. "Real-time estimation of the parameters of long-range dependence," IEEE/ACM Trans. on Networking, Vol.8, Aug. 2000, pp 467-478.

[25] R. Riedi, M. S. Crouse, V. Ribeiro, and R. G. Baraniuk, "A multifractal wavelet model with application to TCP network traffic," IEEE Trans. Info. Theory, Special issue on multiscale statistical signal analysis and its applications, vol. 45, pp. 992-1018, April 1999.

[26] Z. Zhang, V. J. Ribeiro, S. Moon, and C. Diot, "Small-time scaling behaviors of Internet backbone traffic: an empirical study," In Proc. IEEE INFOCOM, April 2003.