Spectral Flatness Measurements for Detection of C-Worms

Rajesh Jaladi [#1], Mr. Rakesh Nayak [#`2]

#1M.tech Student,Dept of CSE, [1] Sri Vasavi Engineering College, Tadepalligudem, Andhra Pradesh,

#2Assoc.Professor,Dept of IT, Sri Vasavi Engineering College, Tadepalligudem, Andhra Pradesh.

***Abstract*-- Active Worms such as Morris(1988), CodeRed(2001)[1], Nimda(2001), Slammer worm(2003),Blaster(2003) and Witty(2004) had always caused large parts of the Internet to be temporarily inaccessible, costing both public and private sectors millions of dollars. We identified a hard to detect new class of worms like C-worms that has the ability to camouflage its propagation by intelligently manipulating its scanning traffic volume over time so that its propagation goes undetected by the existing worm detection schemes and file-sharing worms which propagate within a relatively smaller community. Genuine traffic and these worms traffic are almost identical and barely noticeable in time-domain. So we propose a new Digital Signal Processing (DSP) scheme that uses power spectral density (PSD) distribution and energy spectral density (ESD) distribution in frequency domain for tracking spectral flatness measure of C-worms and tracking high CPU usage of most other worms with similar varying energy patterns. Performance evaluations on our proposed detection scheme confirm our claims.**

***Keywords*—PSD, ESD, SFM.**

I INTRODUCTION

A virus or worm is a computer program that can spread across computers and networks by making copies of itself, usually without the user's knowledge. Viruses can have harmful effects. These can range from displaying irritating messages to stealing data or giving other users control over your computer. A computer worm is a standalone malware computer program that replicates itself in order to spread to other computers. Often, it uses a computer network to spread itself. This is due to security shortcomings on the target computer. Unlike a computer virus [3], it does not need to attach itself to an existing program. Worms [2] almost always cause at least some harm to the network, even if only by consuming bandwidth, whereas viruses almost always corrupt or modify files on a targeted computer.

Worm Detection Schemes[6] such as statistical detection schemes, distributed attack detection using cooperating end-hosts scheme, traffic monitoring schemes[7] do well in detecting these worms. But a new class of smart-worm called CWorm, which has the capability to camouflage its propagation and further avoid the detection. The investigation showed that, although the C-Worm successfully camouflages its propagation in the time domain, its camouflaging nature inevitably manifests as a distinct pattern in the frequency domain.

Based on observation, a novel spectrum-based detection scheme to detect the C-Worm. The evaluation data showed that a scheme achieved superior detection performance against the C-Worm in comparison with existing representative detection schemes. So better detection schemes were required. Genuine traffic and c worm's traffic are almost identical and barely noticeable in time-domain. So a new Digital Signal Processing (DSP) scheme that uses power spectral density (PSD) distribution or energy spectral density (ESD) distribution with a pre-defined specific energy pattern to detect c worms was developed.

Power spectral density (PSD), or energy spectral density (ESD), is a positive real function of a frequency variable associated with a stationary stochastic process, or a deterministic function of time, which has dimensions of power per hertz (Hz), or energy per hertz. It is often called simply the spectrum of the signal. Intuitively, the spectral density measures the frequency content of a stochastic process and helps identify periodicities. Power Spectral Density distribution of worm detection data needs to transform from time domain to frequency domain.

In this paper, we propose to extend Digital Signal Processing (DSP) scheme that uses power spectral density (PSD) distribution and energy spectral density (ESD) distribution with many pre-defined specific energy patterns to detect c worms and many other smart worms, file sharing worms, internet worms alike. The result is an integrated systems that can detect a variety of computer worms including the smart ones such as C worms.

II RELATED WORK

Active worms [4][5] are similar to biological viruses in terms of their infectors and self-propagating nature. They identify vulnerable computers, infect them and the worm-infected computers propagate the infection further to other vulnerable computers.

The basic form of active worms [8] is the Pure Random Scan (PRS) worm, where a worm infected host continuously scans a set of random Internet IP addresses to find new vulnerable hosts. There are several variants of the PRS worm such as local subnet scan worm and hitlist scan worm. Both of these worms attempt to speed up their propagation by increasing the probability of successful scanning. However, it is hard to achieve large scale of worm propagation using pure local subnet scan or hit-list scan strategy due to their limitations in finding large number of vulnerable hosts. Consequently, PRS scan strategy is still widely adopted in worms and other strategies are used to speed up the worm propagation at different stages during the propagation. In addition, worms [9] use different scan strategies during different stages of propagation. In order to increase propagation efficiency, they use a local network or hitlist to infect previously identified vulnerable computers at the initial stage of propagation. They may also use DNS, network topology, and routing information to identify active computers instead of randomly scanning IP addresses.

On July 19, 2001, more than 359,000 computers connected to the Internet were infected with the Code- Red (CRv2) worm in less than 14 hours. The cost of this epidemic, including subsequent strains of Code-Red, [1] is estimated to be in excess of $2.6 billion. Despite the global damage caused by this attack, there have been few serious attempts to characterize the spread of the worm, partly due to the challenge of collecting global information about worms. We also qualified the effects of DHCP on measurements of infected hosts and determined that IP addresses are not an accurate measure of the spread of a worm on timescales longer than 24 hours. The C-Worm studied in this paper aims to elude the detection by the worm defense system during worm propagation. The C-Worm also shares some similarity with stealthy port-scan attacks. Such attacks try to find out available services in a target system, while avoiding detection. Due to the nature of self propagation, the C-Worm must use more complex mechanisms to manipulate the scan traffic volume over time in order to avoid detection.

III PROPOSED SCHEME

The C-Worm is different from traditional worms because of its ability to intelligently manipulate its scan traffic volume over time. Thereby, the C-Worm camouflages its propagation from existing worm detection systems based on analyzing the propagation traffic generated by worms. The C-Worm adapts their propagation traffic patterns in order to reduce the probability of detection, and to eventually infect more computers. The C-Worm is different from polymorphic worms that deliberately change their payload signatures during propagation. The Major Advantage of the C- Worm is,it scans all the IP Present in the Network first then identifies the number of protected systems, number of Worm Affected Systems, number of Vulnerable Systems. The Main aim of C-Worm is the overall scan traffic for the C-Worm should be comparatively slow and variant enough to not show any notable increasing trends over time. The C-worm and non worm network traffic is need to analyze.

In this paper, we propose to extend Digital Signal Processing (DSP) scheme that uses power spectral density (PSD) distribution or energy spectral density (ESD) distribution with many pre-defined specific energy patterns to detect c worms and many other smart worms, file sharing worms, internet worms alike. The proposed scheme is used to distinguish the non worm and the C-worm traffic. The Power Spectral Density and its corresponding Spectral Flatness Measure is used, the PSD distribution for worm detection data the data need to transform data from the time domain into the frequency domain. The C-worm can be detected only in frequency domain. The SFM values are comparatively very small than the SFM values of normal non worm scan traffic. Thus the worm is identified and alerts the system. Each and every time of scan it scan the unoccupied IP address. When the worm is detected the patch file is used to clear the worm. The IP trace back is used to find the source node which propagates the worm and eliminates such type of system from the network.

A) Power Spectral Density (PSD)

To obtain the PSD distribution for worm detection data, data in time domain is transformed into the frequency domain. The PSD function of the scan Transform (DFT) of its autocorrelation function. As the PSD inherently captures any recurring pattern in the frequency domain, the PSD function shows a comparatively even distribution across a wide spectrum range for the normal nonworm scan traffic. The PSD of C-Worm scan traffic shows spikes or noticeably higher concentrations at a certain range of the spectrum.

$$\Phi(R[L],k) = \Sigma_{n=0}^{N-1} (E[X(t)X(t+L)]). e^{-j2\pi Kn/N}$$

Where k=0,1,…,N-1.X(t) is the random process to model the worm detection data and R[L] is the correlation of worm detection data in an interval.

PSD tells that at which frequency ranges variations are strong and that might be quite useful for further analysis. The concept and use of the power spectrum of a signal is fundamental in electrical engineering, especially in electronic communication systems, including radio communications, radars, and related systems, plus passive [remote sensing] technology.

B) Spectral Flatness Measure (SFM)

We measure the flatness of PSD to distinguish the scan traffic of the C-Worm from the normal nonworm scan traffic. For this, we introduce the SFM, which can capture anomaly behavior in certain range of frequencies. The SFM is defined as the ratio of the geometric mean to the arithmetic mean of the PSD coefficients. In statistical signal processing and physics, the spectral density, power spectral density (PSD), or energy spectral density (ESD), is a positive real function of a frequency variable associated with a stationary stochastic process, or a deterministic function of time, which has dimensions of power per hertz (Hz), or energy per hertz.

$SFM = [\pi_{k=1}^{N} PSD(f)]^{1/n} / {}_{[(1/n)(} \Sigma_{k=1}^{N} PSD(f))]$ where f is frequency.

It is often called simply the spectrum of the signal. Intuitively, the spectral density measures the frequency content of a stochastic process and helps identify periodicities. SFM is a widely existing measure for discriminating frequencies in various applications, such as voiced frame detection in speech recognition. In general, small values of SFM imply the concentration of data at narrow frequency spectrum ranges.

C) Worm Detection

As the SFM value can be used to sensitively distinguish the C-Worm and normal nonworm scan traffic, the worm detection is performed by comparing the SFM with a predefined threshold Tr. In the worm detection systems, monitors collect port-scan traffic to certain area of dark IP addresses and periodically reports scan traffic log to the data center. Based on the continuous reported data, the value of Tr will be tuned and adaptively used to carry out worm detection. If we can obtain the PDF of SFM values for the C-Worm through comprehensive simulations and even real-world profiled data in the future, the optimal threshold can be obtained by applying the Bayes classification.

If the PDF of SFM values for the C-Worm is not available, based on the PDF of SFM values of the normal no worm scan traffic, we can set an appropriate Tr value.

SCHEME	VARIANCE	TREND	MEAN	PROPOSED SCHEME
DETECTION RATE(DR)	48%	0%	14%	96.4%
MAXIMAL INFECTION RATIO (MIR)	14.4%	100%	7.5%	4.4%
DETECTION TIME(DT) IN MINUTES	2367	α	1838	1707

Table: Detection results

D) Energy spectral density

The energy spectral density describes how the energy (or variance) of a signal or a time series is distributed with frequency. If f(t) is a finite-energy (square integrable) signal, the spectral density Φ(ω) of the signal is the square of the magnitude of the continuous Fourier transform of the signal (here energy is taken as the integral of the square of a signal, which is the same as physical energy if the signal is a voltage applied to a 1-ohm load, or the current).

$$\Phi(\omega) = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t}\, dt \right|^2 = \frac{F(\omega) F^*(\omega)}{2\pi}$$

Where ω is the angular frequency (2π times the ordinary frequency) and F(ω) is the continuous Fourier transform of f(t), and F* (ω) is its complex conjugate.

If the signal is discrete with values f_n, over an infinite number of elements, we still have an energy spectral density:

$$\Phi(\omega) = \left| \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n e^{-i\omega n} \right|^2 = \frac{F(\omega) F^*(\omega)}{2\pi}$$

Where (ω) is the discrete-time Fourier transform of f_n.

If the number of defined values is finite, the sequence does not have an energy spectral density per se, but the sequence can be treated as periodic, using a Discrete Fourier Transform (DFT) to make a discrete spectrum, or it can be extended with zeros and a spectral density can be computed as in the infinite-sequence case.

IV PERFORMANCE

Performance of our proposed scheme is estimated by comparing its performance with three existing popular worm detection schemes. The first scheme is the volume mean-based (MEAN) detection scheme [10] the second scheme is the trend-based (TREND) detection scheme [6]; and the third scheme is the victim number variance based (VAR) detection scheme[11]. Our Proposed detection schemes also achieve good DT (detection time) performance in addition to the high DR (detection rate) values as shown in table . In contrast, the DT of existing detection schemes have relatively larger values. Since the DR values for the existing detection schemes are relatively small, obtaining low values of Maximal infected ratio for those schemes are not as significant as those for proposed scheme.

V CONCLUSION

In this paper, we propose to extend Digital Signal Processing (DSP) scheme that uses power spectral density (PSD) distribution or energy spectral density (ESD) distribution with many pre-defined specific energy patterns to detect C-worms and many other smart worms, file sharing worms, internet worms alike. The result is an integrated systems that can detect a variety of computer worms including the smart ones such as C worms. Genuine traffic and C-worms traffic are almost identical and barely noticeable in time-domain. So, the proposed scheme used the *Power Spectral Density* (*PSD*) and Energy spectral density distribution of the C-Worm scan traffic volume and its corresponding *Spectral Flatness Measure* (*SFM*) as the key detection feature to distinguish the C-Worm scan traffic from the normal non-worm scan traffic. The evaluation data showed that our scheme achieved superior detection performance against the C-Worm in comparison with existing worm detection schemes.

VI REFERENCES

[1] D. Moore, C. Shannon, and J. Brown, "Code-Red: A Case Study on the Spread and Victims of an Internet Worm," Proc. Second Internet Measurement Workshop (IMW), Nov. 2002.

[2] R. Vogt, J. Aycock, and M. Jacobson, "Quorum Sensing and Self- Stopping Worms," Proc. Fifth ACM Workshop Recurring Malcode (WORM), Oct. 2007.

[3] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," Proc. 11th USENIX Security Symp. (SECURITY), Aug. 2002.

[4] Z.S. Chen, L.X. Gao, and K. Kwiat, "Modeling the Spread of Active Worms," Proc. IEEE INFOCOM, Mar. 2003.

[5] M. Garetto, W.B. Gong, and D. Towsley, "Modeling Malware Spreading Dynamics," Proc. IEEE INFOCOM, Mar. 2003.

[6] C. Zou, W.B. Gong, D. Towsley, and L.X. Gao, "Monitoring and Early Detection for Internet Worms," Proc. 10th ACM Conf. Computer and Comm. Security (CCS), Oct. 2003.

[7] W. Yu, S. Chellappan, C. Boyer, and D. Xuan, "Peer-to-Peer System-Based Active Worm Attacks: Modeling and Analysis," Proc. IEEE Int'l Conf. Comm. (ICC), May 2005.
[8] D. Ha and H. Ngo, "On the Trade-Off between Speed and Resiliency of Flash Worms and Similar Malcodes," Proc. Fifth ACM Workshop Recurring Malcode (WORM), Oct. 2007.
[9] W. Yu, X. Wang, D. Xuan, and D. Lee, "Effective Detection of Active Worms with Varying Scan Rate," Proc. IEEE Int'l Conf. Security and Privacy in Comm. Networks (SECURECOMM), Aug. 2006.
[10] S. Venkataraman, D. Song, P. Gibbons, and A. Blum,"New streaming algorithms for superspreader detection," in Proceedings of the 12-th IEEE Network andDistributed Systems Security Symposium (NDSS), SanDiego, CA, Febrary 2005.
[11] J. Wu, S. Vangala, and L. X. Gao, "An effective architecture and algorithm for detecting worms with various scan techniques," in Proceedings of the 11-thIEEE Network and Distributed System Security Symposium (NDSS), San Diego, CA, Febrary 2004.