Spectrum Based Detection Scheme on Worm

[1]Susarla Valli Kameswari, [2]A. S. K. Maha Lakshmi

[1,2]Dept. of CSE, Kakinada Institute of Engineering and Technology, Korangi, AP, India

Abstract

There are several worm attacks in the recent years, this leads to an essentiality of producing new detection technique for worm attacks. In this paper we present a range based smart worm detection scheme, this is based on the idea of detection of worm in the frequency domain. This scheme uses the power spectral density of the scan traffic volume and its corresponding flatness Measure to distinguish the smart worm traffic from background traffic. This scheme showed better results against the smart worms and also for the c-worm detection. Motivated by our observations, we design automatic detection of C-worm and specifying scope for the entire network and the scope for the particular region. We are proposed a method that scans the entire network globally and defines the regions to model and detect the C Worm and detect the different worms over the internet and the proposed system automatically applied on the worm detection according to the regions of over the internet without disturbing their real work. The proposed method is optimized method to differentiate the region results over the entire network maintained by the system. Our scheme uses the Power Spectral Density (PSD) distribution of the scan traffic volume and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from background traffic. Using a comprehensive set of detection metrics and real-world traces as background traffic, we conduct extensive performance evaluations on our proposed range -based detection scheme. The performance data clearly demonstrates that our scheme can effectively detect the C-Worm propagation.

Keywords

Networks, Camouflage, Anomaly Detection, C-Worm.

I. Introduction

C- Worm is quite different from traditional worms in which it camouflages any noticeable trends in the number of infected computers over time. The camouflage is achieved by manipulating the scan traffic volume of worm-infected computers. Such a manipulation of the scan traffic volume prevents exhibition of any exponentially increasing trends or even crossing of thresholds that are tracked by existing detection schemes. We note that the propagation controlling nature of the C-Worm (and similar smart worms, such as “Attack”) cause a in the propagation speed. However, by carefully controlling its scan rate, the C-Worm can: (a) still achieve its ultimate goal of infecting as many computers as possible before being detected, and (b) position it to launch subsequent attacks. We comprehensively analyze the propagation model of the CWorm and corresponding scan traffic in both time and frequency domains. We observe that although the C-Worm scan traffic shows no noticeable trends in the time domain, it demonstrates a distinct pattern in the frequency. Specifically, there is an obvious concentration within a narrow range of frequencies. This concentration within a narrow range of frequencies is inevitable since the C-Worm adapts to the dynamics of the Internet in a recurring manner for manipulating and controlling its overall scan traffic volume. We adopt frequency domain analysis techniques and develop a detection scheme against wide-spreading of the C-Worm. Particularly, we develop a novel range -based detection scheme that uses the Power Spectral Density (PSD) distribution of scan traffics volume in the frequency domain and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from non- worm traffic (background traffic). Our frequency domain analysis studies use the real-world Internet traffic traces (Shield logs dataset) provided by SANs Internet Storm Center

(MIR) is the one to quantify the infection damage caused by a worm before being detected. Other metrics include Detection Time (DT) and Detection Rate (DR). Our evaluation data clearly demonstrate that our range based detection scheme achieves much better detection performance against the C-Worm propagation compared with existing detection schemes. Our evaluation also shows that our range -based detection scheme is general enough to be used for effective detection of traditional worms as well.

In this paper Camouflaging worm (c-worm short) is modeled. The C-Worm is different from traditional worms because of its ability to intelligently manipulate its scan traffic volume over time. Thereby, the effects of C-Worm camouflages its propagation from existing worm detection systems based on analysing the propagation traffic generated by worms. Two types of traffic are barely distinguishable in the time domain. However, their distinction is clear in the frequency domain, due to the recurring manipulative nature of the C-Worm. A novel range -based scheme to detect the C-Worm. Scheme uses the Power Spectral Density (PSD) distribution of the scan traffic volume and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from background traffic. Using a comprehensive set of detection metrics and real-world traces as background traffic. The performance data clearly demonstrates that scheme can effectively detect the C-Worm propagation. The generality of our range -based scheme in effectively detecting not only the C-Worm, but traditional worms as well.

II. Related Works

A. Worm Analysing

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. To analyse the C-Worm, we adopt the epidemic dynamic model for disease propagation, which has been extensively used for worm propagation modelling. Particularly, the epidemic dynamic model assumes that any given host is in one of the following states: immune, vulnerable, or infected. An immune host is one that cannot be infected by a worm; a vulnerable host is one that has the potential of being infected by a worm; an infected host is one that has been actually infected by a worm. The simple epidemic model for a finite population of traditional PRS worms can be expressed as, The basic form of active worms 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 [10] and hit list scan worm [9]. Both of these worms attempt to speed up their propagation by increasing the probability of successful scanning.

B. Worm Revealing

The generic worm detection framework that we use in this paper consists of multiple distributed monitors and a worm detection centre that controls the former. The monitors are distributed across the Internet and can be deployed at hosts, router, or firewalls etc. Each monitor passively records irregular port-scan traffic such as connection attempts to a range of invalid IP addresses (IP addresses not being used) and restricted service ports. Periodically, the monitors send traffic logs to the detection centre. The detection centre analyses the traffic logs and determines whether there is suspicious scans to restricted ports or to invalid IP addresses. If such uncommon scans are detected, the detection centre determines that there is wide-spreading worm propagation on the Internet. The worm detection schemes used in the detection centre rely on the analysis of globally collected scan traffic data. Specifically, they study the traffic volume to detect the existence of wide-spreading worms. Some of these schemes use the variance of traffic volume or the exponentially increasing trend of traffic volume to identify Large-scale worm propagations. Besides the above detection schemes that are based on the global scan traffic monitoring, there are other worm detection schemes such as sequential hypothesis testing for detecting worm-infected hosts, DSC (Destination-Source Correlation) for detecting a worm in local networks , content-based worm signature . In contrast, our range -based detection scheme uses frequency domain analytical techniques to capture the wide spreading worm propagation.

C. C-Worm: Modelling of the C-Worm

The C-Worm camouflages its propagation by controlling scan traffic volume during its propagation. The simplest way to manipulate scan traffic volume is to randomly change the number of worm instances conducting port-scans. Worm propagation over the Internet can be considered a dynamic system. When an attacker launches worm propagation, it is vey challenging for the attacker to know the accurate parameters for worm propagation dynamics over the Internet. As more and more computers get infected, they, in turn, take part in scanning other computers. Hence, we consider the C-Worms a worst case attacking scenario that uses a closed loop control for regulating the propagation speed based on the feedback propagation status. In order to effectively evade detection, the overall scan traffic for the C-Worm should be comparatively slow and variant enough to not show any notable increasing trends over time. On the other hand, a very slow propagation of the C-Worms also not desirable, since it delays rapid infection damage to the Internet. Hence, the C-Worm needs to adjust its propagation so that it is neither too fast to be easily detected, nor too slow to delay rapid damage on the Internet. a control parameter called attack probability P(t) for each worm-infected computer. P(t) is the probability that a C-Worm instance participates in the worm propagation (i.e. scans and infects other computers) at time t. Our C-Worm model with the control parameter P(t) is generic. P (t) = 1represents the cases for traditional worms, where all worm instances actively participate in the propagation. For the C-Worm (t) needs not be a constant value and can be set as a time varying function. In order to achieve its camouflaging behaviour, the C-Worm needs to obtain an appropriate P(t) to manipulate its scan traffic. Specifically, the C-Worm will regulate its overall scan traffic volume such that:[a] .There are other approaches to achieve this goal, such as incorporating the Peer-to-Peer techniques to disseminate information through secured IRC channels.

D. Breeding Model of the C-Worm

To analyse the C-Worm, we adopt the epidemic dynamic model for disease propagation, which has been extensively used for worm propagation modelling. Based on existing results, this model matches the dynamics of real worm propagation over the Internet quite well. For this reason, similar to other publications, we adopt this model in our paper as well. Since our investigated C-Worm is a novel attack, we modified the original Epidemic dynamic formula to model the propagation of the C-Worm by introducing the p(t) - the attack probability that a worm-infected computer participates in worm propagation at time t. We note that there is a wide scope to notably improve our modified model in the future to reflect several characteristics that are relevant in real-world practice.

E. Success of the C-Worm

We now demonstrate the effectiveness of the C-Worm in evading worm detection through controlling P(t). Given random selection of Mc, we generate three C-Worm attacks (viz., CWorm1, C-Worm 2 and C-Worm 3) that are characterized by different selections of mean and variance magnitudes for MC. Fig. 1 shows the observed number of worm-infected computers over time for the PRS worm and the above three CWormattacks. Fig. 2 shows the infection ratio for the PRSworm and the above three C-Worm attacks. These simulations.

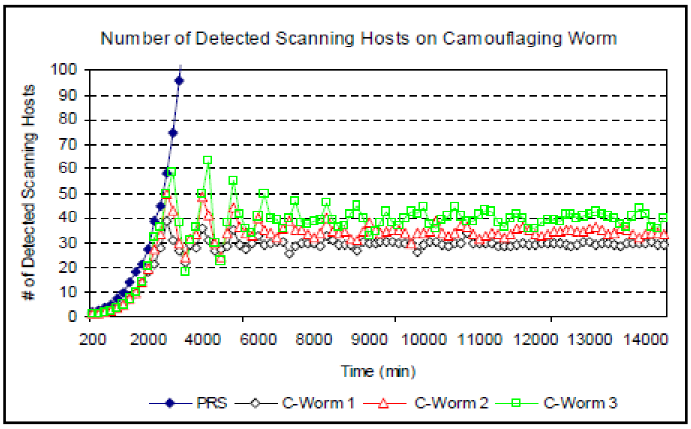

Fig. 1: Observed Infected Instance Number for the C-Worm and PRS Worm

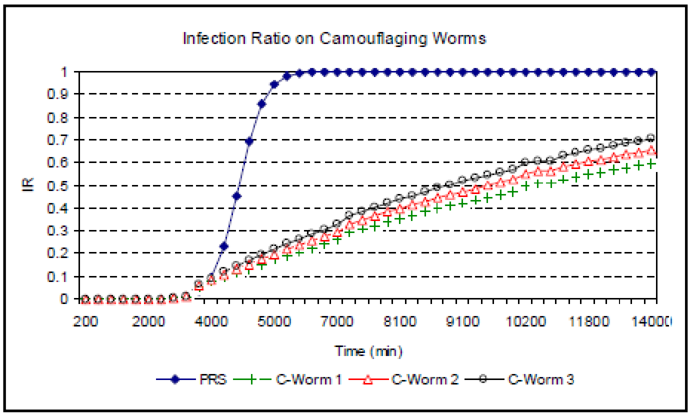

Fig. 2: Infected Ratio for the C-Worm and PRS-Worm

From above figs. 1 and 2, we also observe that the C-Worm is still able to maintain a certain magnitude of scan traffic so as to cause significant infection on the Internet.

III. Proposed System

If worm were found out and cleared user might not know about the source node which sent the worm file. This is major disadvantage in the existing systems. The Worm Behavior is monitored and compared with the Previous Behavior of Worms, so that Traditional Worm Detection Method is adopted to kill the worm from the network. Network Traffic is also monitored so that to identify the Worm presence in the network. Traditional Worms are more threats to the Internet and also would produce lot of Overall Network Traffic. It is very easy to identify the Traditional Worm as it Increases the Overall Traffic of the Network Significantly.

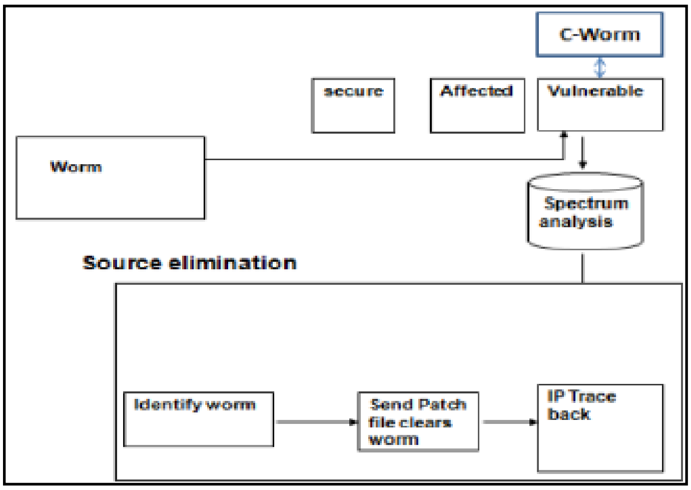

Fig. 3: Source Elimination

The worm infected computer identifies and infect vulnerable computer. This newly infected computer will automatically scan several IP addresses to identify and infect other vulnerable computers. The C-worm is different from traditional worms in which it camouflages any noticeable trends in the number of infected computers. 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. C-Worm rather focusing all the IP, instead it focuses only the Vulnerable Systems, because these systems are the Target of C-Worm. 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. On the other hand, a very slow propagation of the C-Worm is also not desirable, since it delays rapid infection damage to the Internet. Hence, the C-Worm needs to adjust its propagation so that it is neither too fast to be easily detected, nor too slow to delay rapid damage on the Internet. The C-worm and non worm network traffic is need to analyze. In this paper the range based 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.

A. Range Based Analysis

Worm which is the malicious software program that propagate itself on the Internet. It self-replicating computer program which uses a computer network to send copies of itself to other nodes without any intervention. In range based detection, the distribution of PSD and its corresponding SFM are used to distinguish the C-Worm scan traffic from the non worm scan traffic. The C-worm which doesn't show any noticeable trend and detecting c-worm is very difficult. The A range based detection schemes which detect the C-Worm very easily. The Power Spectral Density which shows it work only in time domain but C-Worm can be detected only in frequency domain. Spectral Flatness Measure is the correspond method of Power spectral Density and by using this C-worm and the difference of their traffic level is detected. The central step in devising our source separation algorithm is the choice of a measure describing the complexity of an audio scene. Given such a measure, it is possible to evaluate it for several combinations of input sounds and choose the combination that gives the lowest complexity score. The measure used in the approach of the spectral flatness measure. It measures how much the energy at a given time is spread in the range, giving a high value when the energy is equally distributed and a low value when the energy is concentrated in a small number of narrow frequency bands. The spectral flatness measure is computed from the range as the geometric mean of the Fourier coefficients divided by the arithmetic mean.

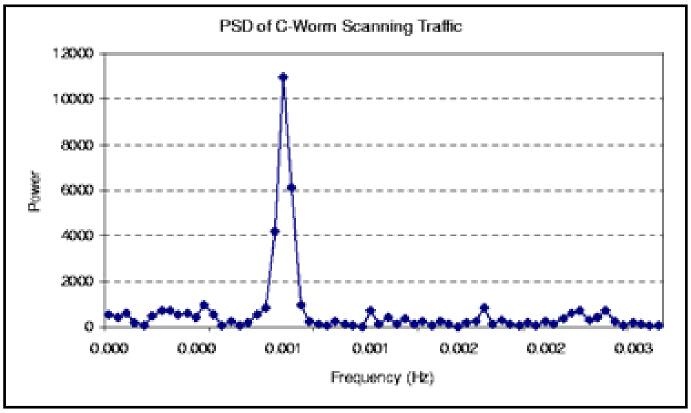

Fig. 4: PDF of C-Worm SFM

B. Power Spectral Density

Power Spectral Density the distribution of worm detection data need to transform from time domain to frequency domain. The C-worm is modeled in such a it increases the CPU usage memory. Using Power spectral Density some time period is added and its correspond method Spectral Flatness Measure which scans the background traffic of C-worm and non worm traffic in that specified time period. PSD describes how the power of time series is distributed I the frequency domain. The SFM of PSD is defined as the ratio of geometric mean to arithmetic mean of the coefficient of PSD. 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. It is often called simply the range of the signal. Intuitively, the spectral density measures the frequency content of a stochastic process and helps identify periodicities. PSD is a very useful tool it identify oscillatory signals in your time series data and want to know their amplitude. For example let assume the operating a factory with many machines and some of them have motors inside. It detects unwanted vibrations from somewhere. It might be able to get a clue to locate offending machines by looking at PSD which would give you frequencies of vibrations. PSD is still useful even if data do not contain any purely oscillatory signals. For example, the sales data from an ice-cream parlor, you can get rough estimate of summer sales peak by looking at PDF of your data. The quite often compute and plot PSD to get a "feel" of data at an early stage of time series analysis.

C. Sub-Region Based Network

In this paper, we focus on a new class of worms, referred to as the camouflaging worm (C-Worm). 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. Recent studies also showed that existing commercial anti-worm detection systems fail to detect brand new worms and can also be easily circumvented by worms that use simple mutation techniques to manipulate their payload.

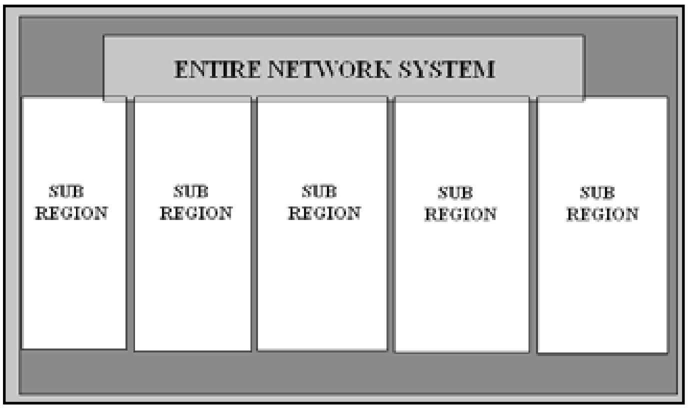

Fig. 5: Sub-Region based Network Architecture

The proposed system works on the worms, the system creates large scans and sub scans of the entire network and sub network for different type of attack over the internet, considering the traffic over the network and the system automatically shows the impact over the internet to reduce the runtime of the worm and eliminates the worms from the internet by applying the above methods. The performance during detection is very high for this method which shows in the result evolution.

In this paper, we are working on the new concept of region based setting of the network to model and detect the worms over the internet and by running our system over the internet according to the Network Rate (NR). The system first incorporates the Power Spectral Density (PSD) distribution of the scan traffic volume and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from background traffic. by considering the PSD and SFM the system automatically detects and assigns distinguishes the effeteness of the worms over the internet. In order to identify the C-Worm propagation in the frequency domain, we use the distribution of Power Spectral Density (PSD) and its corresponding Spectral Flatness Measure (SFM) of the scan traffic. Particularly, PSD describes how the power of a time series is distributed in the frequency domain. Mathematically, it is defined as the Fourier transform of the auto-correlation of a time series. In our case, the time series corresponds to the changes in the number of worm instances that actively conduct scans over time. The SFM of PSD is defined as the ratio of geometric mean to arithmetic mean of the coefficients of PSD. we plot the Probability Density Function (PDF) of SFM for both C-Worm and normal non-worms can traffic as shown in fig. 5 and fig. 6, respectively.

IV. Conclusion

In this paper, we studied a new class of smart-worm called C-Worm, which has the capability to camouflage its propagation and further avoid the detection. Our 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, we developed a novel region based network detection of worms over the entire network. A.range -based detection scheme to detect the C-Worm. Our evaluation data showed that our scheme achieved superior detection performance against the C-Worm in comparison with existing representative detection schemes. This paper lays the foundation for ongoing studies of “smart” worms that intelligently adapt their propagation patterns to reduce the effectiveness of countermeasures.

References

[1] D. Moore, C. Shannon, J. Brown,“Code-red: a case study on the spread and victims of an internet worm”, in Proceedings of the 2-th\ Internet Measurement Workshop (IMW), Marseille, France, November 2002.

[2] D. Moore, V. Paxson, S. Savage,“Inside the slammer worm”, in IEEE Magazine of Security and Privacy, July 2003.

[3] CERT,CERT/CCadvisories, [Online] Available: http://www.cert.org/advisories

[4] P. R. Roberts,"Zotob Arrest Breaks Credit Card FraudRing", [Online] Available: http://www.eweek.com/article2/0,1895,1854162,00.asp.

[5] W32/MyDoom.B Virus, [Online] Available: http://www.us-cert.gov/cas/techalerts/ TA04-028A.html.

[6] W32.Sircam.Worm@mm, [Online] Available: http://www.symantec.com/avcenter/venc/data/w32.sircam.worm@mm.html.

[7] Worm.ExploreZip, [Online] Available: http://www.symantec.com/avcenter/venc/data/worm.explore.zip.html.

[8] R. Naraine, Botnet Hunters Search for Command and Control Servers, [Online] Available: http://www.eweek.com/article2/0,1759,1829347,00.asp.

[9] T. Sanders,"Botnet operation controlled 1.5m PCs Largest zombie army ever created", [Online] Available: http://www.vnunet.com/vnunet/news/2144375/botnet-operation-ruled-million, 2005.

[10] R. Vogt, J. Aycock, M. Jacobson,“Quorum sensing and selfstopping worms”, in Proceedings of 5th ACM Workshop on RecurringMalcode (WORM), Alexandria VA, October 2007.

[11] S. Staniford, V. Paxson, N.Weaver,“How to own the internet in your spare time”, in Proceedings of the 11-th USENIX Security Symposium(SECURITY), San Francisco, CA, August 2002.

[12] Z. S. Chen, L.X. Gao, K. Kwiat,“Modeling the spread of active worms”, in Proceedings of the IEEE Conference on ComputerCommunications (INFOCOM), San Francisco, CA, March 2003.

[13] M. Garetto, W. B. Gong, D. Towsley,“Modeling malware spreading dynamics”, in Proceedings of the IEEE Conference on ComputerCommunications (INFOCOM), San Francisco, CA, March 2003.

[14] C. C. Zou, W. Gong, D. Towsley,“Code-red worm propagation modeling and analysis”, in Proceedings of the 9-th ACM Conferenceon Computer and Communication Security (CCS), Washington DC, November 2002.

[15] Zdnet,"Smart worm lies low to evade detection", [Online] Available: http://news.zdnet.co.uk/internet/security/0,39020375,39160285,00.htm.

[16] J. Ma, G. M. Voelker, S. Savage,“Self-stopping worms”, in Proceedings of the ACM Workshop on Rapid Malcode (WORM), Washington D.C, November 2005.

[17] Min Gyyng Kang, Juan Caballero, Dawn Song, “Distributed evasive scan techniques and countermeasuress”, in Proceedings of InternationalConference on Detection of Intrusions & Malware, and Vulnerability Assessment (DIMVA), Lucerne, Switzerland, July 2007.

[18] Charles Wright, Scott Coull, Fabian Monrose,“Traffic morphing: An efficient defense against statistical traffic analysis”, in Proceedingsof the 15th IEEE Network and Distributed System Security Symposium (NDSS), San Diego, CA, Febrary 2008.

[19] C. Zou, W. B. Gong, D. Towsley, L. X. Gao, “Monitoring and early detection for internet worms”, in Proceedings of the 10-th ACM Conference on Computer and Communication Security (CCS), Washington DC, October 2003.

[20] S. Venkataraman, D. Song, P. Gibbons, A. Blum,“New streaming algorithms for superspreader detection”, in Proceedings of the 12-thIEEE Network and Distributed Systems Security Symposium (NDSS), San Diego, CA, Febrary 2005.

[21] J. Wu, S. Vangala, L. X. Gao,“An effective architecture and algorithm for detecting worms with various scan techniques”, inProceedings of the 11-th IEEE Network and Distributed System Security Symposium (NDSS), San Diego, CA, Febrary 2004.

[22] Dshield.org,"Distributed Intrusion Detection System", [Online] Available: http://www.dshield.org/, 2005.

[23] SANS,"Internet Storm Center", [Online] Available: http://isc.sans.org/.

[24] C. C. Zou, W. Gong, D. Towsley,“Worm propagation modeling and analysis under dynamic quarantine defense”, in Proceedings of the1-th ACM CCS Workshop on Rapid Malcode (WORM), Washington DC, October 2003.

[25] C. C. Zou, D. Towsley, W. Gong,“Modeling and simulation study of the propagation and defense of internet e-mail worm”, IEEETransactions on Dependable and Secure Computing, Vol. 4, No. 2, pp. 105–118, 2007.

[26] C. Zou, Don Towsley, Weibo Gong,“Email worm modelingand defense”, in Proceedings of the 13-th International Conferenceon Computer Communications and Networks (ICCCN), Chicago, IL,October 2004.

[27] W. Yu, S. Chellappan C. Boyer, D. Xuan,“Peer-to-peer systembasedactive worm attacks: Modeling and analysis”, in Proceedings of IEEE International Conference on Communication (ICC), Seoul, Korea, May 2005. 2007.

Ms. A.S.K. Mahalakshmi, an efficient teacher, received M.Tech (CST) from Gitam University – Visakhapatnam, is working as an Assistant Professor in Department of CSE, Kakinada Institute of Engineering and Technology, Korangi. Her areas of Interest include mobile communications, Computer Networks, Database Management Systems, C Programming and Unified modeling language.

Ms. Susarla Valli Kameswari is a student of Kakinada Institute of Engineering & Technology, Korangi. Presently she is pursuing her M.Tech (CSE) from this college and she received her M.Sc(Computer Science) from Andhra University. Her areas of interest includeComputer Networks, Database Management systems and Object oriented Programming languages.