A Novel Mechanism for Modeling and Defensing of Camouflaging Worm

Anand Chalamcharla [#1], D.Sattibabu [*2], Dr.S.Maruthu Perumal [*3]

[#1] IInd M.TECH (CSE) Student, [*2] Associate.Professor, [*3]Professor & HOD
Department of CSE
Godavari Institute of Engineering and Technology (GIET),
Rajahmundry, AP, India.

Abstract

Active Worms wreak havoc by exploiting security loopholes and flaws in software design to propagate from one machine to another. Active Worms are different than a traditional virus in that they don't spread by modifying programs on a single system, but rather by searching for and implanting destructive code onto other systems automatically. A new class of active worms, referred to as Camouflaging Worm (C-Worm in short). 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 characteristics of the C-Worm and conduct a comprehensive comparison between its traffic and non-worm traffic (background traffic). The 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. Motivated by observations, designed a novel spectrum-based scheme to detect the C-Worm. 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 extensive performance evaluations on proposed spectrum-based detection scheme. The performance data clearly demonstrates that our scheme can effectively detect the C-Worm propagation. Furthermore, show the generality of spectrum-based scheme in effectively detecting not only y the C-Worm, but traditional worms as well.

Keywords:

Worm, camouflage, anomaly detection, DDoS, SFM, PSD, C-WORM, Networks.

1. Introduction

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. If a system is affected by worm it is cleared by using antivirus software. But if the operating system of a system gets affected by worm it is impossible to clear it. So we use special technique called spectrum based method to detect the worm and control the worm using Discrete Mathematical model. In most of the existing system, when an operating system is affected by a worm it has to be formatted and a new operating system should be installed. 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 proposed system models the camouflaging worm (C- Worm), in which the behavior is hidden and its action is implicitly kept secret. So this process of detecting the c-worm is not possible using the usual traditional worm detection techniques as well as ip trace back systems. 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, and 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.

Smart worms are malicious software that can self –propagate across the internet, i.e., compromise vulnerable hosts and use them to attack other victims. Since the early stage of the internet, worms have caused enormous damage and been a significant security threat. For example, the Morris worm infected 10% of all hosts in the internet in 1988. The Code Red worm compromised at least 359,000 hosts in one day in 2001 [1], and the Storm botnet affected the tens of millions of hosts in 2007.

Many real-world worms have caused notable damage on the Internet. These worms include ―Code-Red‖ worm in 2001, ―Slammer‖ worm in 2003 [2], and ―Witty‖/ ―Sasser‖ worms in 2004 [3]. Many active worms are used to infect a large number of computers and recruit them as bots or zombies, which are networked together to form botnets [4]. These botnets can be used to:

1. Launch massive Distributed Denial-of-Service (DDoS) attacks that disrupt the Internet utilities [5],

2. Access confidential information that can be misused [6] through large-scale traffic sniffing, key logging, identity theft, etc.,

3. Destroy data that has a high monetary value and

4. Distribute large-scale un solicited advertisement Emails (as spam) or software.

Nevertheless, the attackers are crafting attack strategies that intend to defeat existing worm detection systems. In particular, ―stealth‖ is one attack strategy used by a recently discovered active worm called ―Atak‖ worm [7] and the ―self-stopping‖ worm [8] circumvent detection by hibernating with a predetermined period.

Camouflage worm is modeled and spectrum based approach is used for the detection of C-Worm. The project 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 spectrum-based detection scheme and effectively detecting not only the C-Worm, but traditional worms as well and prevention of C-Worm using mathematical model with particular references of C-Worm.

2. Related Work

Wei Yu, Xun Wang, Prasad Calyam, Dong Xuan, and Wei Zhao [9] proposed a mechanism for detecting C-Worms based on analyzing the propagation traffic generated by worms. They analyzed characteristics of the C-Worm and conducted a comparison between its traffic and non-worm traffic (background traffic). Observations show that 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. They designed a novel spectrum-based scheme to detect the C-Worm. 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.

Zesheng Chen, Lixin Gao, and Chuanyi Ji focuses on effectiveness of defense systems against active worms. It is vital to provide a basic understanding of how efficient the current systems defend against worms, by what way we determine the effectiveness of a defense system, and the guidelines that can be drawn for developing future defense systems. In their work, they have investigated the performance of different host-based defense systems against active worms using a discrete-time (AAWP) model and shown that the ability of worm propagation is constrained by three parameters: number of vulnerable machines, scanning rate, and time to complete infection. Focusing on the Code-Red-v2-like worm, they have performed a quantitative study on how well a system can slow down the propagation of worms.

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,[11] 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. Using a technique that enables global detection of worm spread, we collected and analyzed data over a period of 45 days beginning July 2nd, 2001 to determine the characteristics of the spread of Code-Red throughout the Internet. In this paper, we describe the methodology we use to trace the spread of Code-Red, and then describe the results of our trace analyses. 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. Finally, the experience of the Code- Red worm demonstrates that wide-spread vulnerabilities in Internet hosts can be exploited quickly and dramatically, and that techniques other than host patching are required to mitigate Internet.

A mathematical model derived from empirical data of the spread of Code Red I in July, 2001. We discuss techniques subsequently employed for achieving greater virulence by Code Red II and Nimda. In this context, we develop and evaluate several new, highly virulent possible techniques: hit-list scanning (which creates a Warhol worm), permutation scanning (which enables self-coordinating scanning), and use of nternet sized hit-lists (which creates a flash worm). We then turn to the to the threat of surreptitious worms that spread more slowly but in a much harder to detect "contagion" fashion. We demonstrate that such a worm today could arguably subvert upwards of 10,000,000 Internet hosts. We also consider robust mechanisms by which attackers can control and update deployed worms. In conclusion, we argue for the pressing need to develop a "Center for Disease Control" analog for virus and worm-based threats to national cyber security, and sketch some of the components that would go into such a Center. Slammer worm spread so quickly that human response was ineffective. As it began spreading throughout the Internet, the worm infected more than 90 percent of vulnerable hosts within 10 minutes, causing significant distruption to financial, transportation, and government institutions.

We tracked slammer spreading behavior via network telescope technique a large address, which is monitored for unusual activity. Ideally these address ranges are unused but routed, which eliminates all normal activity. David Moore, Geoffrey volkier and Stefan Savage developed network telescope to understand distributed denial of service attacks, which generate a considerable amount of "backscatter", or response to packets with randomly forged source addresses. Modeling the spread of active worms can help us understand how active worms spread, and how we can monitor and defend against the propagation of worms effectively. In this paper, author present a mathematical model, referred to as the Analytical Active Worm Propagation (AAWP) model,[12] which characterizes the propagation of worms that employ random scanning. We compare our model with the Epidemiological model and Weaver's simulator. Our results show that our model can characterize the spread of worms effectively. Taking the Code Red v2 worm as an example, we give a quantitative analysis for monitoring, detecting and defending against worms. Furthermore, we extend our AAWP model to understand the spread of worms that employ local subnet scanning. To the best of our knowledge, there is no model for the spread of a worm that employs the localized scanning strategy and believe that this is the first attempt on understanding local subnet scanning quantitatively.

2.1. Active Worms

An active worm refers to a malicious software program that propagates itself on the internet to infect other computers. An active worm such as Code Red or the original Morris worm takes advantage of a security hole in a server. It scans through the Internet, looking for machines running that service. Then it tries to break into that service. If successful, it infects the target machine with another copy of itself. Over a period of several hours, it goes from an initial machine to Internet wide contamination. For an active worm to infect a machine, it must first discover that the machine exists. In traditional active worms, each worm instance takes part in spreading worm attack by scanning and infecting other vulnerable hosts in the internet.

2.2 Camouflaging Worm (C-Worm)

Camouflaging worm (C-Worm) is an intricate type of active worm which attempts to remain hidden by sleeping (suspending scans) when it suspects it is under detection. Worms that adopt such smart attack strategies could exhibit overall scan traffic patterns different from those of traditional worms. Since the existing worm detection schemes will not be able to detect such scan traffic patterns, it is very important to understand such smart-worms and develop new countermeasures to defend against them. However, the 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 "Atak") cause a slowdown in the propagation speed. However, by carefully controlling its scan rate, the C-Worm can still achieve its ultimate goal of infecting as many computers as possible before being detected and position itself to launch subsequent attacks.

In this paper, we conduct a systematic study on a new class of such smart-worms denoted as Camouflaging Worm (C-Worm in short). The C-Worm has a self-propagating behavior similar to traditional worms, i.e., it intends to rapidly infect as many vulnerable computers as possible. We employ a Controlled Packet Transmission (CPT) for monitoring the network traffic and thereby detecting the C-Worm.

3. System Architecture

3.1 C-worm

Camouflaging Worm (C Worm). The C-Worm has the capability to intelligently manipulate its scan traffic volume over time, there by camouflaging its propagation from existing worm detection systems. The C-Worm has a self-propagating behavior similar to traditional worms, i.e, it intends to rapidly infect as many vulnerable computers as possible. However, the 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.

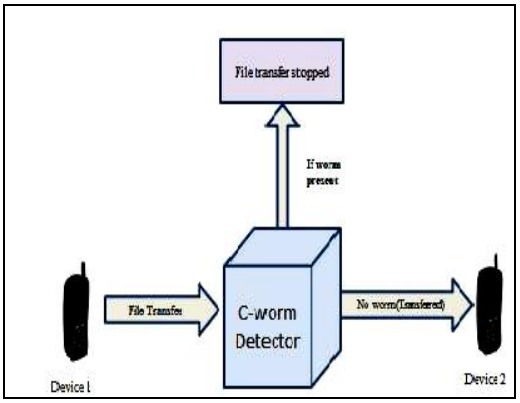

Fig.1 Block diagram of worm detection system

Fig.1 shows the block diagram of worm detection system. It consists of c-worm detector , devices .The device1 will start sending files to device2 ,while transferred c-worm detector is in between them ,which will check the existence of worm , if present stop sending files else send it to other device 2. c worm detector is main system which will wok based on various data mining algorithms of classification. This paper investigates new techniques to detect worms. Initially at first when worm will start affecting, it will find target first in order to attack or spread itself. Once target is find out, it will start propagation by using either any technique of propagation. So device 1 will transfer a file to c-worm detector module. This module will detect the worm, if there is no worm, it will directly pass file to device 2, otherwise if worm is present, it will stop transfer of file.

In this paper new spectrum based detection scheme is introduced. It consists of Centralized data center, Monitor, User. The data center will collect all traffic logs from various network monitors foridentifying worms by their own IP address. The monitors will record all traffic and send it to data center when needed. The data center will collect traffic logs from monitors across internet. The data center then analyzes collected traffic logs and publishes reports to system users.

3.2 Propagation Model of the C-Worm

To analyze the C-Worm, we adopt the epidemic dynamic model for disease propagation, which has been extensively used for worm propagation modeling [10], [13]. Based on existing results [10], [13], this model matches the dynamics of real-worm propagation over the Internet quitewell. 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.

4. Detecting the C-Worm

In most of the existing system, if a system is affected by worm it is cleared by using antivirus software. But if the operating system of a system gets affected by worm it is impossible to clear it. As a result the operating system has to be formatted and a new operating system only should be installed. 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.

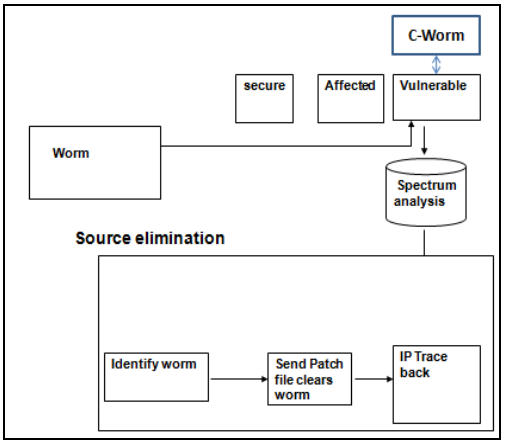

Fig.2 Source Elimination

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 spectrum 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. 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 un occupied 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.

4.1. Spectrum 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 spectrum 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 Spectrum 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 spectrum, 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 spectrum as the geometric mean of the Fourier coefficients divided by the arithmetic mean.

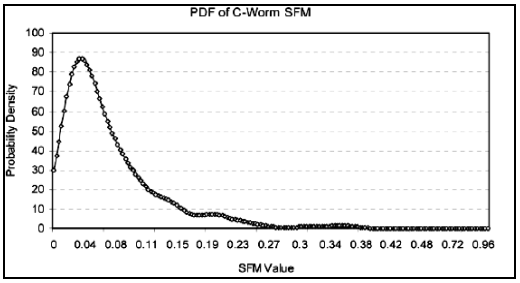

Fig.3 PDF of C-Worm SFM

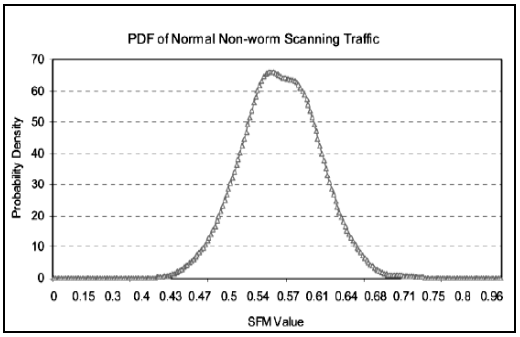

Fig.4 PDF of Normal Non-Worm Scanning Traffic

4.2. 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 spectrum 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 detect 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. Looking at PSD is like looking at simple time series plot except that we look at time series as a function of frequency instead of time. Here, it could say that frequency is a transformation of time and looking at variations in frequency domain is just another way to look at variations of time series data. 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. Much effort has been expended and millions of dollars spent on developing and producing electronic instruments called "spectrum analyzers" for aiding electrical engineers and technicians in observing and measuring the power spectra of signals. The cost of a spectrum analyzer varies depending on its frequency range, its bandwidth (signal processing), and its accuracy. The higher the frequency range (S-band, C-band, X-band, Ku-band, K-band, Ka-band, etc.), the more difficult the components are to make, and the more expensive the spectrum analyzer is. Also, the wider the bandwidth that a spectrum analyzer possesses, the more costly that it is, and the capability for more accurate measurements increases costs as well.

5. Conclusion

In this paper a new class of smart-worm called C-Worm, 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. In this the foundation for ongoing studies of “smart” worms that intelligently adapt their propagation patterns to reduce the effectiveness of countermeasures.

6. References

[1] D. Moore, C. Shanoon, and 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, and S. Savage, ―Inside the Slammer Worm,‖ Proc. IEEE Magazine of Security and Privacy, July 2003.

[3] CERT, CERT/CC Advisories, http://www.cert.org/advisories/, 2010.

[4] P.R. Roberts, Zotob Arrest Breaks Credit Card Fraud Ring, www.eweek.com/article2/0,1895,1854162,00.asp, 2010.

[5]W32/MyDoom.BVirus,http://www.uscert.gov/cas/techalerts/TA04-028A.html, 2010.

[6]W32.Sircam.Worm@mm,http://www.symantec.com/avcenter/venc/data/w32.sircam.worm@mm.html, 2010.

[7] Zdnet, Smart Worm Lies Low to Evade Detection,http://news.zdnet.co.uk/internet/security/0,39020375,39160285,00.htm, 2010.

[8] J.Ma,G.M. Voelker, and S. Savage, ―Self-Stopping Worms,‖Proc.ACM Workshop Rapid Malcode (WORM), Nov. 2005.

[9]. Wei Yu, Xun Wang, Prasad Calyam, Dong Xuan, and Wei Zhao "Modeling and detection of camouflaging worm", IEEE transactions on dependable and secure computing, vol. 8, no. 3, may/june 2011.

[10] A survey of internet worm detection and containment .

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

[12] C.C. Zou, W. Gong, and D. Towsley, "Code-Red Worm Propagation Modeling and Analysis," Proc. Ninth ACM Conf. Computer and Comm. Security (CCS), Nov. 2002.

[13] D. Moore, V.P axsonand S. Savage, Inside the slammer worm,‖ 2003,vol. 1, pp. 33-39.

7. About the Authors

Anand Chalamcharla is currently pursuing his M.Tech (CSE) in Computer Science & Engineering Department, GIET, Rajahmundry. His area of interests includes Networks, Security.

D.Sattibabu is currently working as an Associate Professor in Computer Science & Engineering Department, GIET, Rajahmundry. His research interests include Networks, Data Mining.

Dr. S. Maruthu Perumal is currently working as a Head of the Department for Computer Science & Engineering Department, GIET, Rajahmundry. He is awarded with PhD in related field. His research interests include Image Processing, Data Mining & Warehousing, Networks and Security, Software Engineering.