STAKCERT Worm Relational Model for Worm Detection

Madihah Mohd Saudi, Andrea J Cullen and Mike E Woodward

Abstract— **In this paper, a new STAKCERT worm relational model is being developed based on the evaluation of the STAKCERT worm classification using the dynamic, static and statistical analysis. A case study was conducted to evaluate the effectiveness of this STAKCERT relational model. The case study result analysis showed that the 5 main features in the relational model play an important role in identifying the vulnerability exploited, the damage caused, the expected rate of worm propagation, the chronological flows and the detection avoidance techniques used by the worms. As such, perhaps this new relational model produced can be used as the basis for organizations and end users in detecting worm incidents.**

Index Terms—**dynamic analysis, relational model, static analysis and statistical analysis.**

I. Introduction

Living in a cyber world, email is seen as one of the most important method of communication in our daily life [1]. But since year 2005, there is a new trend on how users communicate with each other. From this year and after, social network sites such as Facebook [2], Myspace,Twitter and Buzz becomes popular and later become very important in users' daily life.

Unfortunately, this kind of media also has been misused and being targeted by worms such as Myspace XSS worm in 2005 [3], followed by Koobface worm in 2008 [4] and XSS exploits in 2009 [5]. These worms have succeeded making their ways due to the vulnerability of these sites. Shepherd [6] explains that this vulnerability is a form of flaw which later been exploited to allow unauthorized access, elevation of privileges and denial of service. These social network sites are using website as its medium of communication and the study and survey conducted by nCircle [7] and IBM [8] shows how worrying these could be. In a survey conducted by IBM in year 2009, shows that by end of 2009, 69% of these websites were still not patched [8]. Therefore these vulnerabilities can be easily exploited by the attacks such as cross site scripting attack or SQL injection. In 2008, based on nCircle survey more than 3000 new web application vulnerabilities were detected and SQL injection errors was the main biggest problem for that year [7], where else cross scripting attack overtook SQL injection in year 2009 [8]. Further discussion on the relationship between vulnerability and email on how the worms infect end user is detailed under section IV of this paper.

Further analysis on the day's vulnerabilities being released and the worms being created is simplified in Figure 1. In Figure 1, if in year 2001, it took about 331 days for the Nimda worm being launched, the numbers of days have now been reduced tremendously in the succeeding years.

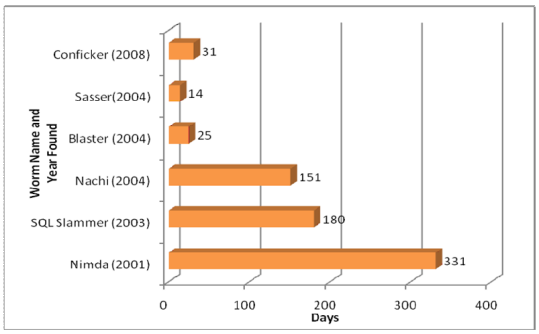

Figure 1. Days between vulnerabilities being released and worms created.

Besides using the vulnerability exploits, removable drive such as USB is another most common way on how the worms can infect victims. In November 2008, Swabey wrote that after the attack of AgentBTZ worm, US Department of Defense had banned the use of the USB unless after it has been scanned and proved free from any worms [9]. Indeed in the same month of the year, three London hospitals' IT systems went down due to Mytob worm attack which caused the administrative process to be done manually [10]. Mytob worm is a mass mailing worm and it spreads via attaching infected file via email and via network shares. Once executed, it opens a backdoor in the victim's machine and blocks access to the security websites [11]. Furthermore in January 2009, 9 million PCs were infected by Downadup worm or also known as Conficker worm [12]. This worm spreads by exploiting Microsoft Windows Server Service vulnerability (MS08-067) in victim's machine [13]. The security incidents listed earlier are only the tip of an iceberg on how the worms spread and its implication thereof. Motivated by these factors, this research has come about.

Initially the researchers for this paper have proposed a framework called STAKCERT framework which stands for Starter Kit for Computer Emergency Response Team to solve the above problems [15]. It is targeted that by the end of the framework, STAKCERT system can be produced to assist

Madihah Mohd Saudi is a lecturer with the Faculty Science and Technology, Islamic Science University of Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. Currently she is a PhD student in School of Computing, Informatics and Media, University of Bradford, United Kingdom. (email: m.b.mohdsa@brad.ac.uk / madihah@usim.edu.my).

Dr Andrea J Cullen is a senior lecturer with the School of Computing, Informatics and Media, University of Bradford, United Kingdom. (email: A.J.Cullen@brad.ac.uk).

Professor Dr. Mike Woodward is a professor with the School of Computing, Informatics and Media, University of Bradford, United Kingdom. (email: M.E.Woodward@brad.ac.uk).

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

user or anyone who is responsible in confronting the worm security incident. This paper is the continuity and explains in detail the testing result part of the procedures in Phase 1 of the framework. The objectives of this paper are to evaluate the relationship between features proposed in the STAKCERT worm classification by using static, dynamic and statistical analysis and to produce new STAKCERT worm relational model. At the same time, the STAKCERT framework proposed earlier being validated and improved to produce a better result. The main reason to evaluate the STAKCERT worm classification relationship is to make sure the features selected are the best features which latter can be used as input for other procedures in the STAKCERT framework.

This research paper consists of the following: Section II contains a discussion on previous work on worm detection and worm relational framework. Section III discusses the methodology used and follows by section IV that explains the findings of the testing result. Section V concludes and discusses the future work for this paper.

II. PREVIOUS WORK

There are many ways on how worm detection and analysis can be conducted. Techniques, tools or software such as Honeynet, Honeymonkey, Sandbox II, static and dynamic analysis are among the most common ways being used for the past few years. Each techniques, tools or software used has its own strengths and drawbacks.

Examples of researches using Honeypot are [16], [17], [18] and [19]. The concept of the Honeypot is to allow the attacker to play around then attack the systems that consists of few machines with different function such as web server and mail server, which are on purposely being left as vulnerable. But it only allows the incoming traffic to the Honeypot and disallows any outgoing traffic from the Honeypot itself. The constraint of the Honeypot lays on its capability to allow the incoming traffic only.

Another example of worm detection techniques is known as HoneyMonkey [20]. It detects and analyses the website that hosting the malicious code. The HoneyMonkey is definitely useful in identifying these malicious websites, but it is not suitable to be implemented in this research. Same goes to firewall and Intrusion Detection System (IDS), which were dedicatedly built to detect the worms' attack.

While Sandbox uses the virtual environment approach to allow worms' replication [21]. The drawbacks of this approach are it is incapable to detect worm in data stream and assuming all worms are executable. Some of the worms would not execute once it identified that it was under emulation mode and if this is the case it will lead to false negative in worm detection.

Ellis presented a framework on a worm relational model that consists of targeting, vulnerability, visibility and infectability [22]. It is a well-structured relational model and the relational model is represented by algebra relational. The improvement that can be made to this relational model is by integrating the avoid detection technique which is being proposed and implemented in STAKCERT worm relational model. In a paper written by Nazario [23], he proposed a network worm framework that consists of 6 components which are reconnaissance capabilities, specific attack capabilities, a command interface, communication capabilities, intelligence capabilities and unused attack capabilities. The framework considers all aspects in confronting the worms but the drawbacks of this framework are that it needs lots of expertise and time consuming as it goes in detail each suggested strategies. However, it is best if we can produce a framework as good as his with less reliance on human and less time consuming.

Based on the gaps found in previous works, this new relational model produced for this research, which is a part of the STAKCERT framework can help to resolve those gaps.

III. METHODOLOGY

As mentioned earlier under section I, the relationship between features selected in STAKCERT worm classification is being evaluated by using the static, dynamic and static analysis as displayed with red highlighted line in Fig. 2.

The main features selected for the relationship testing are infection, propagation, activation, payload and operating algorithm which can be referred in [15]. The lab used for this testing as per illustrated in Fig. 3. It is a controlled lab environment. Almost 80% software used in this testing is open source or free basis. Data is taken from VxHeavens [24]. From 66711 samples taken from the VxHeavens, 5614 were identified as worms. From this figure, 575 which represented 0.86% were identified as the host worm which is the scope for this research. There were 161 variants of worms from the sample taken.

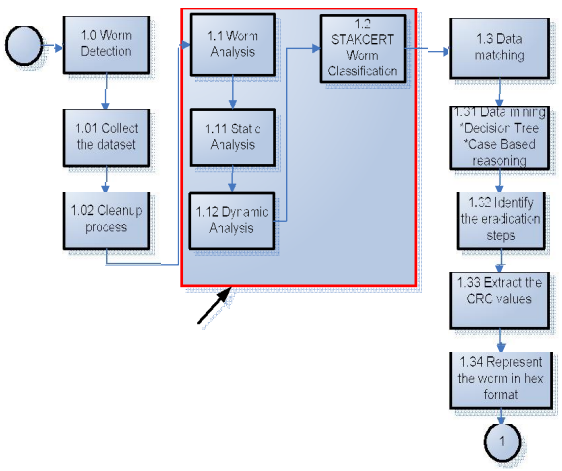

Figure 2. Evaluation phase 1 STAKCERT framework.

The worm analysis techniques used here are the static, dynamic and statistical analysis. Before loading the worm specimen into the machine, the researchers make sure that all preparation and verification have been properly done. When conducting the analysis, all the processes involved were well documented. It is useful in understanding how the worms work so that it can be used for repetitive experiment. The static and dynamic analysis flow is as per illustrated in Fig. 4. Once the static and dynamic analyses were done, the statistical analysis took place. For statistical analysis, the data set have been categorized based on the main features of the STAKCERT classification. The results are recorded based on the worm variant cases. There were 161 variants which represented 161 cases for statistical analysis. The mode for each category being analyzed based on the frequencies. SPSS has been used to conduct this statistical analysis.

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

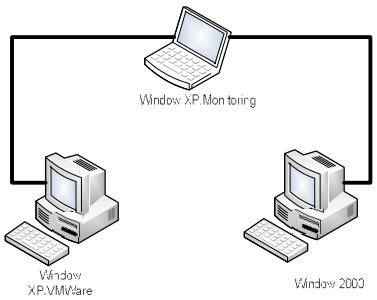

Figure 3. Lab architecture.

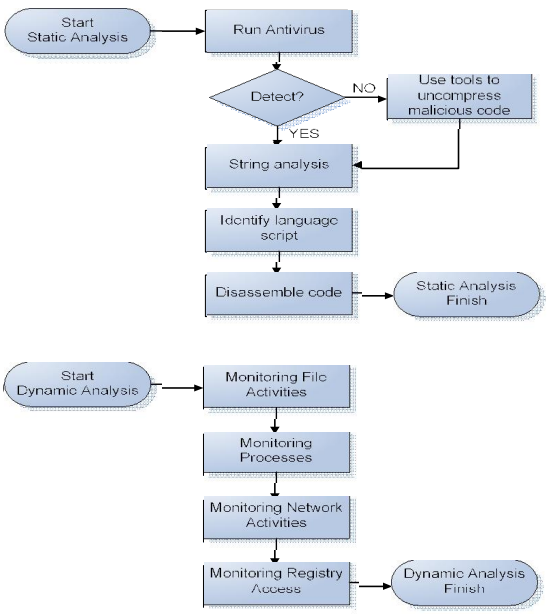

Figure 4. Static and dynamic analysis.

IV. FINDING

A case study was conducted using a sample from VXheavens [24] with the same architecture showed in Fig. 3, which was carried out in a controlled lab environment. A static analysis, dynamic analysis and a statistical analysis had been carried out and the characteristic of the worm were categorized based on the STAKCERT worm classification. Details on how these 3 analyses being conducted can be referred under section III of this paper. The first objective of this case study was to evaluate the infection, activation, payload, operating algorithm and propagation relationship from the worm sample dataset. Based on the testing that had been conducted, a relational model as displayed in Fig. 5 is produced. The statistical analysis results are summarized as follows.

Figure 5. STAKCERT worm relational model.

From statistical analysis, we identified the top 10 ways on how the worms' infect the victim's machine. According to Fig. 6 for infection result analysis, 27.3% is contributed by file and followed by email and vulnerability represented 9.9%. Each of the categories are from the sharing directories, file and sharing directories and file, email and vulnerability represented 8.7%. Only 4.3% represented the vulnerability and 3.1% each for file and vulnerability. Three categories which are the email, chatting channel and sharing directories and vulnerability represented 2.5% each. Others are the combination for different categories in infection.

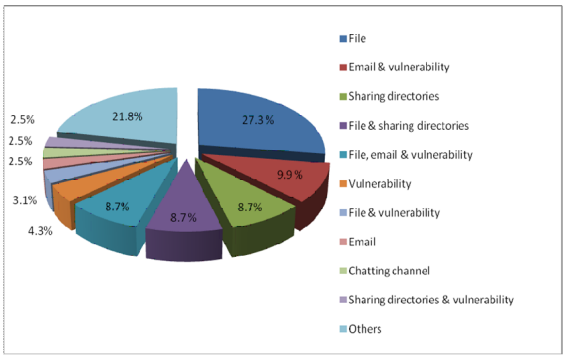

Figure 6. Infection result analysis.

Even though this experiment was conducted with the sample from VxHeavens only, taken up to year 2006, a few interesting association can be made with the current way how the worm's infect. Based on the infection result analysis in Fig. 6, file, email, vulnerability and sharing directories are the most common way how the worms' infect. The top 10 threats for January 2010 by Eset [25] showed that vulnerability, file and email still being used as ways for worms to infect the victim's machine. From this paper, Win32/Conficker worm exploited the vulnerability existed in Windows operating system, while INF/Autorun worm used file autorun.inf to infect and Win32/PSW.OnlineGames worm used phishing attack specifically for game player to steal information related with online game. Bear in mind, phishing can be distributed via email which explains why the email usage can be abused easily. When we make an analysis towards the trend on how the worms spread from year 2001 to 2010, file, vulnerability and email are the most common ways.

Once we conducted this analysis, the relationship between file, vulnerability and email being explored in more depth. There was a scenario where the worms only infect via file, email or vulnerability but not to forget that certain worm use the combination of these 2 or 3 ways to infect victim's machine. Starting from year 1971 until 2010 there are so many methods of worms' infections [26]. Examples of worms exploiting the vulnerability in website or Windows operating systems are such as Code Red worm (2001), Nimda worm (2001) and Conficker worm (2008). But still we cannot simply ignore the other way of the worm's infection. Chatting channel, social network website, removable drives such as USB, P2P (peer-to-peer) and smart phone are among the other alternative way and getting popular nowadays. Worm_Autorun.AZ is an example of worm that spread via chatting channel, P2P network and removable drives.

As for the propagation result analysis, only 10% was through random scanning followed by 3% by sequence scanning. The rest had no scanning implication. This result analysis can be referred in Fig. 7. Once the worm has infected the victim's machine, it needs to spread itself to another

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

machine or network. Based on the testing result with the dataset, more than 50% of the worm did not propagate itself.

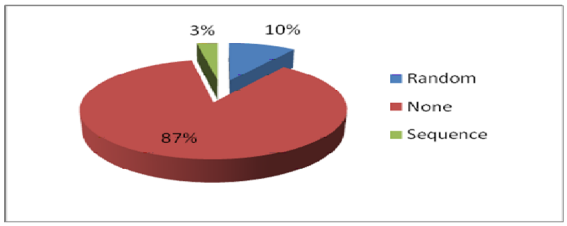

Figure 7. Propagation result analysis.

The question that should be raised here is, should propagation be highlighted as one of the important component in classifying worm? Even though random and sequence propagation represents only 10% and 3%, we cannot underestimate these two ways of propagation. Worms such as Code Red, Nimda, Blaster, Nachi and Sobig.F have their own propagation rate [27]. Moreover, based on the researchers' analysis and experience, propagation is seen as one of the most important element in detecting worms attack.

For the activation result analysis according to Figure 8, more than half was activated via self activation which represents 54.8%. It was followed by the combination of self activation and human trigger which represents 21.7% and 18% accordingly. No activation is represented by 3.7% and the rest presented 0.6% each. Self activation means worm can spread itself to other machine without human intervention such as the Conficker worm where exploits Microsoft vulnerability. The human trigger is caused by few factors such as social engineering technique and logging to certain websites or downloading certain files which leads to file or script execution or opening certain port on the victim's machine.

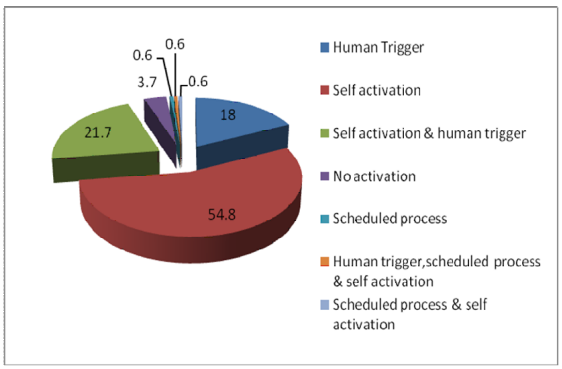

Figure 8. Activation result analysis.

Fig. 9 displayed the first top 10 type of payload. 9.3% represents the destructive implication, followed by performance degradation with 9.3%, autorun registry with 5% and the combination of backdoor and autorun registry with 1.9%. The rest, which are backdoor, infect PE executable, the combination of backdoor and drives infection, the combination of the autorun registry and the creation of infected .exe, and the combination of autorun registry, drive infection and the creation of infected .exe represents 1.2% each. Other payloads that not being discussed here are mostly based on the combination of different payloads. Since our target toward the end of this research is to produce an incident response tool, payload is seen as one of the important element where user can used as the input for this tool. There are so many payload identified based on the testing that had been conducted. In our proposed tool, we decided to use the STAKCERT worm classification as the basis and therefore it is important for us to ensure each component being validated well. It is very interesting to know all the features selected are related with each other based on our static, dynamic and statistical analysis.

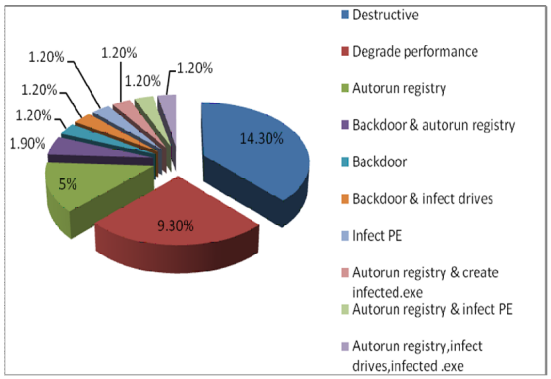

Figure 9. Top 10 payload result analysis.

Last but not least is the operating algorithm. It is the technique used by the worm to avoid detection. Operating algorithm is considered as added feature that should be taken into account when building up the incident response tool. Albanese[14] classified operating algorithm as survival. From the testing conducted, a majority of 96% was categorized as terminate and stay resident (TSR) as displayed in Fig. 10. Stealth represented by 2%, followed by polymorphic and anti anti-virus with 1% each. Each of the operating algorithms has it own ways to keep on spreading and replicating to other machine. Many researchers in worm field are focusing on the polymorphic worm but how about the other techniques? What if in the nearest time the worm uses the combination of polymorphic, stealth, TSR and anti- anti-virus to conceal itself? If a good understanding how each of this technique works being studied, it is not impossible to produce defensive method if the combination techniques being implemented.

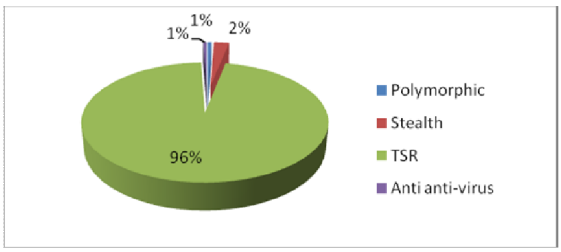

Figure 10. Operating algorithm result analysis.

V. CONCLUSION AND FUTURE WORK

Based on the analysis and testing conducted to the 5 main features of the STAKCERT classification, it can be concluded that each of the features is related with each other. The formation of the STAKCERT relational model is based on the fact that each feature plays an important role for the worm detection and isolation and supported the relevance of the current issues related with worm. This paper is part of a larger research project to tackle the worm detection and isolation issues. Ongoing research includes using the STAKCERT

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

relational model as the input for worm data matching as displayed in Fig. 10 and worm isolation phase which are part of the STAKCERT system formation.

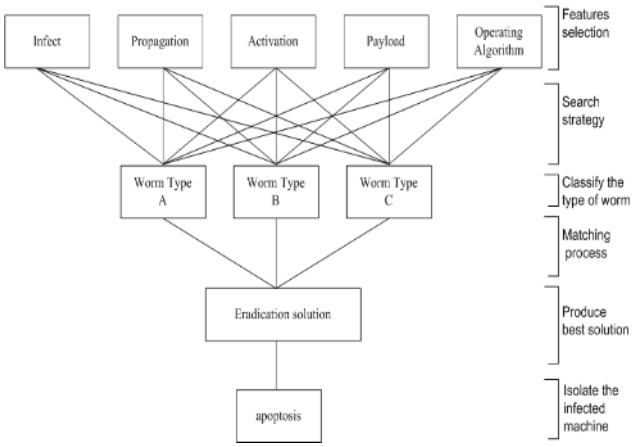

Figure 10. Future work for STAKCERT data matching.

ACKNOWLEDGMENT

The authors would like to express their gratitude to School of Computing, Informatics and Media, University of Bradford and Universiti Sains Islam Malaysia (USIM) for the support and facilities provided.

REFERENCES

[1] Gardner, H. "The Role of Email in Your Communications Mix", July 2009, Available: http: //www.idealware.org/ articles/ email_comm_mix.php

[2] Boyd, D. "Why Youth Social Network Sites: The Role of Networked Publics in Teenage Social Life", *MacArthur Foundation Series on Digital Learning – Youth, Identity, and Digital,* Cambridge, MA: The MIT Press, 2008. pp. 119–142.

[3] Laborger, P. "XSS worm hits Myspace", 19th October 2005, Source: SecurityFocus, Available: http://www.securityfocus.com/brief/18

[4] Vamosi. R. "Koobface virus hits Facebook", 4th December 2008,Source: CNET news: Security, Available: http://news.cnet.com/koobface-virus-hits-facebook/

[5] Robert, A. "Twitter response to XSS worm attack", 4th December 2009, Source: CGISecurity.com, Available: http://www.cgisecurity.com/2009/04/twitter-response-to-xss-worm-attack.html

[6] Shepherd, S. A. "Vulnerability disclosures", 22nd April 2003,Source:Sans Institute Infosec reading room, Available:http://www.sans.org/reading_room/whitepapers/threats/how_do_we_define_responsible_disclosure_932?show=932.php&cat=threats

[7] Westervelt, R. "nCircle statistics show rising Web application vulnerabilities", 2nd July 2009, Source: Search Security.co.UK: Information Security News, Available:http://searchsecurity.techtarget.com/news/article/0,289142,sid14_gci1360879,00.html

[8] Kirk, J. "IBM: Vulnerabilities fell in 2009, but other risks abound", 26th February 2010, Source: Fairfax Media Business Group: Reseller News,Available:http://reseller.co.nz/reseller.nsf/inews/4225F26284474661CC2576D50073ADF9

[9] Swabey,P. "US Department of Defense bans USB drives after worm attack", 20th November 2008, Source: Information Age Today, Available:http://www.information-age.com/home/information-age-today/814827/us-department-of-defense-bans-usb-drives-after-worm-attack.thtml

[10] Swabey,P. "Virus takes down three hospitals", 19th November 2008, Source: Information Age Today, Available: http://www.information-age.com/home/information-age-today/814312/virus-takes-down-three-hospitals-it-systems.thtml

[11] Trend Micro, Inc. "Worm_Mytob.MX", Available:http://threatinfo.trendmicro.com/vinfo/virusencyclo/default5.asp?vname=WORM_MYTOB.MX

[12] Keizer, G. "Amazing' worm attack infects 9 million PCs", 19th January 2009, Source: Computerworld Security, Available: http://www.computerworld.com/s/article/9126205/_Amazing_worm_attack_infects_9_million_PCs

[13] McAfee, Inc. "W32/Conficker.worm",2009, Available: http://vil.mcafeesecurity.com/vil/content/v_153464.htm

[14] Kevin Mandia, Chris Prosise and Matt Pepe. "Incident Response & Computer Forensics", 2nd Edition. McGraw-Hill, 2003.

[15] Madihah Mohd Saudi, Andrea J. Cullen, Mike E. Woodward, "STAKCERT Framework in Eradicating Worms Attack," *In Proceedings International Conference on CyberWorlds 2009.* pp. 257-264,2009.

[16] Levin, J., Labella,R.,. Owen, H., Contis, D. And Culver, B."The Use of Honeypots to Detect Exploited Systems Access Across Large Enterprise Networks". *In Proceedings of the 2003 IEEE Workshop on Information Assurance*, pp.92-99, 2003. Available:10.1109/SMCSIA.2003.1232406

[17] Dagon,D., Qin,X.,Gu,G., Lee,W., Grizzard,J.,Levine,J., and Owen,H. "Honeystat: Local Worm Detection Using Honeypots", *Recent Advances In Intrusion Detection,Vol. 3224/2004, Springer Berlin / Heidelberg, pp.* 39 -58., 2004. Available:10.1007/B100714

[18] Sadasivam, K., Samudrala, B., And Yang,T.A. "Design Of Network Security Projects Using Honeypots", *Journal Of Computing In Small Colleges,*, 20(4), pp. 282-293, 2005. Available: http://Portal.Acm.Org/Citation.Cfm?Id=1047846.1047890

[19] Spitzner, L. "Honeypots: catching the insider threat", *In Proceedings 19th Annual Computer Security Applications Conference*, pp. 170-179, 2003. Available: 10.1109/CSAC.2003.1254322

[20] Wang, Y.-M., Beck, D., Jiang, X., and Roussev, R."Automated Web Patrol with Strider HoneyMonkeys: Finding Web Sites that Exploit Browser Vulnerabilities", Microsoft Research, Technical Report MSR-TR-2005-72, 2005.

[21] K. Natvig. "Sandbox II: Internet". *In Proceedings of the 2002 Virus Bulletin Conference*, pp. 1–18. 2002.

[22] Dan Ellis. "Worm Anatomy and Model". *In Proceedings of the 2003 ACM Workshop*, pp. 42-50, October 2003. Available: http://mason.gmu.edu/~esibley/INFS697F03/ACM%20Worm%20Anatomy%20and%20Model.pdf

[23] J. Nazario, "Defense and Detection Strategies against Internet Worms" (BOOK), *Artech House Inc.*, 2003.

[24] VXHeavens website, "Virus Collection", 2006, Available:http://vx.netlux.org/vl.php.

[25] Eset Company. "Globat Threats Trend", 2010, Available: http://www.eset.ie/threat-center/case_study/GlobalThreatTrendsJanuary2010.pdf

[26] Trend Micro, Inc. "Top 20 Viruses in history", 2008, Available: http://us.trendmicro.com/imperia/md/content/us/pdf/aboutus/20thanniversary/top20_viruses.pdf

[27] Saudi,M.M, "Combating Worms Outbreaks: Malaysia Experience", *International Journal of Learning*, Volume 12, Issue 2, 2005, pp.295-304.