Raspberry Pi-Based Investigating Model for Identifying Intrusion Evidence

Abdulghani Ali Ahmed*, Yee Wai Kit and Amer A Sallam
University Malaysia Pahang, Malaysia
Submission: January 25, 2018; Published: February 02, 2018

*Corresponding author: Abdulghani Ali Ahmed, University Malaysia Pahang, Malaysia, Email: abdulghani@ump.edu.my

Abstract

Nowadays it is very important to maintain an intermediary level of security to ensure safe and trusted communication of information for daily usage. However, a secured data communication over internet and any other network is hard to achieve due to the threat of intrusions and misuses. Malicious traffic exploits the loophole and thus, invades and further sabotages the entire network. In order to combat this issue, users are entrusted with the network security software for safeguarding their own system against unauthorized users. Unfortunately, none of the existing systems proved to be flawless, though various approaches being utilized for thwarting network intrusion activities. The aim for this research is to identify and review the given loophole within network security in contemplate of pinpointing the common network intrusion behavior. This study then proposes an investigation model for collecting network intrusion evidence.

Keywords: Network Intrusions; Honey bot System; Cloud Computing; Forensic Investigation; Raspberry Pi

Introduction

Network vulnerability has always been an issue, given a new breath over the times. Its presence possesses far more potential threat that it seems, unbeknown to household users, or even an avid web browser. However, lack of understanding regarding network intrusion behavior constitutes to unauthorized data mining, for unknown intention. In return, making it an unruly plague, which continually rotten and likewise, infiltrate the integrity/confidentiality of user information, circulating the Internet? In conjunction with increasing number of Internet users worldwide, exposure to network intrusion activities also incremented proportionally. As community become more dependent on the online environment, the emergence of a massive growth in malware activities all across the globe can be foreseen. Nevertheless, there are numerous ways to access and exploit vulnerable systems [1], and yet it remains nonchalant among users. Thus, awareness is needed as to arouse cautiousness in recognizing malware infection in order to prevent and defend personal systems from malicious software.

On the other hand, scheduled system scanning is crucial for identifying weaknesses in network security for an organization [2-5]. This halts progression from unknown users before an attack is initiated. The aim of running a vulnerability scanner or conducting an external vulnerability assessment is to diagnose devices exists on user or corporate network, which is exposed to known vulnerabilities without compromising the systems operation as a result. Though performing a vulnerability scan is an excellent start, the real dilemma emerges from the aspect on how the users should implement recovery action once network breach is detected. This is where risk-based management comes in aid. It testifies the ability of system to counteract threat imposed in real scenario. It serves as a way in helping community by discovery and mitigates any weaknesses on network before they can be exploited. Although several researches and solutions are proposed [6], identifying and collecting evidence of network intrusions still have several challenges.

Handling large amount of data from all connected devices in a network is difficult. Some of the traffic, especially the malicious one will exploit the loophole and thus, invades and further sabotages the entire network. Lack of awareness among users about the importance of preventive measure on personal devices makes network system vulnerable against unauthorized access. Moreover, using only firewall for network protection system is not enough to refrain from the unauthorized access. Network intrusion can be happening at any time without being noticed by users. Thus, their system will be defenseless against those malicious threats if the network security approach is implemented at scheduled time only [7]. However, it consumes
system resources when running network security software for long duration. This research is a proponent for an investigation model to identify and collect possible evidence of intrusions in computer networks. This research focuses on analyzing the behavior of network intrusion activities through experiment, for the means of pinpointing the network security breach to achieve system fortification against unauthorized access.

This research reviews the existing loop hole within network security, albeit household and organization for identifying network intrusion behavior as first objective. Another objective for this research is to investigate and seek suitable methodology to be imposed for improving existing network security software in malicious network. To implement this proponent, an experiment which involves the use of rule-based approach and Raspberry Pi model as a honey pot system to aid in data collection will be conducted. Rule-based statement instilled is executed for identifying malicious activities. In response, those activities will be logged into database. From the collected information, users will observe the trend and activity rates for given intrusion activities by viewing the statistical report. Defense approach can be made before attack is initiated. The rest of paper is organized as follows. Section 2 reviews related works. Sections 3 provide methodologies of intrusion detection. A comparison between the detection methodologies is summarized in section 4. Architecture for the proposed model is conducted in section 5. Finally is to conclude this paper review in Section 6.

Related Work

Massive growth of the Internet offers overall improvement for data coordination and transmission, especially the accessibility to enormous valuable data storage. However, this phenomenon has indeed exposed users to numerous vulnerabilities too in accordance to this instantaneous network expansion, which provokes network safety issue among users. They could be compromised and fall into malevolent scam without even realizing by themselves. As for hackers, their intention is to get financial benefits through their nasty plot from large pool of compromised hosts. This horrifying threat is worsen with the appearance of botnet, which propagates in the manner of Internet worms, remain hidden within the victim system and launches attack after receiving command from master system. A few researches like [8] have been conducted on the methodological analysis about the bot and botnet such as their behaviors, statistics, and traffic measurements. Studies conducted by Hyunsang, [9,10] address the limitation for given signatures [13]. Last but not least, action/reflex against detected threat. Figure 1 shows the main functions of intrusion detection system.

Intrusion Detection Architectural Model

Basically, this is the overview of components within the vicinity of an intrusion detection system environment. Starting from data collection, followed by feature selection and analysis of given signatures [13]. Last but not least, action/reflex against detected threat. Figure 1 shows the main functions of intrusion detection system.

Data collection: This module acts as the initial phase for intrusion detection system. It captures and passes relative data from the monitored system to the neighboring module for further operation, which is done automatically. The collected data is mostly sent to a designated file before being analyzed.

Feature Selection: Next, users’ action is dynamically monitored once they logged into the system [14]. The necessity of this phase is to sort out the distinct feature of large data captured from the network. This helps in evaluation for intrusion activities. For example, the Internet Protocol (IP) address of the source and target system, protocol type, header length and size could be taken as a key for intrusion [15]. Thus, users must deploy set of rules for governing the alerts to reduce associate false positive and false negative response. Figure 2 illustrates the general process of anomaly features selection.
How to cite this article: Abdulghani Ali Ahmed, Yee Wai Kit, Amer A Sallam. Raspberry Pi-Based Investigating Model for Identifying Intrusion Evidence. J Forensic Sci & Criminal Invest. 2018; 7(3): 555715. DOI: 10.19080/JFSCI.2018.07.555715

Analysis phase: From here, the collected data is analyzed to find the suspicious threat. Profiling and pattern recognition techniques also have been used to analyze the data collected and presented to an intrusion detection system [16]. In response, those captured activities which significantly deviate from the applied rules, are referred as anomalous behavior and be flagged as potential intrusions.

Action/Reflex against Threat: This is where the intrusion detection system responsive mechanism implied. It can be resolved by two ways, either sending alert to system administrator with data evidence or directly impose action against detected threat. For example, selectively drop packets to prevent system penetration or close the targeted vulnerable port.

Methodology of Intrusion Detection System

Due to the immense network vulnerability, intrusion detection systems (IDS) have become an asset in securing data integrity and confidentiality of an information system. They are designated to monitor, analyze, and respond to certain security violations against computer and network systems in real time event [17]. These violations result from break in attempts by unauthorized intruders, either remotely or internally. Obviously, they intend to compromise the system for personal gain. This can be also misconduct from internal privileged users that are misusing their authority. Regardless of the evolution in intrusion detection field by days, the underlying methodologies should not be treated lightly as they hold the keys in maintaining and improving efficiency of an intrusion detection system.

In fact, the use of suitable methodologies within intrusion detection system for handling different situation must be practiced in order to achieve optimal performance. These are the three major intrusion detection system methodologies being applied currently: Anomaly Based Methodology, Signature Based Methodology and Stateful Protocol Analysis Based Methodology.

Anomaly Based Methodology

This methodology works by having comparison between probed activities against a baseline profile. The baseline profile is developed during the learning period where the intrusion detection system learns about the environment and creates a normal profile of the monitored system, which can be networks, users and other systems as well. The profile can be fixed or dynamic. Zero-day attacks to environment can be detected...
without any updates to the system. Figure 3 illustrates the general diagram of Anomaly-based protocol. Which are anomaly detection, knowledge/data-mining and machine learning based? The statistical anomaly techniques are meant to build the two required profiles, one during the learning phase which is then used as the baseline profile and the current profile which is compared to the baseline profile and any differences that found a marked as anomalies depending on the threshold settings of the monitored environment [18].

**Figure 3: Anomaly based methodology architecture.**

As for the knowledge/data-mining technique, it automates the process involving searches for anomalies. However, it causes high overhead on the system and degrades overall system efficiency. Last but not least, machine learning technique, which works by analyzing the system, calls for both normal system behavior and suspicious activities. It is given tasks to audit records used in order to determine the feature definitions for generating intrusion detection rules.

**Signature-Based Methodology**

Signature-based methodology works by comparing observed signatures to the signatures stored on database or a list of known attack signature. It works in a similar way to that of a virus scanner. Any signature observed on the monitored environment that matches the signatures on file is deemed as a violation of the security policy or as an attack. Its implementation involves less overhead on the system. This is due to the fact that it does not inspect every single activity or network traffic on the monitored environment. In return, it only requires searching for known signatures stored in the database or file. Compared to the anomaly based methodology, the signature based methodology system is easy to deploy since it does not require prior learning of the working environment. This methodology works by circulating around the process of searching, inspecting and comparing the contents of captured network packets for known threats signatures. Figure 4 illustrates the general diagram of Signature-based protocol.

**Figure 4: Signature based methodology architecture.**

**How to cite this article:** Abdulghani Ali Ahmed, Yee Wai Kit, Amer A Sallam. Raspberry Pi-Based Investigating Model for Identifying Intrusion Evidence. J Forensic Sci & Criminal Invest. 2018; 7(3): 555715. DOI: 10.19080/JFSCI.2018.07.555715
Signature based methodology is effective when it comes to known attacks/violations [19]. However, it cannot detect new attacks until it is updated with new signatures, which is time-consuming. Moreover, it can be easily evaded since they are based on known attacks and are dependent on new signatures to be added before new attacks can be found.

**Stateful Protocol Analysis Based Methodology**

The Stateful protocol analysis methodology works by comparing predefined profiles of how protocols should behave against the observed behavior to identify deviation. Figure 5 illustrates the general diagram of Stateful protocol. Vendors are responsible for designing and establishing the protocol profiles. It explores and has deep understanding on the interactive behavior between the protocols and applications. It differs from the signature based methodology which only compares observed behavior based on given list. On the contrary, this approach of understanding/analysis depicts high overhead on the systems and further degrades its performance. Figure 5 illustrates the stateful protocol analysis based methodology architecture.

Nevertheless, its attributes in analysis for deep understanding of how protocol should behave serves as a base for developing an intrusion detection system which will understand web traffic behavior. Thus, it is more effective at websites protection manner. Though it has deep understanding on monitored protocols, however, it still can be easily evaded by attacks that follow and stay within the acceptable behavior of protocols. It operates based on protocol standard from software vendors and renowned standard bodies. For example, Internet Engineering Task Force (IETF) and Request for Comments (RFC). Variance do apply in each protocol’s implementation. So, the protocol models also typically take this factor into account during its implementation. It does arouse confusion as many standards do not have completion in explaining the details of the protocol. As a result, it causes variations among implementations for the same standard used.

**Methodologies Comparison**

As shown in Table 1, the given three methodologies are evaluated based on the following criteria, which are high accuracy rate in intrusion detection, overall performance against security threat, protection against new attacks, overhead on monitored system and also their scheduled maintenance.

| Attribute                  | Anomaly based | Signature- based | Stateful Protocol Analysis |
|----------------------------|---------------|------------------|----------------------------|
| High Accuracy Rate         | Medium        | Medium           | Medium                     |
| Performance                | Medium        | High             | High                       |
| Protection against new attacks | High         | Low              | Medium                     |
| Overhead on Monitored System | Medium      | Low              | Low                        |
| Maintenance                | Low           | Medium           | Medium                     |

**High accuracy rate:** This feature is essential for detecting and analyzing possible threats against monitored system. Though the rating for all methodologies falls under the same category, however, anomaly-based methodology outweighs the others as it can detect previously known threats. Unlike signature-based methodology, this refers to known threat signature only while stateful protocol analysis is based on user-defined rule list. Both of the methodologies have limitation when new threat surfaces.

**Performance:** It is crucial for performing at peak performance under any circumstances without leading to bottleneck or reducing system efficiency. Overall, the signature and Stateful protocol analysis based methodologies offers better performance than anomaly-based methodologies since they only check for well-defined signatures by vendors which require minimal resource usage. As for anomaly-based methodology, it involves tedious data-mining process in order to identify and categories foreseen events correctly.
Protection against new attacks: For anomaly-based methodology, it does detect new attacks without any updates by referring to both fixed and dynamic profile established. Unlike the signature-based and stateful protocol analysis, which require their signatures database to be updated before they can detect previously unknown threats. The procedure may take up to one week, referring to the complexity/severity of the threat, where the system is already being infiltrated by that time without knowing the origins of threat.

Overhead on Monitored System: Due to its complicated task and vast area covered for intrusion detection phase, the anomaly-based methodology places the most overhead on the targeted system, followed by signature-based and stateful protocol analysis. Least overhead is exerted on the system for both signature-based and stateful protocol analysis by the fact that less resource is consumed for the operation of methodologies in carrying out their pivotal roles of handling intrusion detection issue.

Maintenance: Requirement for maintenance imposed to anomaly-based methodology is less than others. It does not involve updates to initiate identification against new threats. However, the other two methodologies require constant signature updates as to keep track of new security defects. Additional constant update of signatures to resource is required in order to maintain the flow of methodology. As a whole, each methodology has their own edge compared to others, as well as shortcomings which limit their efficiency. For dealing with real-time threat, appropriate use of suitable methodology is the best way in counteracting security threat where the system is exposed to. In conjunction, it fortifies the monitored system as well as optimizing performance for given methodology in reducing unauthorized intrusion.

System Architecture
Design is the gist for turning all requirements into detailed specifications that covers all the aspect of the system. This testifies the feasibility of the proposed system against malicious activities. The selection of suitable intrusion detection mechanism (anomaly-based methodology, signature-based methodology and stateful protocol analysis based methodology) is taken into account as well. The chosen methodology will affect the accuracy and performance of the proposed intrusion detection system used for this research. Figure 6 is considered as blueprint on how the system will be architected and constructed. It also encompasses for the steps needed to construct the proposed system, such as the installation of related software, graphical displaying tools for report viewing and a secure database module for safeguarding the generated log files. In total, it accounted for six phases, which is data collection phase, detection phase, investigation phase, reporting phase, evidence collection phase and also maintenance phase. For data collection phase, Raspberry Pi model is used as honey pot system in capturing network traffic. It acts as victim in real scenario.

For False positive rate (FP), it is calculated as the ratio between the numbers of normal connections that are incorrectly classified as intrusions and the total number of normal connections. Then, it proceeds with the detection phase where captured network is filtered with rule-based statement implemented in the system. For example, comparing the source and destination IP address as well as ports to the rules defined. If it matches, it will notify the users about the intrusion activities. At the same time, it also logs those activities into database, which triggers the investigation phase for observing network intrusion behavior later, such as activity rates for the given intrusion activities. As for the evidence collection phase, it works concurrently with the maintenance phase. The evidence collection phase involves storing of log files for malicious activities inside database, while the maintenance phase is for instilling new rules in detecting new threat based on the collected log files.

Another experiments will be conducted to further deduct whether the log activities is malicious or not, before new rules is instilled to increase the efficiency of the rule-based system. The log files can be viewed in graphical format such as bar chart. This is where the reporting phase takes place. From the statistical
report, users may be able to foresee and predict the upcoming attack before it actually occurred. Preventive measure can be imposed before intrusion attacks being initiated. The research explores the use of genetic algorithm in detecting phase for malicious activities. The algorithm imitates biological evolution as a strategy for problem-solving. It is based on Darwinian’s principle of evolution and survival of fittest to optimize a population of candidate solutions towards a predefined fitness. It is used to predict intrusion types detected with network audit data (logged files) as input. In addition, the algorithm also includes detection rate (DR) and false positive rate as factors in calculating output (activity rate for intrusion events).

Detection rate (DR) is calculated as the ratio between the number of correctly detected intrusions and the total number of intrusions exerted. The algorithm is envisioned to aid the research in the sense of identifying any anomaly activities from captured traffic for investigation purpose. The flowchart shown in Figure 7 depicts on the general operation for the proposed system which aids in the study of network intrusion behavior and also security patterns/events. This gives an overview about the gist of research study and better understanding about the implementation of the proposed system. This diagram is a workflow of stepwise actions. Firstly, user will access the system and start running SNORT software by powering up the Raspberry Pi, which is used as the honey pot for surveillance purpose. Next, SNORT will monitor for any activities that deem harmful or suspicious in the network. The captured network traffic is then compared with the rule-based statement exerted for detecting anomaly activities.

If threat is detected, it will log the event and saved it to the database as record. This includes the following details, such as which port the attack went through, timestamp for its occurrence and report the total alarm engaged. Else wise, it continues its monitoring task. Last but not least, the result can be viewed on web via Base Analysis and Security Engine (BASE) to observe the patterns of the attacks executed in graphical format such as bar chart. The system will undergo further analysis to determine whether it is a new found threat. Experiments will be carried out to verify the assumption of detecting new threats if needed. New rule is created for detecting the given threat whenever it resurfaces at monitored system in future. This may help in prediction for future attacks by taking into account the timestamp of occurrence and also the targeted port number in network environment. Countermeasure can be instilled to fortify the system before being exploited by the malicious threats.

Implementation and Testing

A DiniB (Detection and Investigation of Network Intrusion Behavior) system is constructed to meet the scope and requirements stated and also further support the gist of this research. SNORT is used to demonstrate the viability of the research. Collection of rule-based statement is done for network packet and traffic filtering. It is used to determine whether a captured activity is malicious or not. For example, to deny access of certain number of hosts from given subnet IP address or to alert the users when malicious threat is detected. The procedure involves configuration of Raspberry Pi as honey pot for logging malicious activities. Raspbian software is downloaded and used as operating system for Raspberry Pi model prior execution of the system. The logged activity retrieved from the database is displayed with the aid of an external monitor, linked to the
Raspberry Pi model by VGA cable. Statistics generated from the logged files can be viewed via Base Analysis and Security Engine (BASE).

This gives an overview of the intruder’s activities rate at certain period of time, such as during daytime or night time. And thus, users may be able to foresee and predict the upcoming attack before it actually occurred. Preventive measure can be imposed beforehand to prevent the supposed intrusion attacks. To evaluate the constructed system whether it fulfills the requirements stated in the design. Several activities are carried out to test for the efficiency of the given software in intrusion detection domain. The tasks involved are as follows: test for planning document, unit testing, module testing, systems integration testing, regression testing, user acceptance testing and traffic load performance testing. It also links to requirements needed for defect logging, tracking and resolution. It begins by using a laptop to initiate scanning/probing activities against Raspberry Pi model. From there, Raspberry Pi model will filter by using the pre-installed SNORT software.

This is to check whether the logged activities have been kept inside the correct database location which is the MySQL server, hosted by Apache web server. Statistical report can be viewed with BASE (Base Analysis and Security Engine) via online platform for learning intrusion behaviour. From the generated report, it indicates the presumed intrusion activities conducted in real scenario. The assumption is further supported by comparing the Detection Rate (DR) and False Positive Rate (FP) from the conducted experiments. This goes by the hypothesis where increase in false positive rate is proportional to the increment of detection rate in intrusion acts. Moreover, it also aids in detecting new threat. For example, new malware where its signature is not stored in the database before. Thus, new rule can be made in order to identify and halt the given threat from harming the system in future.

Conclusion

Implementation of DiniB (Detection and Investigation of Network Intrusion Behavior) system helps in studying the network intrusion behavior within a host-based network. It aims to solve the mentioned problem statement for this research and also lessens the impact from network intrusion behavior circulating the network. In response, prediction could be made for speculating the occurrence of the intrusion events in advance. It is solely based on activity rates and patterns observed from the data collected by using Raspberry Pi model, treated as a honeypot system. New rules will be updated in the rule set to identify and take action against the newly-discovered threat in future. This acts as a milestone for improving network security and also to mitigate unwanted system intrusion by others. As a whole, it is presumed that prevention, detection and response are the fundamental components of network security from the research finding.

Therefore, they deem as the requisites for an effective security programs. Thus, they should be deployed carefully in order to achieve betterment in aspect of network security. As a future work, contribution of this study will be widen to detect and classify new attack from the captured network. In the implementation part, an experiment will be conducting using Raspberry Pi as a honey pot investigation technique for collecting the possible evidence of network intrusions.

References

1. Al Janabi R J S (2010) Malware Avoidance Using Redirection Technique. Journal of Al-Nahrain University 13(3): 178-184.
2. Ahmed AA, Sadiq AS, Zolkibli M F (2016) Traceback model for identifying sources of distributed attacks in real time. Security and Communication Networks 9(13): 2173-2185.
3. Ahmed A A, Mohammed M F (2017) SAIRF: A similarity approach for attack intention recognition using fuzzy min-max neural network. Journal of Computational Science, Elsevier.
4. Ahmed AA, Xue Li C (2017) Analyzing Data Remnant Remains on User Devices to Determine Probative Artifacts in Cloud Environment. Journal of Forensic Sciences 63(1): 112-121.
5. Ahmed AA, Kit YW (2016) MICIE: A Model for Identifying and Collecting Intrusion Evidences. In Signal-Image Technology & Internet-Based Systems (SITIS) 12th International Conference on pp. 288-294, Italy.
6. Ozturk T (2008) The five phases of a successful network penetration. IT Security, CBS Interactive.
7. Abdulghani AA, Aman J, Tat-Chee Wan (2015) Filtration model for the detection of malicious traffic in large-scale networks. Computer Communications 82: 59-70.
8. Ahmed AA, Aman Jantan, Rasmi M (2013) Service Violation Monitoring Model for Detecting and Tracing Bandwidth Abuse. Journal of Network and Systems Management 21(2): 218-237.
9. Choi H, Lee H, Lee H, Kim H (2007) Botnet detection by monitoring group activities in DNS traffic. In Computer and Information Technology CITE 7th IEEE International Conference pp. 715-720.
10. Binkley JR, Singh S (2006) An Algorithm for Anomaly-based Botnet Detection. SRUTI 6: 7-7.
11. Hossein Jadidolestamy (2011) Hierarchical Intrusion Detection Architecture for Wireless Sensor Networks. International Journal of Network Security and its Applications (IJNSA) 3(5): 131-154.
12. Kaur G, Chawla R (2015) Anomaly Based Intrusion Detection System. International Journal of Advance Research in Computer Science and Management Studies 3(8): 99-111.
13. Ahmed AA, Jantan A, Wan T C (2011) SLA-based complementary approach for network intrusion detection. Computer Communications 34(14): 1730-1749.
14. Ahmed AA, Jantan A, Wan TC (2013) Real-time detection of intrusive traffic in QoS network domains. IEEE Security & Privacy 11(6): 45-53.
15. Vijayarani S, Maria Sylvia S (2015) Intrusion Detection System – A Study. International Journal of Security Privacy and Trust Management (IJSPTM) 4(1): 35-36.
16. Rowland C H (2002) U S Patent No 6,405,318. Patent and Trademark Office, Washington DC, USA.
17. Snapp SR, Brentano J, Dias GV, Goan T L, Heberlein LT, et al. (1991) DIDS (distributed intrusion detection system)-motivation, architecture, and an early prototype. In Proceedings of the 14th national computer security conference 1: 167-176.
18. Pedro García Teodoroa, Jesus E Díaz-Verdejoa, Gabriel Macía Fernández, Enrique Vázquezb (2009) Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers Security 28 1-2: 18-28.

19. Mudzingwa D, Agrawal R (2012) A study of Methodologies used in Intrusion Detection and Prevention Systems (IDPS). In Southeastcon Proceedings of IEEE p. 1-6.

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

https://juniperpublishers.com/online-submission.php

This work is licensed under Creative Commons Attribution 4.0 License

DOI: 10.19080/JFSCI.2018.07.555715

How to cite this article: Abdulghani Ali Ahmed, Yee Wai Kit, Amer A Sallam. Raspberry Pi-Based Investigating Model for Identifying Intrusion Evidence. J Forensic Sci & Criminal Invest. 2018; 7(3): 555715. DOI: 10.19080/JFSCI.2018.07.555715