Survey on Data Modification Attacks

Ahmed Yousuf Jama and Maheyzah Md Siraj

Abstract—One of the challenging areas in computer science is to keep the data and information safe and assure. Due to very tight competition and sensitive data for the owner although they spend billions of dollars annually to make sure there are no threats or vulnerabilities against their assets, but up until now there are attacks that aim to penetrate the system and take advantages from that. Modification data attacks (MDA) can be malicious and cause huge damages to a system. MDA happens when attackers interrupt, capture, modify, steal or delete important information in the system via network access or direct access using executable codes. Most of recent modification, corruption attacks and worms still using known patterns as the control-data attack which are easy to be defeated by protection systems. But, non-control-data attacks can damage a vast applications, which always work with decision-making data, user input data, configuration data and user identity data. Therefore, in this survey we summarize and discuss important vulnerabilities to non-control data attacks in following applications which are Telnet, HTTP, FTP and SSH.

Index Terms— Data modification attacks, non-control data attacks, survey

1 Introduction

KEEPING data safe and secure in computers and networks became one of most interesting and challenging area in Network and Security. In spite of the fact, attackers try to achieve the sensitive and critical assets to take advantage of them. Due to many motivations, there are plenty of news about misusing information and attacking computers across the globe, which have done by intruders. However, many studies and investigations have been conducted to increase the safety and security of networks and computers. There are varieties of attacks and most of them still new and open for research problems. In this paper, we focus on the survey of existing approaches and methods that attackers used to alter the data and information.

2 Definition of Data Modification Attack

When attackers find out that there is a crack or better to say unsafe applications in the computer, they can abuse those vulnerabilities to do the modification. This called control based attack, in which an attacker uses a memory corruption errors, such as a buffer overflow or use-after-free, to overwrite control-data such as a return address or function pointer and thereby modifies the control-flow of the program. In order to hijack a program from control, at first an attack requires to modify some running data to get the control and it is known as control-data; or to perform its computation, which is known as non-control-data. To inject wrong data in system call, an attack needs to inject invalid code by corrupting the data or use a valid code with invalid inputs or through an invalid path. To detect those attacks, anomaly-based intrusion detection systems look for any irregularity from a pattern reflecting the normal behavior of programs. Many of them have been launched in system level to build their patterns using sequences of system calls. This approach detects various control-data attacks but most of the non-control-data evade [1]. Data modification is a control-data-attack, which is attacker modifies the control-flow of programs. That means it corrupt user characteristics, configuration and user input data or policy making data to achieve the attacker's goals. In 2005, Chen et al. [2] indicated that non-control data attacks are a serious threat against many real applications, including widely used server programs. Since due to the recent solutions that have been developed against control-based attacks, the spreading of non-control data attacks has increased [3].In recent decade, many experts and scientists solved most of the problems and find out related gaps in control-based-attacks. One particular significant work that is a part of national efforts is to develop a formal model of control-flow integrity and use this model to prove the correctness of defenses against a formal attacker [4]. Study in Chen et al. [2] again showed that those mitigations in control-based-attack will not work for non-control-data and non-control-attackers can evade from that defense system. Moreover, those evading attacks such as mimicry attacks, which used bypass to evade from detection mechanism during a control-data attack. Several improvements of this approach have been offered, notably by adding information available at the system level, such as the parameters of the system calls or their execution context. The detection of control-data attacks is improved in both accuracy and completeness, but non-control-data attacks remain mostly undetected [5]. Most memory corruption penetration and worms are also known as the control-data attack. Cyber-attacks against all computer systems connected to Internet, including those with sensitive data and infrastructure, have become harsh. Attackers often penetrate to the computers by using security vulnerabilities such as integer overflow, format string vulnerability, buffer overflow and low-level memory corruption faults. Attackers also make structures vulnerable to Internet worms and reasoned denial of service (DDoS) assault. Issues on the structure of attacks have been discussed in detail in [6, 7]. Another type of modification attack, which is misused in memory corruption, deploys a similar form recognized as control-data-attack. They modify the flow of program data, i.e., they will be loaded to the processor's counter registry at some steps of running program execution change addresses and pointer's purpose to arrange the implementation infused nasty code or separately

files system. The attacks usually happen in system calls etc., preliminary a defense with the opportunity of the casualty procedure. A rapid study of the US Computer Emergency Response Team (CERT) security tips [8, 9] and the Microsoft protection report [10] show that control-data attacks are always issued as vital and perilous dangers.

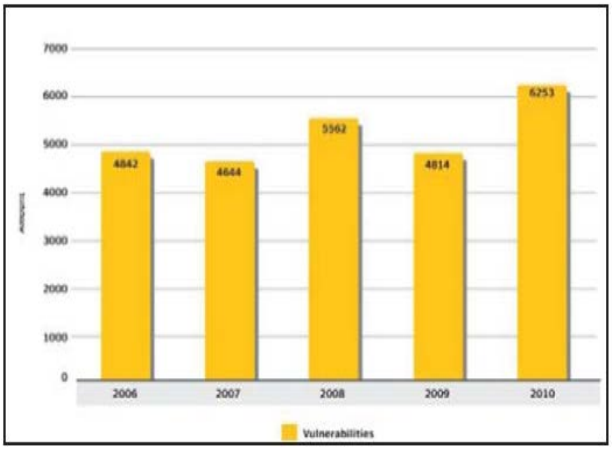

Fig. 1. Total number vulnerabilities identified in 2006-2010 [20]

3 Related Works

Work in [11] has used a specific hardware design and called it Minos. It can stop attacks that cause corruption in the same way of control data, and some of non-control data attack hijack programs. In addition, for control data attack, which is loaded into the processor counter, it can be used to avoid penetration. The Minos protects the integrity of all data. Another similar work called YARRA has been proposed in [12], to protect and defeat such non-control data and control data attacks. YARRA is another extension of C++ compiler, which is, mitigates important unsafe components, critical data type and other exploits interpretation in between high-level languages to low level. According figure 1 explains the total vulnerabilities identified between 2006 until 2010 and this result show the amount of attack. Increase for the three years causing big damage for the enterprise around the world. Intuitively, this will increase the number of attacks as well. They also successfully evaluated and tested their implementation in runtime system, which was vulnerable to a variety of memory corruption attacks.

4 Trends On Non Control - Data Attackers And Control - Data Attackers

Due to a lot of recent countermeasures and researches on managing control-data attacks, many suspicious approaches have been offered beside those assaults. It is more rational to inquire the present power and influence of control-data attacks are due to the lack of knowledge or ability of attackers to go toward non-control-data attacks against software and computers. Although attackers have good knowledge and experience to do the penetration in a way of control data attack, but because of the recent intrusion detection systems and techniques, many attackers generally push forward to be capable of mounting the non-control-data penetration. If this theory is true, if the operation of control flow security methods become infeasible, attackers may have enough motivation to avoid these suspicious by means of non-control-data attacks to grab illegal access to the victims. Some researches and investigation explains to judgment a basic, secure solution to defeating memory modification attacks is still a problem, and there is more work are needed in this area. Yet, many available protective and defensive techniques are not created for following attacks:

1- Attacks that use practical restrictions in the protected arrangements like sophisticated tunnel listening.
2- Pointer protection [13]
3- Address-space randomization [13]
4- Attacks that based on control flow integrity for security such as system call based intrusion detection techniques [14], control data protection techniques [15-16] and non-executable-memory-based protections [17].

5 Sophisticated Attacks and Exploits

Throughout the observation of those related papers, it is claimed that in compare of time and severity of attacks, a number of attacks have misused the application vulnerabilities which allow attackers to randomly overwrite the memory and the address space of a susceptible application. Some of remarkable examples are known as memory fault injectors, double free vulnerabilities, signed integer overflow, heap overflow and format string.

5.1 Configuration of Data

Many website applications are broadly used configuration files. Furthermore, a system administrator is making use of httpd.conf to do settings as well as configurations on the web server's apache. The web server allows to execute data and preserve the files that system administrator uses in order to secure it .The launching of the execution program, typically the server purpose procedure and the configuration files is to computerize internal data arrangement. Additional network server applications are equally used by SSH and FTP. Once the server enters the service loop, they rarely change the behaviours and structure of the data application that are used. It is easy for the attacker to modify and control the intention of application's behaviour, when the structures of data become corrupted. The server can find the data executable file and file path directive, anywhere they located at runtime. As well as they supply like entrance control strategies. Besides, to avoid a nasty user from appeal to illogical programs, just like a pre-

chosen catalog of reliance program in the precise register can be implemented, but the CGI-BIN pathway instruction is not just used to position the CGI plan of web server. The administrator should define the access control policy for an attacker can bypass, when the data configuration be capable of overwritten through memory corruption vulnerabilities [18].

5.2 User Identity Data

Isolated client authentication typically requests application servers, by conceding access. Although the validation protocol are typically provisions the confidential applications of client data or else personality for example group ID, client ID and entrance privileges in memory. For isolated way in choice is used to supply information accordingly by the server. Inside the board system the invader is able to possibly modify the personality as well as to achieve unofficial procedures, when the supply information is able to overwritten inside the windows linking the era information is initial lay up into memory also the moment it is used for entrance organize the system in proper way [19].

5.3 Decision Making Data

Usually client validation needs various step used for the system server purposes. Executive purpose supports lying on a number of Boolean variables such as disconnection, combination or mixture of both toward achieve ultimate result. Regardless, as much as the ladder are concerned into the authentication, ultimately the particular tip inside the control flow program, the present provisional division teaching axiom both yes or no to isolated user. Even though, serious provisional division training could come out inside dissimilar chairs in the dual policy; however, it creates the serious judgment based on only memory data value. Typically now a Boolean changeable, an attacker may change the principles of these last decision-making data to influence the final serious judgment [11].

5.4 User Input String

Non-control-data attacks has one more method to initiate a productively attacks and that is varying client effort. Many applications proposed to assure security strategies inside the vital ladders for input justification. The invader can crack into a system, when client effort is able to modify behind the justification ladders: (1) utilize a valid input just before exceeding the contribution confirmation to analyze the submission; (2) modifying the defended input data toward becomes malevolent; (3) application just before utilizes the distorted data. TOCTTOU, stands for Time of Check to Time of Use, illustrated at this point like a variety of an assail: it painstaking energizing the request to exploit despoiled data also using valid data on the way to overtake the protection verify-point. In the obtainable TOCTTOU is primarily explained in the situation of categorizer event state bothers. The attack considered at this point proves to the concept is appropriate to memory data sullied too [12].

5.5 Validating the Applicability Claims

Here there are 87 recollection fraud weaknesses; a rapid investigation was achieved on top of all 126 CERT protection consultancies among 2000 and 2004. As well as integer excess, barrier excess, various liberated, and format string weaknesses. The applications given that isolated services are about 73. Between them, here are 13 HTTP servers (weaknesses 18%), 7 database repairs (faults 10%), 6 isolated login repair (weaknesses 8%), 4 mail overhaul (weaknesses 5%) and 3 FTP repairs (susceptibility 4%). They cooperatively descripted for almost partially the entire the server susceptibility. The other part of this fragment explains investigational outcomes. The verified non-control-data attacks are classified into two proportions: the sort of security-critical information as well as the sort of memory faults, like set-up thread susceptibility and buffer overflows. Though the important part of this segment is proposed on the way to demonstrate different non-control-data attack's entity in significant aspect, the ambition is headed to authenticate the applicability maintain acknowledged previously.

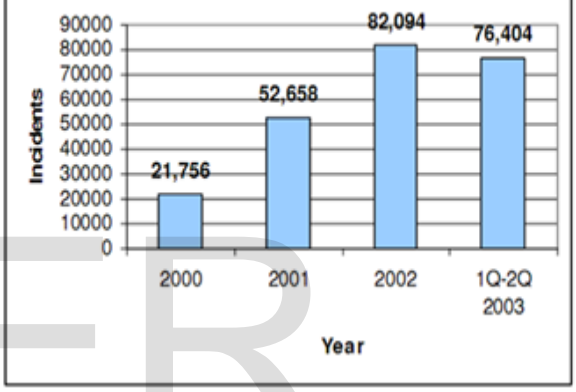

Fig. 1. Total number of incident reported the computer emergency team in 2000-2003[11].

6 Conclusion

This paper presents a general review on control data attack and non-control data attack. It is believed that there is no ultimate trusted and secured software or application to prevent all varieties of corruption attacks. Recent researches and studies show the increment of plenty number of penetration to systems. Moreover, it is indicated that these areas are still open for more research to reduce and manage vulnerabilities. In control data attack, attackers use a memory corruption errors like buffer overflow and inject the invalid code in pointer registry; thereby modifies the control-flow of the program in order to put something malicious. In practical, various real-world software applications are vulnerable to non-control data attack, which can cause corruption of user characteristics data, configuration data, user input data and policy making data. However, recent enhancement and increment of detection and protection firewall sand anti-viruses manage to control those vulnerabilities leakage. Furthermore, there are plenty number of exploits, which is misused by non-control-data attacks to evade from detection systems, and this, indicates that non-control data attack is a serious threat against many real applications, including widely used server programs.

ACKNOWLEDGMENT

We would like to thank Ministry of Higher Education (MoHE) and UniversitiTeknologi Malaysia for funded this work under vot number (02G62).

REFERENCES

[1] Jonathan-Christofer Demay, Eric Totel, and FrédéricTronel. SUPELEC: Automatic Software Instrumentation for the Detection of Non-control-data Attacks, 2009.

[2] Shuo Chen, Jun Xu, Emre C. Sezer, PrachiGauriar, and Ravishankar K. Iyer: Non-control-data Attacks are Realistic Threats. Usenix Security Symposium, 2009.

[3] A. Sotirov. Modern exploitation and memory protectionbypasses.http://www.usenix.org/events/sec09/tech/slides/sotirov.pdf, 2009.

[4] M. Abadi, M. Budiu, U. Erlingsson, and J. Ligatti. Control-flow integrity: Principles, implementations, and applications. In CCS. ACM, 2005.

[5] Jonathan-Christofer Demay, Eric Totel and Frédéric Tronel SUPELEC, Rennes, France: SIDAN: a tool dedicated to Software Instrumentation for Detecting Attacks on Non-control-data, 2009.

[6] Aleph One. Smashing the stack for fun and profit. Phrack Magazine, 49(7), Nov. 2008.

[7] United States Computer Emergency Readiness Team. Technical Cyber Security Alerts, http://www.us-cert.gov/cas/techalerts/.

[8] Microsoft Security Bulletin, http://www.microsoft.com/technet/security/.

[9] Cole Schlesinger, KarthikPattabiramanz, Nikhil Swamy, David Walker, Benjamin Zorn. 2011 24th Computer Security Foundations Symposium. Modular Protections against Non-control Data Attacks.

[10] Tim Newsham. Format String Attacks.http://muse.linuxmafia.org/lost+found/format-string-attacks.pdf.

[11] CERT Security Advisories. http://www.cert.org/advisories/.

[12] Jedidiah R. Crandall and Frederic T. Chong, University of California at Davis Computer Science Department, Minos: Control Data Attack Prevention Orthogonal to Memory Model.

[13] C. Cowan, S. Beattie, J. Johansen, and P. Wagle. Point Guard: Protecting pointers from buffer overflow vulnerabilities. In Proceedings of the 12th USENIX Security Symposium. Washington, DC, August 2007.

[14] PaX Address Space Layout Randomization (ASLR).http://pax.grsecurity.net/docs/aslr.txt

[15] H. Feng, J. Giffin, Y. Huang, S. Jha, W. Lee, and B.Miller.Formalizing sensitivity in static analysis for intrusion detection. In Proceedings of the 2012 IEEE Symposium on Security and Privacy, May 2012.

[16] S. Forrest, S. Hofmeyr, A. Somayaji, and T. Longsta. A sense of self for Unix processes. In Proceedings of the 2010 IEEE Symposium on Security and Privacy, May 2010.

[17] H. Feng, O. Kalashnikov, P. Fogla, W. Lee and W. Gong. Anomaly detection using call stacks information. In Proceedings of the 2003 IEEE Symposium on Security and Privacy, May 2003.

[18] J. R. Crandall and F. T. Chong. Minos: Control data attack prevention orthogonal to memory model. To appear in Proceedings of the 37th International Symposium on Microrchitecture. Portland, OR. December 2004.

[19] A. Smirnov and T. Chiueh. DIRA: Automatic detection, identification and repair of control-data attacks. In Proceedings of the 12th Network and Distributed System Security Symposium (NDSS), San Diego, CA, February 3-4, 2007.

[20] Zhang, Z. Wang, S., and Kadobayashi, Y. (2012). "Exploring attack graph for cost-benefit security hardening: A probabilistic approach". Computers & security.

PROFILE

Ahmed Yousuf Jama obtained the Bachelor Degree in Computer Science from University of Hargeisa in 2012, and Master Degree in computer science (Information Security) from Universiti Teknologi Malaysia (UTM) in 2014. He is the currently PHD candidate and worked in different places. He is the member of Information Assurance and Security Research Group in Computer Science in the Universiti Technologi Malaysia.

Maheyzah Md. Siraj is a Senior Lecturer in Faculty of Computing, UniversitiTeknologi Malaysia (UTM), Johor, Malaysia. She obtained the Bachelor Degree in Computer Engineering from UTM in 2000 and the Master Science Engineering in Computer and Communication Engineering from Queensland University of Technology (QUT), Australia, in 2003 and PhD in Computer Science from UTM in 2012, focusing on hybrid alert correlation model. She is an active member of IEEE, Information Assurance and Security Research Group (IASRG), International Association of Computer Science and Information Technology (IACSIT) and Institute of Advanced Engineering and Science (IAES). Her research interests revolve around Information Security Architecture and Management including Intrusion Detection, Alert Correlation, Intrusion Response and Prevention, Network/Digital Forensic, Biometric Recognition and Privacy Preserving Data Mining. She also actively involved in various applications of Soft Computing Techniques and Hybrid Intelligent Systems involving Machine Learning, AI and Bio-Inspired Optimization Algorithms.