Advanced Persistent Threat Detection using Malware Infection

Tajagn Jagani[1], Sachin Todkari[2]

[1]M.E (Computer) Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India -411007

[2]M.E Prof (Computer) Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India -411007

Abstract: ***APT (Advanced Persistent Threat) is a genuine risk to the Internet. With the help of APT malware, attackers can remotely control infected machine and steal the personal information. DNS is well known for malware to find command and control (C&C) servers. The proposed novel system placed at the network departure guide that points toward effectively and efficiently detect APT malware infections based on malicious DNS and traffic analysis. To detect suspicious APT malware C&C domains the system utilizes malicious DNS analysis method, and afterward analyse the traffic of the comparing suspicious IP utilizing anomaly-based and signature based detection innovation. There are separated features in view of big data to describe properties of malware-related DNS. This manufactured a reputation engine to compute a score for an IP address by utilizing these elements vector together.***

Keywords: APT; Intrusion Detection; Malware Infections; DNS.

1. Introduction

The Advanced Persistent Threat (APT) defects or attacks are spreading on the web these days. Unfortunately, they are critical to detect an APT. It is a continuous or persistent attacking processes and collection of mainly focusing on a particular entity with high-value information, for example, government, military and the monetary business. The aim of an APT assault is to steal the information instead of to make harm the association or system. First installing so as to hack into the system has been accomplished, APT malware on the contaminated machine by attacker. For a while, APT malware is, Trojan horse or backdoor secondary passage, is customized for firewalls and anti-virus software of the target network. It is not only used for remotely controlling the traded off machine in the APT assault, additionally to steal touchy information from extended period of time.

APT malware is altogether different than other malware like the worms and bots. The basic intension of APT malware is to remotely control the machines and to steal private data of user, rather than to launch denial-of-service attacks, which causes damage or send spam emails. For example, in the case of those worms and bots, the attackers need to utilize the C&C servers to remotely control thousands of infected host. But APT attackers do not use the same Command & Control server to remotely control so many infected end user machines because it increases the risk of exposure.

For identify malicious domains that involved in APT malware activity is a challenge. The crafted malware in APT attack do not use DGA domains or malicious flux service. The domains for APT malware were registered by the attacker. Compared with these bots and worms the crafted malware requires high degree of stealth. Because of this reason the DNS behavioural features of APT malware are inconspicuous. It is too hard to analyse large volumes of outbound and inbound traffic in a large network, such as an ISP and a large enterprise. Detection of APT malware infections in a big network is another challenging problem.

2. Literature Survey

[8] Recent Botnets such as Conficker, Kraken and Torpig have used DNS based “domain fluxing” for command-and-control, where each Bot queries for existence of a series of domain names and the owner has to register only one such domain name. In this paper, developed a methodology to detect such “domain fluxes” in DNS traffic by looking for patterns inherent to domain names that are generated algorithmically, in contrast to those generated by humans. In particular, we look at distribution of alphanumeric characters as well as bigrams in all domains that are mapped to the same set of IP-addresses. We present and compare the performance of several distance metrics, including KL-distance, Edit distance and Jaccard measure. We train by using a good data set of domains obtained via a crawl of domains mapped to all IPv4 address space and modeling bad data sets based on behaviors seen so far and expected. We also apply our methodology to packet traces collected at a Tier-1 ISP and show we can automatically detect domain fluxing as used by Conficker botnet with minimal false positives.

[3] Denial-of-Service (DoS) attacks pose a significant threat to the Internet today especially if they are distributed, i.e., launched simultaneously at a large number of systems. Reactive techniques that try to detect such an attack and throttle down malicious traffic prevail today but usually require an additional infrastructure to be really effective. The paper we shows that preventive mechanisms can be as effective with much less effort: Presents an approach to (distributed) DoS attack prevention that is based on the observation that coordinated automated activity by many hosts needs a mechanism to remotely control them. To prevent such attacks, it is therefore possible to identify,

infiltrate and analyze this remote control mechanism and to stop it in an automated fashion. We show that this method can be realized in the Internet by describing how we infiltrated and tracked IRC-based botnets which are the main DoS technology used by attackers today.

[11]Modern botnet trends have lead to the use of IP and domain fast-fluxing to avoid detection and increase resilience. These techniques bypass traditional detection systems such as blacklists and intrusion detection systems. DNS is one of the most prevalent protocols on modern networks and is essential for the correct operation of many network activities, including botnet activity. For this reason DNS forms the ideal candidate for monitoring, detecting and mitigating botnet activity. In this paper a system placed at the network edge is developed with the capability to detect fast-flux domains using DNS queries. Multiple domain features were examined to determine which would be most effective in the classification of domains. This is achieved using a C5.0 decision tree classifier and Bayesian statistics, with positive samples being labelled as potentially malicious and negative samples as legitimate domains. The system detects malicious domain names with a high degree of accuracy, minimising the need for blacklists. Statistical methods, namely Naive Bayesian, Bayesian, Total Variation distance and Probability distribution are applied to detect malicious domain names. The detection techniques are tested against sample traffic and it is shown that malicious traffic can be detected with low false positive rates.

[2] The performance and operational characteristics of the DNS protocol are of deep interest to the research and network operations community. In this paper, we present measurement results from a unique dataset containing more than 26 billion DNS query-response pairs collected from more than 600 globally distributed recursive DNS resolvers. We use this dataset to reaffirm findings in published work and notice some significant differences that could be attributed both to the evolving nature of DNS traffic and to our differing perspective. For example, we find that although characteristics of DNS traffic vary greatly across networks, the resolvers within an organization tend to exhibit similar behavior. We further find that more than 50% of DNS queries issued to root servers do not return successful answers, and that the primary cause of lookup failures at root servers is malformed queries with invalid TLDs. Furthermore, we propose a novel approach that detects malicious domain groups using temporal correlation in DNS queries. Our approach requires no comprehensive labeled training set, which can be difficult to build in practice. Instead, it uses a known malicious domain as anchor, and identifies the set of previously unknown malicious domains that are related to the anchor domain. Experimental results illustrate the viability of this approach, i.e. , we attain a true positive rate of more than 96%, and each malicious anchor domain results in a malware domain group with more than 53 previously unknown malicious domains on average.

[20] Now a day's Intrusion Detection systems plays very important role in Network security. As the use of internet is growing rapidly the possibility of attack is also increasing in that ratio. People are using signature based IDS's. Snort is mostly used signature based IDS because of it is open source software. World widely it is used in intrusion detection and prevention domain. Basic analysis and security engine (BASE) is also used to see the alerts generated by Snort. In the paper we have implementation the signature based intrusion detection using Snort. Our work will help to novel user to understand the concept of Snort based IDS.

3. Proposed Systems

IDnS is designed to detect malicious domains used for malware in APT attacks and for detection of infected machine. For this purpose analysis of large volumes of DNS traffic which can be called big data. And also analyse the network traffic of large numbers of suspicious malware Command & Control servers. The features extracted from very big data for detection consist of malicious DNS and network traffic features. By studying the behaviours of the benign applications and crafted malware, they achieved to extract distinguishable network traffic features are able to define the APT malware Command & Control traffic.

Our system IDnS uses signature-based detection and anomaly-based detection together to provide the maximum defense for the monitoring network.

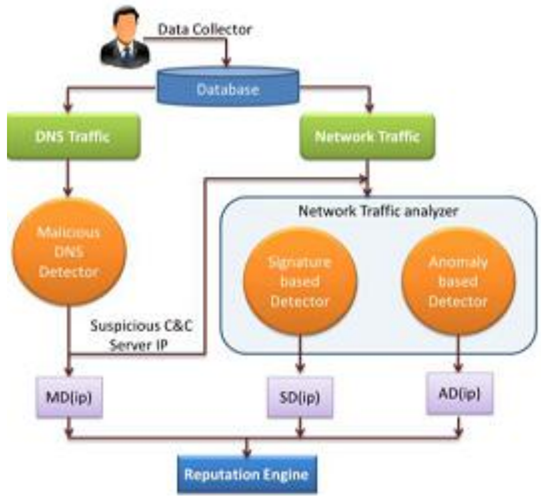

Figure 1: Architecture

4. Advantages of Proposed System

1) It is a useful intrusion system that can help to fight against cyber-crime such as theft of information from infected host.
2) Build a reputation engine for decides whether an IP address is infected or not infected.
3) Malicious DNS analysis is first performed to find out suspicious IP addresses of command and control servers in this approach.
4) The system can significantly increase the detecting accuracy as well as efficiency.

5. Mathematical Model

Let W is the set of whole of system which consists:
W= {input, process, output}.
Input={D, MDNS, RE, NTA}

Where, D is the set of data collector.

1. MDNS is the set of malicious DNS detector which detects the malicious IP at DNS server traffic.
2. NTA is the network traffic analyser which detects the network traffic.
3. RE is the reputation engine which calculates the reputation score of an IP address.

6. Comparative Result

Parameters	Existing (%)	Proposed (%)
Protocol anomaly	65.8	76
Statistical anomaly	80.9	79.5
Application anomaly	89.5	96.5

7. Conclusion

In this, a proposed a system IDnS placed at the network egress points to detect malware infections inside the network combined with DNS traffic analysis and network traffic analysis. Extracted new features and built a reputation engine based on big data which is calculated from two detections methods one is signature based detection and anomaly based detection. Further the anomaly based detection system is divided into three phases named protocol anomaly, statistically anomaly and application anomaly detection. The system processes advantages of high efficiency and accuracy. The experimental results show that this security approach is feasible for improving the sustainability of the system and is good at detecting APT malware infections.

8. Acknowledgement

The authors would like to thank the researchers as well as publishers for making their resources available and teachers for their guidance. We are thankful to the authorities of IJSR, organized by, for their constant guidelines and support. We are also thankful to the reviewer for their valuable suggestions. We also thank the college authorities for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

References

[1] V. Kumar and D. O. P. Sangwan, “Signature based intrusion detection system using SNORT," Int. J. Comput. Appl. Inf. Technol., vol. 1, no. 3, pp. 3541, 2012.
[2] H. Gao et al., “An empirical reexamination of global DNS behavior," in Proc. ACM SIGCOMM Conf. SIGCOMM, 2013, pp. 267_278.
[3] F. C. Freiling, T. Holz, and G. Wicherski, “Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks,"
[4] Lect. Notes Comput. Sci., vol. 10, no. 2, pp. 319_335, 2005.
[5] A. Karasaridis, B. Rexroad, and D. Hoe_in, “Wide-scale botnet detection and characterization," in Proc. 1st Conf. 1st Workshop Hot Topics Understand. Botnets, 2007, p. 7.
[6] J. Jung, M. Konte, and N. Feamster, “Dynamics of online scam hosting infrastructure," in Proc. 10th Int. Conf. Passive Active Netw. Meas., 2009, pp. 219_228.
[7] H. Porras, H. Saïdi, and V. Yegneswaran, “A foray into Con_cker's logic and rendezvous points," in Proc. USENIX Conf. Large-Scale Exploits Emergent Threats, Botnets, Spyware, Worms, More, 2009, p. 7.
[8] S. Yadav, A. K. K. Reddy, A. L. N. Reddy, and S. Ranjan, “Detecting algorithmically generated malicious domain names," in Proc. ACMSIGCOMM Conf. Internet Meas., 2010, pp. 48_61.
[9] J. Wolf. (2008). Technical Details of Srizbi's Domain Generation Algorithm. Online.. Available: http://tinyurl.com/6mdasc
[10] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi, “EXPOSURE: Finding malicious domains using passive DNS analysis," in Proc. NDSS, 2011.
[11] E. Stalmans and B. Irwin, “A framework for DNS based detection and mitigation of malware infections on a network," in Proc. Inf. Secur. South Africa (ISSA), Aug. 2011, pp. 1_8.
[12] M. Antonakakis, R. Perdisci, D. Dagon,W. Lee, and N. Feamster, “Building a dynamic reputation system for DNS," in Proc. 19th USENIX Secur. Symp., 2010, pp. 273_290.
[13] LASTLINE. (2015). Using Passive DNS Analysis to Automatically Detect Malicious Domains. Online.. Available: https://www.lastline. com/papers/dns.pdf
[14] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou, II, and D. Dagon, “Detecting malware domains at the upper DNS hierarchy,” in Proc. USENIX Secur. Symp., 2011, p. 27.
[15] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, “Measuring and detecting fast- ux service networks," in Proc. NDSS, 2008.
[16] N. Brownlee, K. Claffy, and E. Nemeth, “DNS measurements at a root server," in Proc. IEEE Global Telecommun. Conf. (GLOBECOM), vol. 3.2001, pp. 1672-1676.
[17] S. Castro, D. Wessels, M. Fomenkov, and K. Claffy, “A day at the root of the Internet," ACM SIGCOMM Comput. Commun. Rev., vol. 38, no. 5,pp. 41-46, 2008.
[18] H. Gao et al., “An empirical reexamination of global DNS behavior," in Proc. ACM SIGCOMM Conf. SIGCOMM, 2013, pp. 267-278.
[19] [19] R. Perdisci, I. Corona, D. Dagon, and W. Lee, “Detecting malicious ux service networks through passive analysis of recursive DNS traces," in Proc. Annu. Comput. Secur. Appl. Conf. (ACSAC), Dec. 2009, pp. 311-320.
[20] H. Choi, H. Lee, H. Kim, and H. Lee, “Botnet detection by monitoring group activities in DNS traf c," in Proc. 7th IEEE Int. Conf. Comput. Inf.Technol. (CIT), Oct. 2007, pp. 715-720.
[21] Source re. (2015). Snort Network Intrusion Detection System Web Site. Online.. Available: https://www.snort.org/
[22] M. Roesch, “Snort Lightweight intrusion detection for networks," in Proc. 13th LISA, 1999, pp. 229-238.

[23] V. Kumar and D. O. P. Sangwan, “Signature based intrusion detection system using SNORT," Int. J. Comput. Appl. Inf. Technol., vol. 1, no. 3, pp. 35-41, 2012.

[24] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, “Anomaly-based network intrusion detection: Techniques, systems and challenges," Comput. Secur., vol. 28, nos. 1-2, pp. 18-28, 2009.

[25] F. Gong, “Deciphering detection techniques: Part II anomaly-based intrusion detection," McAfee Security, White Paper, 2003.

[26] N. Villeneuve and J. Bennett. (2012). Detecting apt activity with network traf c analysis. Trend Micro Inc. Online. Available: http://www.trendmicro.com/cloud-content/us/pdfs/securityintelligence/white-papers/wp-detecting-apt-activity-with-network-traf c-analysis.pdf, accessed Oct. 31, 2013.

[27] Kaspersky Lab. (2015). Targeted Cyberattacks. Online.. Available: https://apt.securelist.com/

[28] K. Koutroumbas and S. Theodoridis, “Pattern recognition," in Encyclopedia of Information Systems. 2003, pp. 459-479.

[29] (2015). Malware Domains List. Online.. Available: http://www.malwaredomains.com//

Author Profile

Mr. Tajagn N Jagani is currently pursuing M.E (Computer) from Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India. Savitribai Phule Pune University, Pune, Maharashtra, India. He received his B.E (Information Technology) Degree from BVM college of engineering VV Nagar, Gujarat India. Sardar Patel University, VV Nagar Gujarat India. His area of interest is Network security.

Porf. Sachin V Todkari BE CSE college of Engg Ambajogai ME IT MIT COE KOTHRUD PUNE Area: Wireless Network