INFORMATION SYSTEMS IN MANAGEMENT

Information Systems in Management (2016) Vol. 5 (4) 473−480

MANAGEMENT SYSTEM FOR DYNAMIC ANALYSIS OF MALICIOUS SOFTWARE

KRZYSZTOF CABAJ

Institute of Computer Science, Warsaw University of Technology

In the recent years, one can observe the increase in the number of malicious software (malware) samples analyzed by the antivirus companies. One explanation is associated with attacker's antivirus systems hider tactic, which modifies hostile programs form, without changing it functionality. In effect, the first step of analysis is associated with the check if a given sample is a new threat or modification of existing. Very often such simple test can be performed automatically by dedicated information system. Paper describes information system, which allows dynamic analysis of malicious sample. Presented system was developed and deployed in the Institute of Computer Science, Warsaw University of Technology. During performed security research concerning ransomware threats system proves its usefulness. Additionally, the system become a knowledge base of known malware recently analyzed by our security team.

Keywords: malware, ransomware, dynamic analysis, expert system

1. Introduction

Today, the first step of analysis is associated with the check if a given sample is a new threat or modification of existing. Malicious samples can be analyzed using one of two methods - static or dynamic analysis. In the first one, analyzed sample is manually decompiled and dissembled. In effect person, who performs analysis, poses assembler code of the whole program. Using these instructions, analyst can manually reconstruct what actions perform malicious software. This is

not an easy task, due to the complexity of modern malicious software and intentionally introduced by the attacker parts of code, which disable modern software used for reverse engineering. For example, some assembler instruction in malicious program can jump into middle of other instruction - such program works; however disassembler did not properly decode further part of the code. On the other hand, analyst can perform dynamic analysis. In this kind of analysis malicious sample is executed in controlled environment and its activity is observed. This type of analysis is faster; however some malware functionalities can be omitted. Due to our previous experience with this kind of analysis and rapid rise of number of gathered samples, we decided to use the dynamic approach. However, even dynamic analysis of such number of samples without dedicated support software becomes almost impossible. For example, during analysis of CryptoWall ransomware family we gathered more than three hundreds of samples. The most important information found during analysis concerns detection of Command and Control Servers (C&C servers). Shortly number of such machines rise over one thousand. Due to this fact we have decided to develop and deploy in the network of the Institute of Computer Science, Warsaw University of Technology dedicated system which supports our work. The system allows automatic dynamic analysis of new samples and initial analysis of used by tested samples Command and Control servers. The system was called ARTA (Automatic Ransomware Traffic Analyzer). To easy work system has Web interface which allows addition of new samples and search in all gathered data. The first weeks of work with ARTA system shows that it becomes a knowledge base concerning analyzed malware samples.

The paper is organized as follows. The next section introduces currently most active malicious type called ransomware. This kind of malware locks victim's computer and demand some money for enabling access. The third section describes ARTA system in details. The fourth section presents preliminary results of dynamic analysis concerning CryptoWall 3.0 ransomware. The last, fifth section contains conclusions.

2. Ransomware malware

In the last decade the main motive of attackers' actions is associated with money. In the previous years the most precious treasures are credit cards numbers or data used for accessing to e-banking systems. However, it is worth mentioning that as a reaction to these threats from, financial organizations made such attacks harder. From the last few years, more and more popular are attacks that lock victim's computers and demand some ransom for enabling access to the infected machines. Due to fact of ransom request malware used during these attacks is called ransomware. Reports prepared by antivirus companies' show huge increase in this kind of attacks in the last two years. For example, McAfee shows that only

in the first quarter of the 2015 year, the number of observed ransmoware samples rose by 165% [1]. Symantec shows even more horrific data, accordingly to its report number of ransomware which encrypts files in the hard drive rise almost 45 times, from 8274 samples observed in 2013 to the 373342 in 2014 [2].

The first generation of ransomware only locks access to the computer, preventing logging to the machine. For skilled users these threats can be easily overcome. In the most severe cases full system reinstallation is needed; however, all user's data stored in the infected machine can be restored. Due to this fact, shortly second generation of ransomware will become popular which works in more hostile fashion. In this generation malware encrypts various types of files associated with user precious data generated by, for example, word processor, spreadsheet or game - yes, some ransomware encrypts games' saves files. As such program in most cases uses modern encryption algorithms, like AES (Advance Encryption System), the decryption without key is almost impossible. The first examples of this new generation utilize symmetric-key algorithms, which use the same key for encryption as well decryption. In effect, this key can be extracted from poor implementation of the malware (key was not deleted after encryption of the whole data) or during its transfer from victim to the attacker. However, the most sophisticated ransomware family called CryptoWall 3.0, which appears at the beginning of the 2015 starts using asymmetric-key encryption algorithm. Such algorithm uses two separate keys: public used for encryption and private used for decryption. In this situation both keys are generated somewhere in the Internet and only public key used for encryption of users' data is transferred to the infected machine. Private key used for decryption newer appears in the victims' machines. Considering that this malware uses 2048 bit RSA asymmetric-key algorithm, decryption of victim's data without private key is unfortunately impossible. Detailed analysis of various ransomware families can be found in [3].

At the end of March 2015 our security group clean infected machine in the Institute of Computer Science. Gathered from victim's machine CryptoWall 3.0 sample was examined using dynamic analysis. Performed analysis reveals very interesting behavior concerning network activity of the ransomware. After infection the sample contacts attacker's Command and Control server using list of prior infected Web servers. We called this kind of Web server a CryptoWall proxy. What should be emphasized, these servers are innocent victims, too. A few analyzed samples show that these proxies' lists contain many infected servers, and are centrally managed by attackers. Detailed description of this network activity and results from initial analysis can be found in [4]. During continuation of this research many samples of CryptoWall family are soon gathered. Manual analysis of all samples soon became almost impossible. We decided to develop and deploy automatic system which can support our research. The ARTA system is a result of this work.

3. ARTA system

The most important and interesting results gained from dynamic analysis of gathered CryptoWall 3.0 samples are connected to CryptoWall proxy servers list. The initial research shows, that attacker prepares a list of previously compromised Web servers and uses it for a while in several distinct variants of malware. When the numbers of active proxy servers, these which still are not cleaned by administrator, fall below some level, the attacker starts using the new list. As was described in the previous chapter, soon the number of analyzed samples increased and its' manual analysis became almost impossible. The importance of reviled from dynamic analysis data and huge numbers of new samples lead to the development of support system called ARTA. The system combines various systems developed and deployed in the network of Institute of Computer Science, for example, Maltester and WebHP. The ARTA system not only uses other security systems for automation of analysis process, but also contains Web interface enabling easy management and access to gathered data. The first observations from work with ARTA system prove that it supports the analyst. The usage of the system decreases analysis of single sample from about ten minutes to less than three minutes. Moreover, it shortly becomes main knowledge base and source of all information concerning analyzed malware samples in our security team.

Dynamic analysis of malware sample is performed using, deployed in the network of Institute of Computer Science, Maltester environment. The Maltester system allows execution of malicious samples using a clean virtual machine in controlled environment that protects other network users. Moreover, after preconfigured time execution is automatically stopped and machine is investigated, for example, for downloaded malicious executables. More details concerning Maltester environment can be found in [4]. During manual analysis, ransomware samples are passed to the Maltester system using dedicated management software from command line interface. ARTA system directly contacts Maltester using dedicated XML-RPC interface. The analyst simply adds new sample in Web interface, and results are provided in less than three minutes, where two minutes are used for execution of analyzed sample.

When samples are analyzed manually all traffic related to the DNS service, which is used for resolving human readable names to the IP address used in Internet, was forwarded to custom DNS server. This server records all queries, and responds that domain is unknown. This behavior leads analyzed sample to try all proxy servers contained in hardcoded proxies list. Gathered from the custom DNS logs contains server list. Further the list was manually sorted and compared to the previous ones. ARTA system automates this process as well. Custom DNS server is used, but instead blocking of all queries, it forwards traffic to our dedicated Web HoneyPot called WebHP [5, 6]. The WebHP system records all details concerning

received connections and stores them in the data base. Simple python script extracts all used by the analyzed sample domains, sorts them and compare results with previously analyzed samples. In effect ARTA system gives to the analyst information whether this sample uses new or previously observed list of CryptoWall proxy servers.

All accesses to the data base and executions of external programs are hidden from the analyst, which interacts with the ARTA system using dedicated Web interface. Web interface is used for management of analysis processes and access to the results. It is implemented using Python language and Django framework. Figure 1 presents sample screen of ARTA system concerning detected CryptoWall proxy lists.

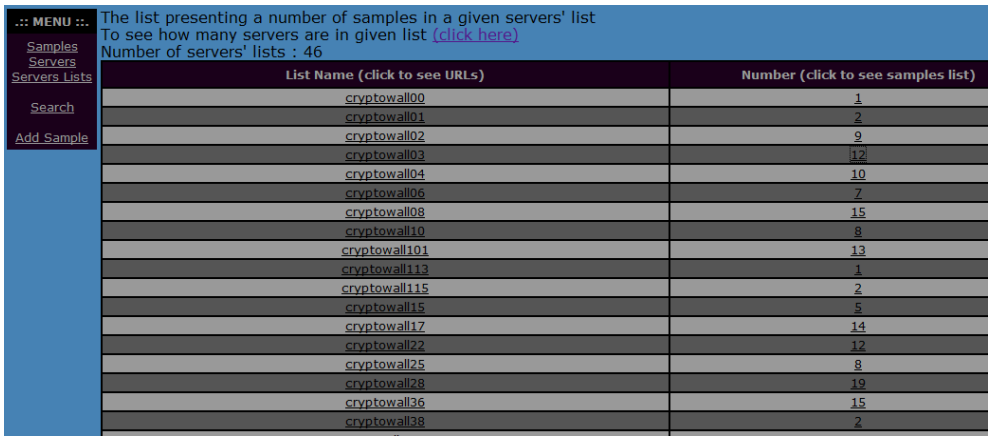

Figure 1. Sample screen of ARTA system - detected CryptoWall proxy lists

Developed interface is interactive; where it is possible presented data are connected using hyperlinks to other views. For example, in the view from figure 1 presenting CryptoWall proxy servers' list user can simply navigate to another view containing list of URLs or samples associated with particular name. All this actions can be simply accessed by ne click using appropriate hyperlinks. As first experiences shows, this simple solution eases further analytical work which uses results from previous analyses.

Figure 2 presents all previously described elements which in cooperation forms ARTA system. All elements are installed on five dedicated virtual machines running in one physical machines controlled by the Xen hypervisor. For security purposes, two virtual networks which connects all elements are introduces. The first one contains potentially hostile traffic generated by infected machine. This traffic is recorded for further analysis and firewalled before forwarded to the Internet. The second network is used for the management of analysis process. Using this network analyst access Web interface if the ARTA system.

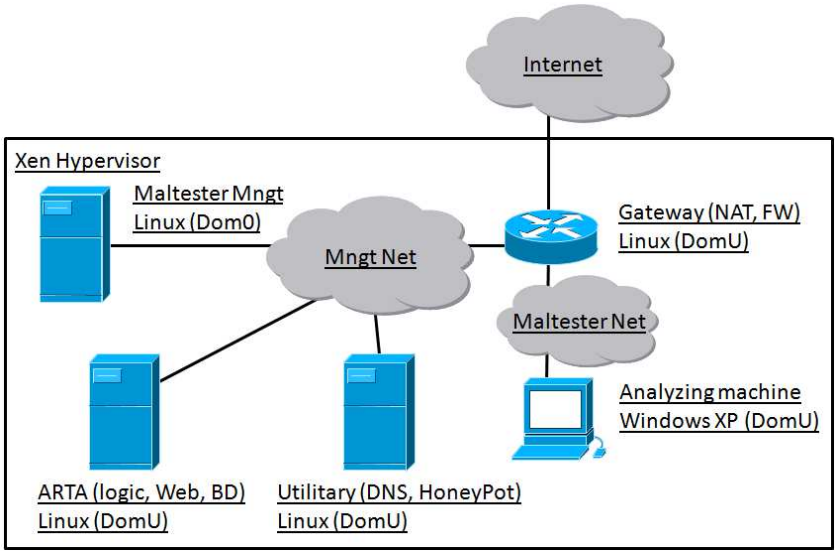

Figure 2. Network infrastructure used by the ARTA system

4. First results

The ARTA system was developed and deployed due to the increase in number of malicious samples makes manual analysis almost impossible. As soon as prototype was implemented our analysis are supported by it and first feedbacks from analyst introduced some project changes and improvements. Following paragraph contains some numbers concerning performed automatic analysis.

Using ARTA system 332 samples of CryptoWall 3.0 ransomware from February 2015 to the middle of November 2015 are analyzed. Samples are gathered from freely available sources [7, 8]. Analysis reveals usage of 1397 unique URLs in 1356 distinct domains. Forty five distinct CryptoWall proxy list are detected. Table 1 presents automatically gathered information concerning 27 detected proxy lists, which are used by at least five analyzed samples. What should be emphasized these data is directly accessed from ARTA Web interface (compare Figure 1).

The usage of Web interface which enables interactive access speed up further analysis which are based on automatically gathered data from dynamic analysis. Using ARTA system and analyst knowledge some new, before unknown facts concerning CryptoWall 3.0 activity was discovered. The first behavior is associated with re-usage of CryptoWall proxies in various lists. Research shows that in most cases addresses is used only once. From 1397 URLs 1119 URLs are used only once and only 31 are used more than twice. Additional investigation shows that in

most cases these reused addresses are associated with servers that probably have poor management and are not cleaned for longer period of time.

Table 1. Automatically gathered from ARTA system information concerning CryptoWall 3.0 proxy lists (list name, number of URLs contained in list and number of samples which uses this list)

List name	URL-s no.	Samples no.	List name	URL-s no.	Samples no.
cryptowall542	37	25	cryptowall565	32	9
cryptowall41	25	22	cryptowall10	43	8
cryptowall28	51	19	cryptowall25	53	8
cryptowall511	39	16	cryptowall06	33	7
cryptowall08	67	15	cryptowall79	37	7
cryptowall527	29	15	cryptowall48	43	6
cryptowall17	41	14	cryptowall72	43	6
cryptowall101	58	13	cryptowall85	49	6
cryptowall03	36	12	cryptowall89	58	6
cryptowall22	41	12	cryptowall15	40	5
cryptowall47	36	12	cryptowall57	40	5
cryptowall04	29	10	cryptowall579	28	5
cryptowall601	41	10	cryptowall596	23	5
cryptowall02	24	9			

The second discovered behavior shows some changes in attackers' tactic. From the beginning of this threat, the activity in compromised server used later as CryptoWall 3.0 proxy hostile script was placed only once. But at the end of October 2015 we observed that the same server hosts more than one hostile script. Currently, this activity concerns minority of URLs but a new trend was observed. In most cases one CryptoWall proxy list contains few doubled URLs, however one of observed domain have hostile script in five distinct places. This trend was detected by one of analyst during periodic manual scan of recently discovered proxy lists.

5. Conclusions

The ARTA system was developed due to limitations of pure manual analysis caused by increase number of malware samples. Usage of the ARTA prototype shows system usefulness. The main advantage from usage of the ARTA system

was a decrease in time spends by the analysts for sample analysis. Previously utilized manual analysis takes about ten minutes per sample, using ARTA system analysis is performed automatically and takes no longer than three minutes. The second big advantage of the ARTA system was associated with Web interface which ease access to full analysis results. Usage of interactive views which are connected one to other using hyperlinks is very beneficial. The analyst can change one view to another associated with presented data just after one click, for example, from samples list he or she can go directly to detailed information concerning interesting one. Shortly after deployment of the ARTA prototype it become a knowledge base of all analyzed to date samples and theirs features. Search functionality enabled in the ARTA system allows easy access to interesting information. For example, when some pieces of information concerning CryptoWall 3.0 activity are founded in the Internet, a quick search for sample hash or used domains gives clues if this sample was analyzed by our security team. The last observation from works with ARTA systems shows that this kind of system is a good approach to automation of analyst work. It is planned to be improved and enhanced in the future.

REFERENCES

[1] McAffe Labs, Threats Report, May 2015, URL: www.mcafee.com/us/resources/reports/rp-quarterly-threat-q1-2015.pdf

[2] Symantec, Internet Threat Report, April 2015, URL: https://www4.symantec.com/mktginfo/whitepaper/ISTR/21347932_GA-internet-security-threat-report-volume-20-2015-social_v2.pdf

[3] Kharraz, A., Robertson W., Balzarotti, D., Bilge L.; Kirda E. (2015) *Cutting the Gordian Knot: A Look Under the Hood of. Ransomware*, 12th Conference on Detection of Intrusions and Malware & Vulnerability Assessment, Milan, Italy

[4] Cabaj K., Gawkowski P., Grochowski K., Osojca D. (2015) *Network activity analysis of CryptoWall ransomware*, Przegląd Elektrotechniczny, Vol 91, No 11

[5] Cabaj K., Denis M., Buda M. (2013) *Management and Analytical Software for Data Gathered from HoneyPot System, Information Systems in Management*, WULS Press Warsaw, vol. 2, nr 3, 182-193

[6] Cabaj K., Gawkowski P. (2015) *HoneyPot systems in practice*, Przegląd Elektrotechniczny, Vol 91, No 2

[7] Malwr.com, https://malwr.com/ (last access November 2015)

[8] Reverse.It, http://www.reverse.it/ (last access November 2015)