Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

2entFOX: A Framework for High Survivable Ransomwares Detection

Mohammad Mehdi Ahmadian , Hamid Reza Shahriari
Department of Computer Engineering and Information Technology
Amirkabir University of technology
Tehran, Iran
{mm.Ahmadian, Shahriari}@aut.ac.ir

***Abstract*— Ransomwares have become a growing threat since 2012, and the situation continues to worsen until now. The lack of security mechanisms and security awareness are pushing the systems into mire of ransomware attacks. In this paper, a new framework called 2entFOX[1] is proposed in order to detect high survivable ransomwares (HSR). To our knowledge this framework can be considered as one of the first frameworks in ransomware detection because of little publicly-available research in this field. We analyzed Windows ransomwares' behaviour and we tried to find appropriate features which are particular useful in detecting this type of malwares with high detection accuracy and low false positive rate. After hard experimental analysis we extracted 20 effective features which due to two highly efficient ones we could achieve an appropriate set for HSRs detection. After proposing architecture based on Bayesian belief network, the final evaluation is done on some known ransomware samples and unknown ones based on six different scenarios. The result of this evaluations shows the high accuracy of 2entFox in detection of HSRs.**

Keywords—component; ransomware; malware detection; malware analysis; behavioral detection; high survivable ransomware ;

I. INTRODUCTION

In general, a ransomware is a type of malware that after infecting a computer system, restricts the access to the system or its resources, and demands a ransom paid to the creator(s) of the malware for the restriction to be removed [3]. Considering the economic incentives, it can be concluded that the diversity and complexity of ransomwares have been increased significantly. Dangerous ransomwares are a relatively new form of malwares that have grown strongly at the end of 2012. On the other hand, it seemed that 2012 is the most important year for ransom, but this trend is continuing [3].

Based on our experiments what can strengthen anti-malware systems in order to detect and identify the ransomwares is using combination of static and dynamic analyze methods to extract behavioral information. As is emphasized in [4] and other malware detection resources, behavioral detection is one of the most effective methods for malware detection. Fig.1 contains reports released by McAfee [8-5], demonstrates ransomwares have grown since 2012. In 2013, ransomwares became a growing security challenge and the challenge became worse year by year.

Unfortunately, little research work has been done in this type of malware detection and no specific detection framework for ransomwares has been proposed in scholarly circles so far. Our previous work [3] and Kharraz's work [2] are two new and excellent research in this field. In [3], we proposed just one feature for HSR detection which leads to low detection accuracy and high false positive rates. This paper aims to provide a framework for high-precision detection uses 19 other features and Bayesian belief network.

The strength of [2] is taking advantage of the Lastline labs facilities which lead to automatic analysis of 1359 ransomware samples. Reference [2] just introduces the features without any implementation of detection system and any evaluation. Reference [2] found only limited features which has caused too much emphasis on abnormal access feature and high false positive rate in implementation which will be explained in section 2.

[1] This framework called 2entFOX because twenty chosen features are proposed in the first version of it with 2 highly efficient ones ("2ent "implies that). "F" implies Feature or Framework. "O" is one of the ransomware's characters. "X" implies that the number of chosen effective features compulsorily are not fixed to twenty and according to the various conditions these features can be increased or decreased (X is a variable). Totally this framework with the help of twenty foxes can fox (confuse or perplex) high survivable ransomwares and detects them with high accuracy.

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

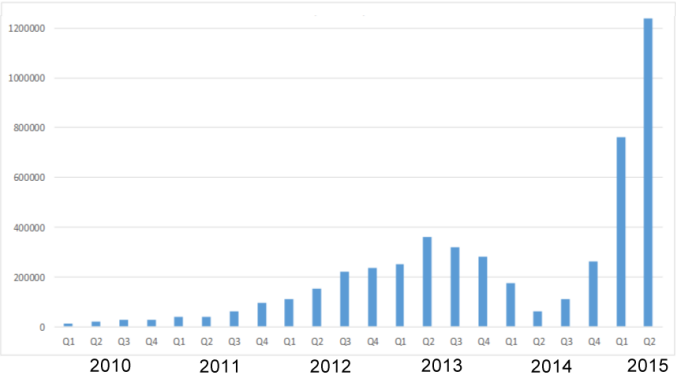

Fig. 1. New Ransomware Samples

In the second section, the 2entFOX (framework architecture, feature extraction and Bayesian network) will be introduced. In the third section, 2entFOX evaluation will be explained. And finally in fourth section conclusions about the achievements will be discussed.

II. Proposed Framework

In this research project based on the taxonomy presented in [3], after considering the general nature of ransomware, static and dynamic analysis on various ransomwares (especially static and dynamic behavior analysis), studying the most of the malware analysis and detection methods, we faced the challenge that detection of all ransomwares is outside the scope of this project (due to their inherent diversity and overlapping some features with other malware).

According to the explanations given in [3] and the importance of HSRs, 2entFOX targets the most dangerous and most destructive ransomwares means HSRs. Meanwhile, the detection of the low survivable ransomwares (LSR) is not so important. Although we can provide a system for all ransomwares detection, but this system would have some challenges. The most significant challenge is for detecting all type of ransomwares we should change the top feature weights in conditional probability tables (CPT) that causes some weaknesses:

- It increases false negative and false positive rates in detecting HSRs.
- It increases false positive rate with detecting other type of malwares which is one of our goal to detect none of them in 2entFOX.
- It increases static and run-time overhead for collecting feature values and making decision when the number of features become too many.

First of all, the studies and experimental analysis were done on the 20 suitable available ransomware samples and many other type of malwares and benign software. 2entFOX architecture was designed according to malware detection systems. Based on the architecture portrayed in Fig.2 Many various features identified and extracted, then 20 desired enough features were chosen among of them.

After that the final detection engine contains CPTs in Bayesian network based on analysis of 7 HSR samples calculated. In the learning phase for CPTs calculation, the analysis information of four benign software and three other type of malware samples were also used.

2entFOX is similar to the specification-based detection system but with a difference; in 2entFOX, the targeted HSRs' illegal behaviors as twenty general features were set and based on this knowledge the detection engine has been designed in the learning phase. During the test phase, any program that has high compliance with these features and relationship between them will be known as HSR.

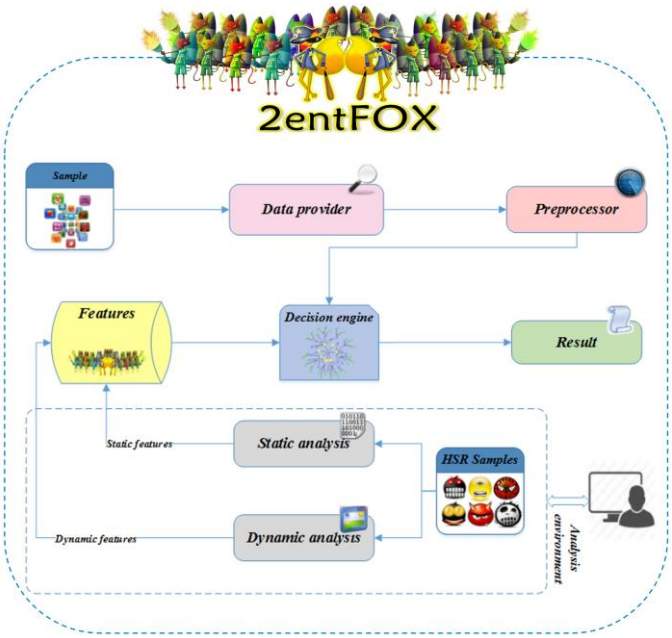

Fig. 2. The 2entFOX architecture

As will be explained in third section this method has a low false positive rate, but extracting and choosing a complete and accurate set of features with relation dependencies among them was time-consuming and cumbersome. On the other hand, since the behavior of the programs comes out as some features this method has some similarities with signature-based methods. Since the behavior of these HSRs is extracted as general form of features this detection system is capable in detection of both known and unknown HSRs unlike signature-based methods.

It should be noted that the extraction of accurate features of HSR is not easy and the most time in this research is dedicated to the best feature extraction. In the surface although some features may seem simple but extracting sufficient features, from hordes of behavioral and nonbehavioral features, is a very challenging work and is the 2entFOX strengths. At the

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

detection time 2entFOX provides the input program (under the supervision of malware expert) into preprocessor and delivers it to the detection engine. Detection engine, according to the extracted features bank and set threshold, pulls out its response regarding detection or no detection of HSR.

A. *Extracted features*

In this section, the twenty extracted features for HSR detection is presented, meanwhile brief explanation of some cases, their full list is given in table 1 and a detailed explanation is available in [1]. The key-exchange step feature is the first one which is explained in details in [3].After numerous analyzes on different ransomware samples and specially HSRs we found that HSRs such as Cryptolocker and Cryptowall are trying to remove VSS[2] data while running that the victim system loses its ability to recover data after data encryption. This feature due to its unique usage in ransomwares rather than other type of malwares so far. Based on our taken reviews have not been considered in any formal document as an extracted feature for ransomware detection systems neither in the commercial and the nor in academic sections. In This project, after extraction of this feature has been used as the most important feature for HSR detection.

HSRs to achieve their sinister goals are trying to eliminate the most important Windows mechanism for storing the users' hot backup files, make changes in the service, that the victim has no way to recover its encrypted data except paying ransom. One of the behaviors that we encountered in analyzing ransomwares was getting access to special directories. Of course, this feature can be seen among the majority of malware because generally access to these directories does not require special credential to store information. We believe that although this feature does not worth a lot but along with other features in 2entFOX can play a good role. One of the other behaviors that we faced in analyzing the ransomwares was accessing to the particular registry paths. Another feature which we have dealt with while analyzing the CGR[3] ransomware [3] is accessing some of the ransomwares to cryptographic library.

TABLE I. EXTRACTED SELECTED FEATURES

#	Feature name	Analysis type	Detection stage	variable value
1	Access to cryptographic libraries	Dynamic	Triggering	Yes/No
2	Decryption help file key words	Static	On disk	Yes/No
3	Decryption help file key words	Dynamic	Dormant	Yes/No
4	Decryption help file key words	Dynamic	Execution	Yes/No
5	Targeted files search key words	Static	On disk	Yes/No
6	Targeted files search key words	Dynamic	Dormant	Yes/No
7	VSS	Static	On disk	Yes/No
8	VSS	Dynamic	Dormant	Yes/No
9	VSS	Dynamic	Triggering	Yes/No
10	Specific registry paths key words	Static	On disk	Yes/No
11	Specific registry paths key words	Dynamic	Dormant	Yes/No
12	Access to specific registry paths	Dynamic	Triggering	Yes/No
13	Access to specific registry paths	Dynamic	Execution	Yes/No
14	Specific directories access key words	Static	On disk	Yes/No
15	Specific directories access key words	Dynamic	Dormant	Yes/No
16	Access to specific directories	Dynamic	Dormant	Yes/No
17	decryption help file content key words	Static	On disk	Yes/No
18	decryption help file content key words	Dynamic	Dormant	Yes/No
19	Abnormal access to the paths and files	Dynamic	Execution	Yes/No
20	Key-exchange step feature	Dynamic	Triggering	Yes/No

B. *Data providing*

This part of the framework, according to the introduced features class is responsible for providing various data corresponding to 2entFOX features. In This section, with the help of tools the process of monitoring of evaluated program and its static and dynamic behavior analysis takes place under the expert supervisor; desired data is provided with the feature recognition stage model is sketched in Fig.3 for HSRs.

C. *Preprocessing*

This section of 2entFOX is designed with goal of preprocessing the obtained data from various extracted features classes that are have high volume. By eliminating additional data, a reduction in unrelated features attempts to enhance speed, generality and simplicity in detection engine. However, the chosen data for elimination by the preprocessor is specified by the expert knowledge of ransomware analyst.

D. *Detection engine*

Extracted features in the detection process have different value and weight. On the one hand it is possible to face with incomplete information or probabilities for some of the characteristics according to involved challenges including the intractable problem of malware detection [9]. Therefore, in detection of ransomwares using these approaches the decision-making system is faced with the problem of uncertainty. However, to solve this problem and providing a useful decision system a probabilistic approach is needed; Bayesian networks

[2] Volume Shadow copy Services
[3] Cryptographic Ransomware

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

is a perfect choice for a detection engine with these requirements.

Bayesian network consists of two components [10]:

1) A directed graph which shows the variables dependencies in set S.

2) Set P which includes the set of A_i probabilities which is called CPT. In CPT there are possible combination of different parents' conditions. Parents of a child node, which is denoted by the symbol I, is a set of nodes that has an edge to children in the graph. As a result, the probability of A is calculated by (1)

$$P(A) = \prod_{i=1}^{n} p(A_i | \pi_i) \qquad (1)$$

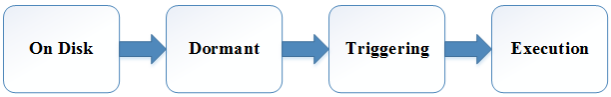

Fig. 3. Model of HSRs feature recognition stage

According to the provided explanations, Bayesian network model is appropriate to identify the malicious nature of a HSR which is a prediction problem based on existing evidence. In 2entFOX twenty chosen features in HSRs detection are planned as the Bayesian network inputs to detect that the evaluated sample destructiveness. Fig.4 illustrates the graphical model of Bayesian network -the decision engine- based on the number column is shown in table 1.

This framework helps us to evaluate and determine which sample can be a HSR. Although we suggest these twenty features for HSR detection but it is not obligated to only use these twenty features for HSR detection in 2netFOX; the feature set can be decreased or increased according to the security countermeasures, safeguards and HSRs evaluations, but every module of 2entFOX framework can be used in other driven systems.

III. Evaluation

To evaluate the proposed approach, six different scenarios are used [1], which only one scenario is expressed in this paper. The first scenario was considered in this project was after extracting each of the features independently and step by step; the value of each feature should be examined to determine the effect of each feature in HSRs detection. The result of this approach was multiple experiments and many technical reports which lead to a growing number of extracted features. As we said, after experimental review on features and dependability relationships among them we have chosen twenty features. however, the principle of minimum features for optimal performance in execution is also considered. for 2entFOX implementation in Windows operating system we have used various tools and modules. At first to have an appropriate evaluation each of the extracted features were examined on the available ransomware samples. Furthermore, in the form of the second scenario, all twenty features which are tabulated in the Table 1 were evaluated in interconnected class.

Despite numerous attempts for HSR sample gathering, because of some access limitations according to political boycott we could collect twenty suitable known samples. The process of feature evaluation is done using these twenty samples. Moreover for 2netFOX evaluation against new and unknown ransomwares, a new ransomware set produced by Tox Virus[4] have been used. Meanwhile to determine the detection rate and the least false positive, a set of benign software - which have the most similar features with some features and we call them "cross-border software"- and number of other types of malware have been used. The final evaluation results can be seen in table 2 (appendix). 2entFOX with presented samples has been able to detect every seven HSRs.

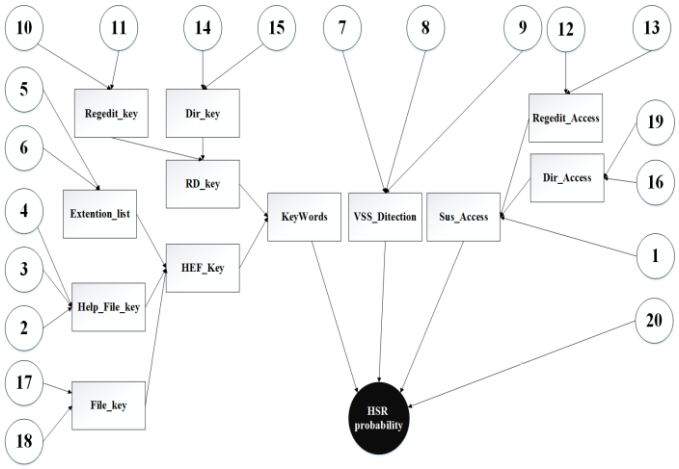

Fig. 4. Graphical Bayesian network model of decision System

According to the obtained statistics from the HSRs, cross-border software and other types of malware (including botnets that uses the key-exchange step feature with the other aim) we suggest the threshold of 85 for 2entFOX. The reason for selecting this threshold is based on a logical test: although to our knowledge the probability that a benign software has both feature classes number one and two is impossible, but if there is such a benign software sample with only these two features

[4] This service was found by McAfee labs as the first ransomware as a service platform which is free dependent on invisible web by TOR

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

2netfox will reach the approximate probability of 84.9. Therefore, we believe that the threshold of 85 is a good option.

It is worthy to note that 2entFOX does not detect cross-border software which are similar in some features with HSR by mistake. We could use the non- cross-border software that have the least common features with HSR but it is clear that such this software is the least probability of detection and take into account very little calculated probability. The amount of recall (R), precision (P) and F-Measure of this evaluation is calculated in (2, 3 and 4).

$$R = \frac{TP}{TP + FN} \times 100 = \frac{7}{7+0} \times 100 = 100 \quad (2)$$

$$P = \frac{TP}{TP + FP} \times 100 = \frac{7}{7+1} \times 100 = 87.5 \quad (3)$$

$$F_{measure} = \frac{2}{\frac{1}{P} + \frac{1}{R}} = 93.33\% \quad (4)$$

As calculated in (2), based on evaluated samples the recall amount of detection system is 100%. It can be seen in (3) that the precision of the approach according to a detection of a LSR among the samples was 87.5%; Thus the amount of F-Measure based on (4) is 93.33%. Although according to this equation, the precision is reduced, but the point is that, 2entFOX in such cases will be able to detect some LSRs; This problem cannot be considered as a disadvantages of the system, but it can be considered as the strength of the system.

Extraction and selection suitable features, selection of appropriate detection target, Bayesian network structural design, appropriate probability distribution for nodes and selection of an appropriate threshold are the reasons of 2entFOX precision and suitable recall. However, it can never be generally claimed that 2entFOX has always this amount of F-Measure; because it would require multiple reviews and evaluations on the numerous samples which unfortunately due to limitations on the number of HSR samples it was not possible for us.

Table 3 (appendix) compares the capabilities of existing ransomware detection products and 2entFOX; Unlike available products to detect ransomwares, 2entFOX has been very powerful in detecting unknown HSRs. Evaluation on unknown ransomwares which were produced by Tox Virus, shows the high detection rate of 2entFox; The reason of this high detection rate is that no HSR is able to bypass all of the extracted features with the help of defensive methods. To our knowledge if HSRs can bypass all of these features, they will not be able to bypass the VSS feature which is monitoring the behavior in the run-time mechanisms.

IV. Conclusion

In this paper we proposed a framework for high survivable ransomwares detection based on twenty appropriate features. In 2entFOX, after providing data and preprocessor step we designed a detection system with the help of Bayesian belief network to use extracted features and their statistical possibilities. 2entFOX have some advantage and disadvantages: The valuable point of 2entFOX is the ability of HSRs detections unlike other detection tools which is related to extractive features. This framework helps us to evaluate and determine which sample can be a HSR. Although we suggest these twenty features for HSR detection but it is not obligated to only use these twenty features for HSR detection in 2netFOX; the feature set can be decreased or increased according to the security countermeasures, safeguards and HSRs evaluations, but every module of 2entFOX framework can be used in other driven systems.

2entFOX has been very powerful in new HSRs detection; Evaluations on new samples produced by Tox Virus show the high detection rate of the proposed framework. The reason of this high detection rate is that no HSR is able to bypass all of the extracted features with the help of defensive methods; To our knowledge if HSRs can bypass all of these features, they will not be able to bypass the VSS feature which is monitoring the behavior in the run-time mechanisms. Significant limitations of 2entFOX, is its weakness in LSRs detection, which in general does not have some heavyweight extracted features of HSRs. LSRs are not considered as a critical cyber threat for computer users due to the weakness in their system; without paying ransom and only with the help of malware analysis or reverse engineering methods also with the help of decrypted tools, encrypted data can be decrypted [3, 11, 12].

References

[1] Ahmadian Mohammad Mehdi, A Framework for Ransomware Detection Based on Behavioral Analysis (M.Sc. thesis), Dept. of Computer Eng. and IT, Amirkabir University of Technology (Tehran Polytechnic), 2015, pp.117-203.

[2] Kharraz, Amin, William Robertson, Davide Balzarotti, Leyla Bilge, and Engin Kirda. Cutting the gordian knot: a look under the hood of ransomware attacks. Detection of Intrusions and Malware, and Vulnerability Assessment. Springer International Publishing, 2015. 3-24.

[3] Ahmadian M M, Shahriari H R and Ghaffarian S M, "Connection-monitor & connection-breaker: A novel approach for prevention and detection of high survivable ransomwares". Proceedings of 12th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Rasht, (2015) September 8-10 DOI: 10.1109/ISCISC.2015.7387902.

[4] Greoigre Jacob, Herve Debar, Eric Fillol, Behavioral detection of malware: from a survey towards an established taxonomy, Springer-Verlag France 2008.

[5] McAfee Threats Report: Third Quarter 2013, By McAfee Labs, Page 19, 2013.

[6] McAfee Threats Report: Second Quarter 2014, By McAfee Labs, Page 21,2014

[7] McAfee Threats Report: February 2015, By McAfee Labs, Page 38, 2015.

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

[8] McAfee Threats Report: August 2015, By McAfee Labs, Page 35, 2015.

[9] Adleman, Leonard M. An abstract theory of computer viruses. Proceedings on Advances in cryptology. Springer-Verlag New York, Inc., 1990.

[10] Koller, Daphne, and Nir Friedman. Probabilistic graphical models: principles and techniques. MIT press, 2009.

[11] Gazet, Alexandre. Comparative analysis of various ransomware virii. Journal in computer virology 6.1 -2010: 77-90.

[12] Young, Adam, and Moti Yung. Malicious cryptography: Exposing cryptovirology. John Wiley & Sons, 2004.

TABLE II. THE 2ENTFOX EVALUATION RESULTS ON VARIOUS SAMPLES.

#	Software name	Software type	Detection (based on feature classes in table 4)							Calculated probability	Final result (detection)
			1	2	3	4	5	6	7		
1	Cryptolocker	Known HSR	✓	✓	✓	✓	✓	✓	✓	98.5	✓
2	Cryptowall 2	Known HSR	✓	✓	✓	✓	✓	✓	✓	98.4	✓
3	Cryptowall 3	Known HSR	✓	✓	✓	✓	✓	✓	✓	97.7	✓
4	TorrentLocker	Known HSR	✓	✓	✓	✓	✓	×	✓	97.0	✓
5	CryptoFortress	Known LSR	×	✓	✓	✓	✓	✓	✓	96.3	✓
6	ToxV_sample1	Unknown HSR	✓	✓	✓	✓	✓	✓	✓	96.9	✓
7	ToxV_sample2	Unknown HSR	✓	✓	✓	✓	✓	×	✓	97.0	✓
8	ToxV_sample3	Unknown HSR	✓	✓	✓	✓	✓	✓	✓	96.9	✓
9	DropBox	Benign software	×	×	×	×	✓	×	×	6.13	×
10	Adobe Reader	Benign software	×	×	×	×	×	✓	×	7.85	×
11	IDM	Benign software	×	×	×	×	✓	×	×	6.2	×
12	Restore Point Creator	Benign software	×	✓	×	×	✓	×	×	75.1	×
13	Total Commander	Benign software	×	×	✓	×	✓	×	×	18.2	×
14	DellTouchPad	Benign software	×	×	×	×	✓	×	×	6.0	×
15	Flash Renamer	Benign software	×	×	✓	×	×	×	×	10.0	×
16	TweakNow	Benign software	×	×	✓	×	×	×	×	10.1	×
17	Torpig	Malware(botnet)	✓	×	×	✓	×	×	×	50.3	×
18	Cbeplay	Malware(botnet)	✓	×	×	✓	×	×	×	44.9	×
19	Sality	Malware(worm)	×	×	×	✓	✓	✓	×	9.66	×
20	Gamarue	Malware(worm)	×	×	×	✓	✓	×	×	8.1	×
21	Beta_bot	Malware(Trojan)	×	×	×	✓	✓	×	×	7.85	×
22	Viko.W32.Autorun	Malware(virus)	×	×	×	✓	✓	×	×	7.7	×

TABLE III. COMPARISON BETWEEN EXISTING RANSOMWARE DETECTION PRODUCTS

Product/Framework	Known LSR detection	Known HSR detection	Unknown LSR detection	Unknown HSR detection	Detection Approach
HitmanPro kickstart	✓	✓	×	×	signature-based
BitDefender AntiCryptoWall	✓	✓	×	×	signature-

Proceedings of 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), IEEE, Iran, Tehran, (2016) September 7-8,Ahmadian M M and Shahriari H R.(Link) ©This article is published in author personal webpage under the terms of the Creative Commons Attribution Non-Commercial License. DOI: 10.1109/ISCISC.2016.7736455.

					based
HitmanPro CryptoGuard	✓	✓	×	×	signature-based
Padvish CryptoProtect	×	✓ (but only 4 families)	×	×	signature-based
2entFOX	× (but can detect some families)	✓	× (but can detect some families)	✓	behavioral Specification-based

TABLE IV. FEATURE CLASSES.

Name	Number
Key-exchange step feature	1
VSS (all related features)	2
Abnormal access to the paths and files	3
Access to specific directories	4
Access to specific registry paths (all related features)	5
Access to cryptographic libraries	6
key words (all related features)	7