Global Journals LaTeX JournalKaleidoscope™
Artificial Intelligence formulated this projection for compatibility purposes from the original article published at Global Journals. However, this technology is currently in beta. *Therefore, kindly ignore odd layouts, missed formulae, text, tables, or figures.*

Review of Contemporary Literature on Machine Learning based Malware Analysis and Detection Strategies

G. Bala Krishna[1], V. Radha[2] and K. Venugopala Rao[3]

[1] KMIT/JNTUH

Received: 12 December 2015 Accepted: 4 January 2016 Published: 15 January 2016

Abstract

Abstract: malicious software also known as malware are the critical security threat experienced by the current ear of internet and computer system users. The malwares can morph to access or control the system level operations in multiple dimensions. The traditional malware detection strategies detects by signatures, which are not capable to notify the unknown malwares. The machine learning models learns from the behavioral patterns of the existing malwares and attempts to notify the malwares with similar behavioral patterns, hence these strategies often succeeds to notify even about unknown malwares. This manuscript explored the detailed review of machine learning based malware detection strategies found in contemporary literature.

Index terms— malware detection, malware signature, API call sequence, anomalies, static analysis, dynamic analysis, machine learning.

1 I. introduction

he term "Malware" stands for malicious software, and it usually specifies as hostile software application. According to G. Mc Graw et al., [1] there are multiple causes as code added, changed, or removed from the software it get corrupt and it deliberately causes harm and disrupt normal computing activity. A virus had a broad range of destructive software applications such as viruses, Trojans, Spywares and other intrusive code [2].

The malware can discriminate by the capability of replication, propagation, self-execution and corruption of the operating system. If the computer system gets extortion it influence on confidential information, integrity and denial of assistance. In malware Replication is a crucial component as it assure its existence.

In some cases Replication generates consumption and continuation of system resources (e. g. hard disk, RAM). If confidential assets are being used by any other malware types other than the user, to conceal themselves from anti-malware detector they use a technique called polymorphic or metamorphic techniques.

The operating system gets corrupt through data transfer from desecrate device to another protected device familiar, such as executable files, boot records of disk drives or exhausting network bandwidth, by using local or network files system. In such case malware makes operating system susceptibility and few software bugs are faults and it starts its life cycle at the same system and infected system simultaneously by remotely control.

According to a McAfee simplified report (year 2013) says that "malware continues to grow" [3] and by G Data and king [4] [5] soft Laboratory declare nnumber of innovative malware will emerge promptly and to build an anti-malware the analyzers and constructors are enhanced by their unique techniques and methods [6]- [10]. To construct a malicious software the techniques which are been used to categorized and estimate in groups such as obfuscation techniques, invocation methods, platform, spreading and propagation techniques.

To actuate a program has a malicious attentive or not, malware detection system is used. In this detection system there are two different functions, detection and analysis [12]. Detection system is a protecting one as it may or may not be prevail in the same system [13] and the tasks can be split into client and server as it analogous in cloud-based antivirus [8,12]. A numerous renovations are made on detection and analysis functions [5], [12], [15]- [19].

In malware detection system specialized solutions are added to expansion in success and achievement. Such as cloud computing [10], network based detection system [20], web, virtual machine [21], [22], agent technology [23]- [29] or by the use of hybrid methods and technologies.

2 II.

3 Review of Contemporary Literature and Contributions

In earlier stage malware had come up with signature based detection. But now in this stage malware signature has introduced an automatic generation and it is pretended to be important and it increased its pattern in similar speed.

The signature based detection system has some imperfection as follows to continue the updates of T signature it requires high maintenance cost. By inclusion such methods it could be evenly avoided by malware in polymorphic form [30]. To conquer the imperfection, it embraces code in normal vision to grab the consecrate original maliciousness. To vary the polymorphic techniques and apply, this grabbed malicious code is used but still it is anemic to detecting obfuscated malware. Apart from this some execution paths can be explored execution **??**31] [30].

Due to certain requirements the malware analysis is all ways conserved the techniques in the prior, consequently dynamic analysis was considered. To identify and execute a complicate malware dynamic analysis methods are used. In dynamic analysis the malware shows how it operates and recognize the unknown malware which is identically operates like a known malware [32]. There are two familiar primary dynamic methods are control flow analysis and API call analysis. **??**33] [34].

API call data display how the malware gets operates and it can be obtained by both static and dynamic approaches. The API list and PE format of the executable files can be derived by the static approach [35] [36] [37] [38]. In dynamic approach. [39] [40] [41], **??**42] [43] [44] API calls can be recognized by running executable files it usually run in virtual machine.

In API call there two familiar ways to evaluate the data accumulated by static approach. The first one implements simple statistical analysis, for example, to count the frequency of API call which is aspect to organize malware [35]. The second approach is to gather the API call data through data mining or machine learning techniques. In another way the API call sequence data which gathered by the dynamic approach are helpful to creates a behavioral patterns. The information accumulated by the dynamic approach also operates simple statistics such as frequency counting [39] and data mining or machine learning [40] [42] [44].

The API call analysis been done with API call approaches. In this abstraction the dynamic method is applied to excerpt API call sequences. To obtain austerity patterns, DNA sequence alignment algorithms (MSA and LCS) are adapted. With API call sequence patterns and the critical API call sequence, we can recognize the unknown malware or its variation with elevated efficiency.

Anderson et al. [49] defined a malware detection strategy that builds a set of graphs from the given instruction set and then analyzes these graphs to notify the proneness of the malware activity. In order to build the graphs the markov chains were defined on 2gram sequences. The graphs defined form the training set further used to build a similarity matrix using graph kernel. The graph kernel is the mix of Gaussian and spectral kernels, which are in use to assess the similarity between graph edges and similarity between graphs respectively. Further the support vector machine that learns from the similarities between graph edges and graphs is used to classify the input call sequences.

By using such liberal malware software the multiple kernel is achieved and learning design used in this work to exhibit selective refined differences occurrence of malware. The inadequacy of this approach is computed consequence is very high, hence the use of this approach is discouraged.

Bayer et al. [50] prospect a technique that groupsthe call sequences generated by Anubis [51]. The behavior adequately of the call sequences is considered as objective to cluster the call sequences by Locality Sensitivity Hashing (LSH) **??**52]. The constraint of the model is that LSH is capable to generate probabilistic clusters.

Biley et al. [53] argued that malware prototyping is not consistent among the notable antivirus products available. In order to this the authors devised a novel classification strategy that classifies the malware according to the changes observed at system state. A strategy that prototypes the behavior of the malware is used, further the malwares are classified according to these behavior prototypes. The distance between a class and a malware is assessed by the distance metric called "normalized compression distance (NCD)". The constraint observed in empirical study of this model is that the behavior prototype definition is static and limited to malwares that are not fall in zero-day category (unknown malwares). park et al. [54] defined a classification strategy that classifies malware based on the graphs generated from the call sequences. Further graph similarities between confirmed malware call sequence graph and unknown call sequence graph will be assessed. The similarity index is the "max number of subgraphs identified in both graphs". The malwares those controls the system privileges without initiating the system call sequences are not traceable by this classification model, which is a significant constraint of this model. Firdausi et al. [55]

4 E

In other way, researchers are analyzed more ways to develop API call sequence information. In earlier research API call had introduced API call graph [45] with various kinds of call graph analysis. To get more consequential features for call graph analysis, the analyzer had espoused the mechanism of social network analysis. [46] According to analyzers the affinity among API call sequences is based on cosine similarity function and lengthy jaccard measure. Due to modern research [33] [34] [47] [48] more information had been added such as control flow information and API argument information to inflate the efficiency in the mining process. by the model called Anubis [51]. Further these observed behavioral patterns will be organized as sparse vectors and learns the behavior prototypes. The malware samples given for testing will be classified, which is based on the behavioral prototypes learned in training phase. The performance of the model is estimated through benchmark classifiers and they are "j48", "multilayer perception neural networks (MLP)" "Naïve Bayes", "Support Vector Machine (SVM)" and "k-Nearest Neighbors (kNN)". The experimental results indicating that the J48 classification delivered much classification accuracy.

Nari et al. [56] devised a network flow behavioral analysis framework for malware detection. The network transactions obtained from PCAP files were considered to extract the network flows. Further a network activity representation graph is drawn from these network flows. The given network flows labeled as malware were used in training phase. Further this framework learns representation of the features such as size of the graph, average, maximum and root level outdegree and count of specific nodes of the network activity graphs of the given input network flows. Further these features specific information uses to classify the input malware samples in testing phase. In order to perform the classification, the WEKA library [57] was used. The experimental study indicating that the J48 is the best classifier among all classifiers available in WEKA library.

Lee et al. [58] explored a machine learning based malwares clustering. The training phase builds the behavioral profiles of the malware samples given as training data and the profile includes the system resources invoked by the system calls and their arguments. Further the similarities between behavioral profiles were considered distance function to cluster the malwares, which was done by k-medoids. The outliers are adjusted to the clusters based on the nearest neighbor strategy. This approach is the combination of static and dynamic clustering strategy that clusters known features by k-medoid and unknown and new features by nearest neighbor approach. This strategy is evincing that hybrid approach is more robust in order to classify the known as well as unknown features effectively.

Another hybrid approach for malware detection was devised by Santos et al. [59]. This approach tracks the known features (static features) through the analysis of the sequence of operational codes in given malicious executable and the unknown features (dynamic features) were noticed from the observation of exceptions and operations in system calls. The experimental study was done under various classifiers and results obtained were evincing the significant accuracy in malware classification Islam et al. [60] explored a similar strategy that extracts static and dynamic features to classify the executables into malevolent or benevolent. The features such as function length, function executable frequency and length of the strings involved are included in known features and the features such as function identity and function arguments are included in unknown features. The experiments were done using the classifiers called Support Vector Machine, Decision Tree and Random Forest and results evincing that the random forest is the best classifier among all considered.

The malware classification method devised by Anderson et al. [61] is using the divergent input sources such as control flow graphs, static call sequences, portioned executables, dynamic call sequences and file signatures. Further this model learns the weight of these input combinations from the given training set of malevolent and benevolent executables. The observed weights of these input combinations are used further to classify the malevolent and benevolent executables during testing phase. The process overhead is the significant constraint of this model observed against dense and high speed network streams.

5 III.

6 Conclusion

The current era of internet and computer systems are prone to serious security threats due to the malicious software which are also referred as malware. Hence the significant research contributions aimed to define malware detection and prevention strategies in contemporary literature. All of these contributions are fall in the categories of either anomaly based, signature based or call sequence analysis based detection. The signature based models are capable to notify and prevent the malwares that are notified earlier. In contrast to this the anomaly models and call sequence analysis models are capable to identify the malwares based on the similarities learned from previous malware attacks. The difference between the anomaly and call sequence analysis models is that the anomaly based learning models can adopt user defined features, whereas the call sequence analysis models notify the similarities learned from the call sequences of 2-gram, 3-gram or ngram. This manuscript aimed to affirm the objectives and limits of the contributions found in recent literature. The conclusion of the review evincing that the machine learning based models that learns from either anomalies or call sequence are tolerable the constraints observed in signature based malware detection strategies. The anomaly and call sequence learning models found in contemporary literature are not adequate to defend the challenges evincing from the vibrant

163 and unjust network data. Hence the significant contributions are in demand to handle the challenges evinced in current era of internet and computer system usage. [1] [2]

Figure 1:

164

[1] © 2016 Global Journals Inc. (US) 1
[2] © 2016 Global Journals Inc. (US)

[Annual Workshop on Cyber Security and Information Intelligence Research] , *Annual Workshop on Cyber Security and Information Intelligence Research* (45) .

[] , 10.1016/j.jnca.2012.10.004. http://dx.doi.org/10.1016/j.jnca.2012.10.004

[Vinod ()] , P Vinod . *Survey on Malware Detection Methods* 2009.

[Berkenkopf and Malware ()] , R B S Berkenkopf , G-Data Malware . 2010. (Report)

[Mcafee ()] , Lab Mcafee . 2013 Threats Predictions. 2013.

[Zhou ()] *A Heuristic Approach for Detection of Obfuscated Malware*, S T A M Zhou . 2009. IEEE.

[Das ()] *A Temporal Logic Based Approach to Multi-Agent Intrusion Detection and Prevention*, Paritosh Das , RN . 2012.

[Garfinkel and Rosenblum ()] *A virtual machine introspection based architecture for intrusion detection*, T Garfinkel , M Rosenblum . 2003. p. .

[Ou and Ou ()] 'Agent-Based immunity for computer virus: abstraction from dendritic cell algorithm with danger theory'. C. -M Ou , C R Ou . *Proceedings of the 5th international conference on Advances in Grid and Pervasive Computing*, (the 5th international conference on Advances in Grid and Pervasive ComputingHualien, Taiwan) 2010. Springer-Verlag. p. .

[Ye ()] *An Agent-Based Framework for Distributed Intrusion Detections*, D Ye . 2009.

[Firdausi et al. (2010)] 'Analysis of Machine Learning Techniques Used in Behavior Based Malware Detection'. I Firdausi , C Lim , A Erwin . *Proceedings of 2nd International Conference on Advances in Computing, Control and Telecommunication Technologies (ACT)*, (2nd International Conference on Advances in Computing, Control and Telecommunication Technologies (ACT)) 2010. Jakarta, 2-3 December 2010. p. .

[Qiao et al. (2013)] 'Analyzing malware by abstracting the frequent item sets in API call sequences'. Y Qiao , Y Yang , L Ji , He . *Proceedings of the 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '13)*, (the 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '13)) July 2013. p. .

[Indyk and Motwani (1998)] 'Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality'. P Indyk , R Motwani . http://anubis.iseclab.org/52 *Proceedings of 30th Annual ACM Symposium on Theory of Computing*, (30th Annual ACM Symposium on Theory of ComputingDallas) 1998. May 1998. p. . (Anubis)

[Mcgraw and Morrisett ()] *Attacking Malicious Code: A Report to the Infosec Research Council*, G Mcgraw , G Morrisett . 2000. IEEE Softw. 17 p. .

[Biley et al. ()] 'Automated Classification and Analysis of Internet Malware'. M Biley , J Oberheid , J Andersen , Z Morley Mao , F Jahanian , J Nazario . 10.1007/978-3-540-74320-0_10. http://dx.doi.org/10.1007/978-3-540-74320-0_10 *Proceedings of the 10 th International Conference on Recent Advances in Intrusion Detection*, (the 10 th International Conference on Recent Advances in Intrusion Detection) 2007. 4637 p. .

[Automated dynamic binary analysis ()] *Automated dynamic binary analysis*, 2007.

[Nari and Ghorbani (2013)] 'Automated Malware Classification Based on Network Behavior'. S Nari , A Ghorbani . *Proceedings of International Conference on Computing, Networking and Communications (ICNC)*, (International Conference on Computing, Networking and Communications (ICNC)San Diego) 2013. January 2013. p. .

[Rieck et al. ()] 'Automatic analysis of malware behavior using machine learning'. K Rieck , P Trinius , C Willems , T Holz . *Journal of Computer Security* 2011. 19 (4) p. .

[Qiao et al. ()] 'CBM: free, automatic malware analysis framework using API call sequences'. Y Qiao , Y Yang , J He , C Tang , Z Liu . *Knowledge Engineering and Management*, (Berlin, Germany) 2014. Springer. p. .

[Cesare and Xiang ()] S Cesare , Y Xiang . *Software Similarity and Classification*, 2012. Springer Science & Business Media.

[Islam et al. ()] 'Classification of Malware Based on Integrated Static and Dynamic Features'. R Islam , R Tian , L Battenb , S Versteeg . *Journal of Network and Computer Application* 2013. 36 p. .

[Abadi et al. (2005)] 'Control-´flow integrity'. M Abadi , M Budiu , U Erlingsson , J Ligatti . *Proceedings of the 12th ACM Conference on Computer and Communications Security*, (the 12th ACM Conference on Computer and Communications Security) November 2005. p. .

[Tian et al. (2010)] 'Differentiating malware from cleanware using behavioural analysis'. R Tian , M R Islam , L Batten , S Versteeg . *Proceedings of the 5th International Conference on Malicious and Unwanted Software (MALWARE '10)*, (the 5th International Conference on Malicious and Unwanted Software (MALWARE '10)Nancy, France) October 2010. p. .

[Kolbitsch ()] 'Effective and efficient malware detection at the end host'. C Kolbitsch . *Proceedings of the 18th conference on USENIX security symposium*, (the 18th conference on USENIX security symposiumMontreal, Canada) 2009. p. . (USENIX Association)

[Venugopal ()] 'Efficient signature based malware detection on mobile devices'. Deepak Venugopal , GH . *Mob. Inf. Syst* 2008. 4 (1) p. .

[Park et al. ()] 'Fast Malware Classification by Automated Behavioral Graph Matching'. Y Park , D Reeves , V Mulukutla , B Sundaravel . *Proceedings of the 6th*, (the 6th) 2010.

[Lagar-Cavilla ()] *Flexible Computing with Virtual Machines*, H A Lagar-Cavilla . 2009.

[Anderson et al. ()] 'Graph Based Malware Detection Using Dynamic Analysis'. B Anderson , D Quist , J Neil , C Storlie , T Lane . 10.1007/s11416-011-0152-x. http://dx.doi.org/10.1007/s11416-011-0152-x *Journal in Computer Virology* 2011. 7 p. .

[Ye et al. (2007)] 'IMDS: intelligent malware detection system'. Y Ye , D Wang , T Li , D Ye . *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining) August 2007. ACM. p. .

[Anderson et al. ()] 'Improving Malware Classification: Bridging the Static/Dynamic Gap'. B Anderson , C Storlie , T Lane . *Proceedings of 5 th ACM Workshop on Security and Artificial Intelligence (AISec)*, (5 th ACM Workshop on Security and Artificial Intelligence (AISec)) 2012. p. .

[Pinz ()] 'Improving the security level of the FUSION@ multi-agent architecture'. C I Pinz . *Expert Syst. Appl* 2012. 39 (8) p. .

[Ye ()] 'Intelligent file scoring system for malware detection from the gray list'. Y Ye . *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, (the 15th ACM SIGKDD international conference on Knowledge discovery and data miningParis, France) 2009. ACM. p. .

[Bijani and Robertson ()] 'Intrusion detection in open peer-to-peer multi-agent systems'. S Bijani , D Robertson . *Proceedings of the 5 th international conference on Autonomous infrastructure, management, and security: managing the dynamics of networks and services*, (the 5 th international conference on Autonomous infrastructure, management, and security: managing the dynamics of networks and servicesNancy, France) 2011. Springer Verlag. p. .

[Shankarapani et al. (2010)] 'Kernel machines for malware classification and similarity analysis'. M Shankarapani , K Kancherla , S Ramammoorthy , R Movva , S Mukkamala . *Proceedings of the International Joint Conference on Neural Networks (IJCNN '10)*, (the International Joint Conference on Neural Networks (IJCNN '10)) July 2010. p. .

[Lee and Mody ()] T Lee , J J Mody . *Behavioral Classification. Proceedings of the European Institute for Computer Antivirus Research Conference (EICAR'06)*, 2006.

[Moser et al. (2007)] 'Limits of static analysis for malware detection'. C Moser , E Kruegel , Kirda . *Proceedings of the 23rd Annual Computer Security Applications Conference (ACSAC '07)*, (the 23rd Annual Computer Security Applications Conference (ACSAC '07)) December 2007. p. .

[Rieck ()] *Malheur A novel tool for malware analysis*, K Rieck . 2012.

[Jang et al. ()] 'Malnetminer: malware classification based on social network analysis of call graph'. J. -W Jang , J Woo , J Yun , H K Kim . *Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion (WWW Companion '14)*, (the Companion Publication of the 23rd International Conference on World Wide Web Companion (WWW Companion '14)) 2014. p. . (International World Wide Web Conferences Steering Committee)

[Ahmed et al. ()] 'Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph'. Ammar Ahmed , E Elhadi , M A Maarof , A H Osman . *American Journal of Applied Sciences* 2012. 9 (3) p. .

[Sami et al. (2010)] 'Malware detection based on mining API calls'. B Sami , H Yadegari , N Rahimi , S Peiravian , A Hashemi , Hamze . *Proceedings of the 25th Annual ACM Symposium on Applied Computing (SAC '10)*, (the 25th Annual ACM Symposium on Applied Computing (SAC '10)) March 2010. ACM. p. .

[Shankarapani et al. ()] 'Malware detection using assembly and API call sequences'. M K Shankarapani , S Ramamoorthy , R S Movva , S Mukkamala . *Journal in Computer Virology* 2011. 7 (2) p. .

[Xufang et al. ()] 'Mechanisms of Polymorphic and Metamorphic Viruses'. L Xufang , P K K Loh , F Tan . *Intelligence and Security Informatics Conference (EISIC)*, 2011 European. 2011.

[Yanfang Ye et al. ()] 'MelihAbdulhayoglu, Combining file content and file relations for cloud based malware detection'. T L Yanfang Ye , Weiwei Shenghuo Zhu , Zhuang , Umesh Egementas , Gupta . *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, (the 17th ACM SIGKDD international conference on Knowledge discovery and data miningSan Diego, California, USA) 2011. ACM. p. .

[Kevadia Kaushal and Prajapati ()] 'Metamorphic Malware Detection Using Statistical Analysis'. P S Kevadia Kaushal , Nilesh Prajapati . *International Journal of Soft Computing and Engineering (IJSCE)* 2012. (2) .

[Gorodetsky ()] *Multi-agent Peer-to-Peer Intrusion Detection Computer Network Security*, V Gorodetsky . 2007. Berlin Heidelberg: Springer. p. .

[Ou ()] *Multiagent-based computer virus detection systems: abstraction from dendritic cell algorithm with danger theory*, C M Ou . 2011. Springerlink.

[Ahmed ()] 'NIDS: A Network Based Approach to Intrusion Detection and Prevention'. M Ahmed . *Computer Science and Information Technology -Spring Conference*, 2009. 2009.

[Okane et al. ()] 'Obfuscation: the hidden malware'. P Okane , S Sezer , K Mclaughlin . *IEEE Security & Privacy* 2011. 9 (5) p. .

[Santos et al. ()] 'OPEM: A Static-Dynamic Approach for Machine Learning Based Malware Detection'. I Santos , J Devesa , F Brezo , J Nieves , P G Bringas . *Proceedings of International Conference CISIS'12-ICEUTE'12, Special Sessions Advances in Intelligent Systems and Computing*, (International Conference CISIS'12-ICEUTE'12, Special Sessions Advances in Intelligent Systems and Computing) 2013. 189 p. .

[Yin ()] 'Panorama: capturing systemwide information flow for malware detection and analysis'. H Yin . *Proceedings of the 14th ACM conference on Computer and communications security*, (the 14th ACM conference on Computer and communications securityAlexandria, Virginia, USA) 2007. ACM. p. .

[Linn et al. (2005)] 'Protecting against unexpected system calls'. C M Linn , M Rajagopalan , S Baker , C Collberg , S K Debray , J H Hartman . *Proceedings of the 14th USENIX Security Symposium*, (the 14th USENIX Security SymposiumBaltimore, Md, USA) August 2005. p. .

[Dong ()] 'Research on adaptive distributed intrusion detection system model based on Multi-Agent'. H Dong . *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, 2011.

[Bayer et al. ()] 'Scalable, Behavior-Based Malware Clustering'. U Bayer , P M Comparetti , C Hlauschek , C Kruegel . *Proceedings of the 16th Annual Network and Distributed System Security Symposium*, (the 16th Annual Network and Distributed System Security Symposium) 2009.

[Christodorescu ()] 'Semantics-Aware Malware Detection'. M Christodorescu . *Proceedings of the 2005 IEEE Symposium on Security and Privacy*, (the 2005 IEEE Symposium on Security and Privacy) 2005. IEEE Computer Society. p. .

[Sathyanarayan et al. ()] 'Signature generation and detection of malware families'. S Sathyanarayan , P Kohli , B Bruhadeshwar . *Information Security and Privacy*, (Berlin, Germany) 2008. Springer.

[Bergeron et al. ()] 'Static detection of malicious code in executable programs'. J Bergeron , M Debbabi , J Desharnais , M M Erhioui , Y Lavoie , N Tawbi . *Proceedings of the Symposium on Requirements Engineering for Information Security (SREIS '01)*, (the Symposium on Requirements Engineering for Information Security (SREIS '01)) 2001.

[Jiang et al. ()] *Stealthy malware detection through vmm-based "outof-the-box" semantic view reconstruction, in Computer and communications security*, X Jiang , X Wang , D Xu . 2007. Alexandria, Virginia, USA: ACM. p. .

[Rajagopalan et al. ()] 'System call monitoring using authenticated system calls'. M Rajagopalan , M A Hiltunen , T Jim , R D Schlichting . *IEEE Transactions on Dependable and Secure Computing* 2006. 3 (3) p. .

[Hall et al. ()] 'The WEKA Data Mining Software: An Update'. M Hall , E Frank , G Holmes , B Pfahringer , P Reutemann , I Witten . *ACM SIGKDD Explorations Newsletter* 2009. p. .

[Ahmed et al. (2009)] 'Using spatiotemporal linformationin API call swithmachinelea rning algorithms for malware detection'. F Ahmed , H Hameed , M Z Shafiq , M Farooq . *Proceedings of the 2nd ACM Workshop on Security and Artificial Intelligence*, (the 2nd ACM Workshop on Security and Artificial Intelligence) November 2009. p. .

[Zeltser] *what is cloud Anti-Virus and how it does work*, L Zeltser .

[Alazab et al. (2011)] 'Zeroday malware detection based on supervised learning algorithms of API call signatures'. M Alazab , S Venkatraman , P Watters . *Proceedings of the 9 th Australasian Data Mining Conference (AusDM '11)*, (the 9 th Australasian Data Mining Conference (AusDM '11)) December 2011. Australian Computer Society. 121 p. .