International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-7, Issue-5, May 2021
ISSN: 2395-3470
www.ijseas.com

A Comprehensive Study on Malware Detection Techniques Using Machine Learning

Vivekanand Kuriyal[1], Dibyahash Bordoloi[2], D.P.Singh[3], Vikas Tripathi[4]
1 2 3 4 Graphic Era Deemed to be University, Dehradun, India

Abstract
In the last decade, there has been a proliferation in the growth of malware. As a result, cyber-crime has emerged as a major issue for computer users. Cyber-attacks are orchestrated in many ways and malware attack is one of them. Sometimes called a malicious program. Malware includes viruses, Trojan horses, worms, rootkits, adwares and ransomware. These are the major weapons of the cyber attackers. Malware detection is a very tedious task for an organization, it has to deploy different techniques to secure its network and system. Malicious programs are traditionally detected by signature-based detection techniques and behavior (dynamic) based detection techniques. Signature based techniques are fast, but they are not able to detect new malicious programs. Hence researchers are more focused on behavior-based techniques. A malicious file is monitored in a virtual environment and certain tools. In this environment a malware sample is tested at run time and its features are captured for analyzing. These features are used for detection of malicious code in the Operating system. In this paper we have made a critical review of static and dynamic techniques and a framework is created for extracting dynamic features.
Keywords: *Malware detection, Machine learning, Malicious file.*

1. Introduction

The Malware and malicious codes are used interchangeably in cyber security. It infects an operating system and its critical services, these code gives power to hackers to erase data, access passwords or even dump bank account numbers. The first virus was detected in 1970. Known as kreeper warm, it was a self-replicated program for experimental purpose. It replicates itself to a remote PC and display a massage "I'm the creeper." After that boot sector virus was discovered at the end of 1980 and this was the beginning of malware. Since then, malware production has not stopped. Malwares are of different types like Trojans, Worms, Viruses, Rootkits, Bots, Spyware, Adware, and Ransomware[1]. These malwares pose a big problem to the regular Internet user. Sometime malwares are also causes internet security issues like spam mails and

Denial of Service attacks. Malware attacks in a computer system occurs in many ways, for example DOS attack, Probing, R2U, R2L virus, port scans, buffer overflow, CGI Attack and flooding etc. We need a platform where a system can be developed for recognition and prevention of these attacks. Researchers use different techniques for malware detection and classification. Machine learning (Deep learning) is one of the most advanced techniques being used [2].

1.1 Taxonomy of malware

Worms: Worms do not rely on any other file to replicate itself. In order to spread itself, computer worms use removable media and computer networks. Some time it is spreads via e-mails and it sends back all contacts from address book to the attacker. Worms are a type of malicious type of software used to damage the system and to gain hidden control over the system.

Viruses: Viruses are computer programs that are executed in the victim's computer that insert their own code and modify other program. Virus need some file or medium to transfer from one system to other. Virus can also delete any file from the computer system.

Trojan horse: Main purpose to create this type of malware is to perform a usual action is a system but at the same time another code used for malicious purposes. Initially it is looks like a safe application. The malicious parts are hidden to the user. Trojan horses perform a number of harmful activities like monitoring network traffic, stealing data, installing other viruses and malicious softwares. It can also perform backdoor function to hijack system for cyber-crime.

Adware: Adware is related to advertisements displayed on an application or website. After clicking on such an advertisement, it can change browser's setting without permission or can install any mischievous program that can harm the system.

Spyware: It is a hidden malicious software that is installed in the user's computer and user is not aware about its presence. That hidden software can record all user activities and all key strokes and send back the information to the creator of spyware or third party. Keylogger is an example of spyware.

Ransomware: It is used to block the access to the operating system or any particular file. To unlock the hacker asks for money or ransom payment.

Rootkit: This type of malware works at the kernel level. It hides presence of malware after modifying API calls in the system and then send system information to the hacker. This types of software resides in the master boot record or system firmware.

Backdoor: it is an unauthorized way of accessing a system. Hacker gets access into the victim's system for stealing data or some confidential information. This takes place without the user's knowledge. Certain points of entries that are unsecured are used to gain access.

1.2. Cyber Attacking techniques

Vulnerable: It is a device or an software's flaw that attacker uses to insert malicious program into the operating system, vulnerability may be in the form of run time error, designing error or some other form of system error. Therefore protection from such attacks users need to update their system on time to time.

Backdoor: Attacker use backdoor to upload malicious code into the system. Such backdoors left open because of systems bug. Sometime backdoor create by Trojan horse. Such malware do not harm the system but they open the point to enter other viruses to infect the system.

Removable Drives: Drive like flash drive, hard drives are the common way to spread the virus among PC and network. it can spread all type of malware like Trojan horse, warm, ransomware etc.

Homogeneity: This is a special type of spreading method in which malicious program get propagated among itself due to networking of PC. Each PC is connected through another PC using network and they must have same type of operating system[3].

2. Ease of Use

2.1 Contributions of this paper

- This literature review presents the evolution of malware detection techniques and available detection techniques till now.
- This literature review consists malware detection techniques, and detailed work done till now.
- Various Malware detection and classification algorithms are described and compared using different feature.
- Recently Static, Dynamic and hybrid techniques are compared with their advantage and disadvantage.

2.2 Organization

This literature review is organized into sections. Section 1 is an introduction into the types of malware. Section 2 describes difference malware analysis tools. Section 3 is a comparison of different machine learning algorithms which are used for malware detection. In Section 4 a survey of static and dynamic approaches is given. In this section malicious files features are compared with machine learning algorithm in a tabular form. Finally, the future scope and conclusion are given.

3. Malware analysis

Malware analysis is used for feature extraction, and these features are used for detection of various malwares. Now days mainly two types of technique are used Signature based (static analysis) and behaviour based (Dynamic analysis), these are described below.

3.1 Static malware analysis

Static analysis is used for malware analysis and extracting features without executing it. After analysis various features are extracted such as N-grams, PE Header, string, hash value etc., and used in the antivirus or IDS. In the static analysis the malware is examined at code level, it is disassembled into assembly language , for this task disassemblers are used[3].

3.2 Dynamic malware analysis

In this approach features are extracted at run time, examples of such features are Memory and register usage, Instruction traces, Network traffic, API call traces and registry changes. It is also called behavior based approach. To study the malware we need a virtual environment for analysis because of security reasons. A virtual environment is created using virtual box or VMware[3]. When a malicious file is executed in it this virtual environment the activities of malware like registry changes, file creation, URL access and downloading file etc. Are observed.

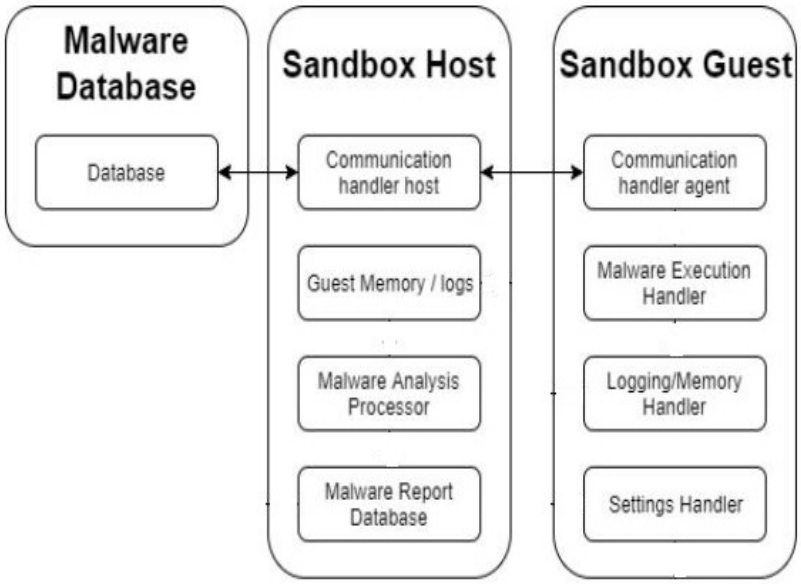

Fig1.Components of Virtual environment for malware detection

3.3. Dynamic analysis environments

Cuckoo sandbox is an open source system that supports all kind of operating systems, but runs very smoothly on Linux (Ubuntu) operating systems. It analysis all window based executable file in a virtual environment and generate a report of its behavior .Researchers use sandbox to monitor how malware operates in victims operating system, which helps the researcher to design a defense system against malicious threats. Sometime malware developer use evading techniques from sandbox but cuckoo sandbox has the ability to disable such defiance at run time and bring out real intention of a malicious program in front of researcher. Figure 2 shows the architecture of cuckoo sandbox .It has one host machine and multiple guest machines, guest machines are installed in a virtual environment for example virtualBox and VMware. Connectivity between host and guest machine is through virtual networking. After installing all necessary supporting software, a Web user interface is generated to input any malicious file for testing. This file is distributed among all guest machines and when analysis is completed report is generated and agent sent this report back to the host machine in JSON format. For virtual environment different type of software are used like VMware, VirtualBox and Hyper-V. Figure 1 show the component of virtual sandbox environment.

4. Machine Learning for malware detection [3].

Machine Learning (ML) is becoming a very useful tool for researchers in the field of Malware detection and analysis, now days lot of research work has been done for classification malware and benign files using ML. ML provides more accurate modelling using more features of a malicious file to decide whether a file is malware or not[4]. To train a model first features are extracted and the features are selected and then suitable classification algorithms are applied for trainingg the malware classifier. It is very difficult to train and update malware classifier frequently, because many malware is created almost every day.To detect new malware, classifier has to be updated and this is a very time consuming task and also expensive.

4.1 Static analysis tools.

PeView: it provides information about header of any executable file. PE (Portable header) consist of the header information.

PEid: obfuscated malware is detected using this tool, because some actual code is packed in other software.

CFF Explorer: This tool gives header information of executable file and displays metadata about the executable file.

PsFile: This tools is used to check whether a file is opened by the remote computer or not. It analyzes whether the local computer is being operated remotely.

Accesschk: this tool provides information regarding access registry for read, write and execute operations. It provide security information at the user level.
Radare: Reverse engineering is done using Radare tools. It can be used in Linux, Window and MAC OS.
Yara: This tool is used for matching malware signature in a malicious file. String matching can be used in PDF, DOC, etc. using this tool.
SSDeep: This tool is used for determining variants of a malware sample. It calculates fuzzy values for analysis.
Disassemblers: IDA pro and Ghidra is used for reverse engineering. The file contain are converted into assembly code in this process and then further used for analysis purposes.

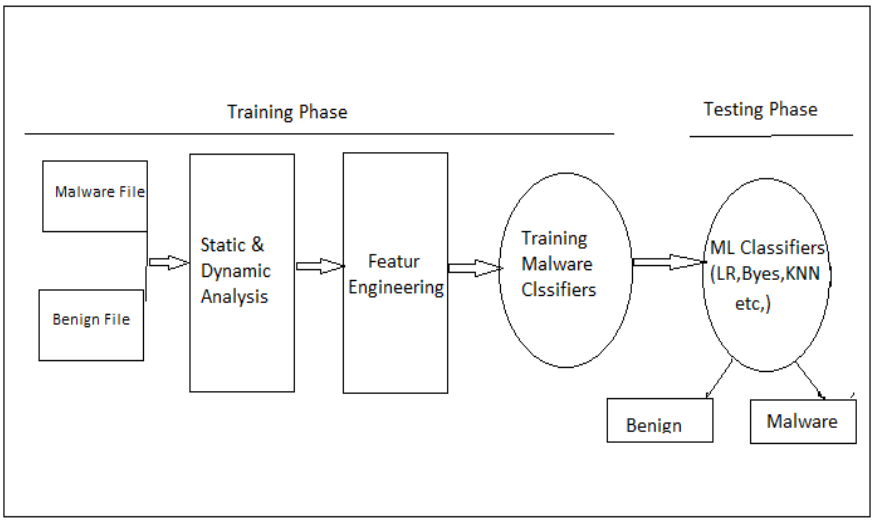

Fig 2. Process of malicious file detection

It extract feature of malware automatically.
• This techniques is better to detect new malware.
• It reduce efforts and time for analyzing malware sample

4.2 Dynamic analysis tools.

Process Explorer It is used to get details of all running process and application in the system, just like task manager.
ProcMon it captures system call accessed by a process or application like file operation, registry, memory operation and network related task. This tool is also called (ProcMon) process monitor.
Wireshark, and Tshark these tools are used to capture incoming and outgoing network packet. Sometimes malwares are designed in such a way that they can access URL and download information automatically, this tool is very handy for analyzing such malwares.
TCPdump Network traffic is analysed by this tool, it can also analyse TCP/IP packet
TCPview This tool help to analyze UDP and TCP at the different end-points.
Regshot Registry changes are analyzed by this tool. Before and after executing a malicious file this tool monitor all changes in registry.
Memoryze it is used for forensics purpose, it uses all memory dump after executing a malware sample.
Volatility It is a framework used for analyzing memory dump. Python programing language is used in this framework. It is used to extract features for the memory dump like injected library, network connection, and registry.
inetsim This tool is used to simulate virtual internet for a malware, for example if a malware requires connectivity to a remote computer then this tool provides virtual internet so that malware can work farther .
FakeDNS this tool is used to reply malware's DNS related query so that malware behaviour can be analyzed.
Sandboxes It is a virtual environment used to evaluate a malware, consisting of most of the above tools for various operations. Sandbox instances include .Cuckoo, Panda, Parsa, Anubis, etc.

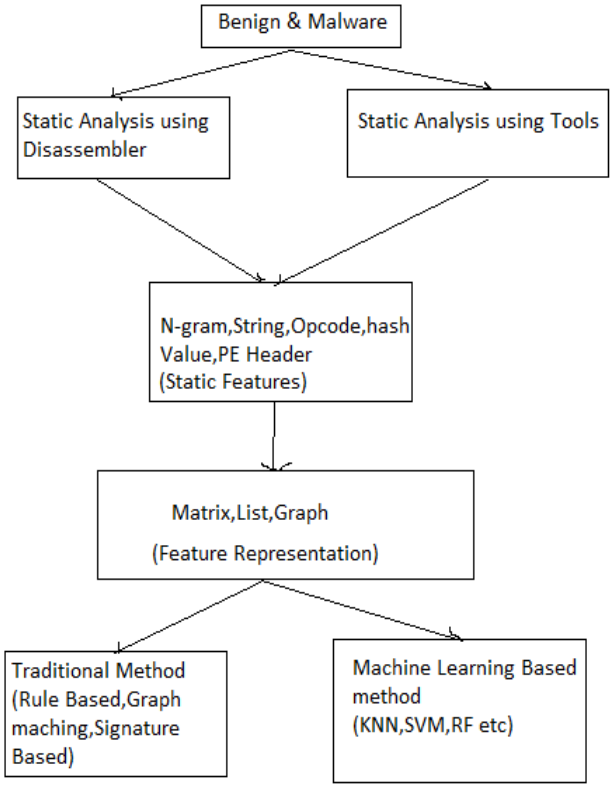

Fig 3 Static Malware analysis

4.3 Difference between Static and Dynamic Analysis Techniques.

For static analysis of malicious file execution of the malicious file is not required, but in Dynamic analysis Malicious file must be executed. Advantage of Static analysis is fast processing and advantage of Dynamic techniques is that new malware can be detected.

5. Malware detection techniques

Mostly three type of malware detection techniques are used: static, dynamic and hybrid. Main purpose of using these techniques is to protect PC from malicious programs that can harm files and services of theoperating system. Research mainly concentrate on extracting the features using static, dynamic and hybrid methods and then process this data to understand the behavior of a malicious file. Processed data is arranged in a proper format and utilized it to train Machine Learning models. Feature extraction is very important for model accuracy. Malicious files features are Opcodes, Byte sequence, String, PE Header, APIs calls etc. If extracted features are appropriate then false positive rate will decrease automatically.

5.1 Signature-based (Static feature) malware detection

In this method, malicious files are detected using signature of a file. It is very fast method to detect malware. Mostly signatures are generated using hash algorithm for example MD5, SHA1 etc. Generated signatures are stored in the dataset for training ML model. It is a very conventional method, in this techniques if a single byte is changed then complete signature will be changed. Some researchers have proposed other pattern such as control flow graph, mnemonic sequence etc.

5.1.1 Work done in Static feature (Signature-based) detection techniques

Under signature based approach various techniques have been developed. J.Saxe et al. Purpose a Malware detection using two dimensional binary program, Using static analysis, Contextual byte features, PE import feature and Score calibration model. Malware detection rate of this model is 95% and FP rate is 0.1 %. Static analysis on malware does not give satisfactory output for classification.[5].

A. Malhotra et al. present a text mining based approach for malicious file detection. DBScan algorithm is used for classification of malware. Signature based method is used for malware classification. This algorithm gives better results compared with other approach like IMDSS and MSPMD. It is a signature based approach so it cannot detect those malware having different signatures [6].

X. Meng et al. purpose a Classification Model for malware Detection Based on Deep Learning. MCSMGS model is use for Malware Classification and model Based on Static features. MCSMGS is a combination of static malware classification method (signature based) with deep learning methods. MCSMGS model is superior to the traditional malware classification method. SVM model give 94.7% accuracy which is lower than the MCSMGS model, it gives 98%.Dataset contain only PE file [7].

5.2. Behavior-based (Dynamic feature) malware detection

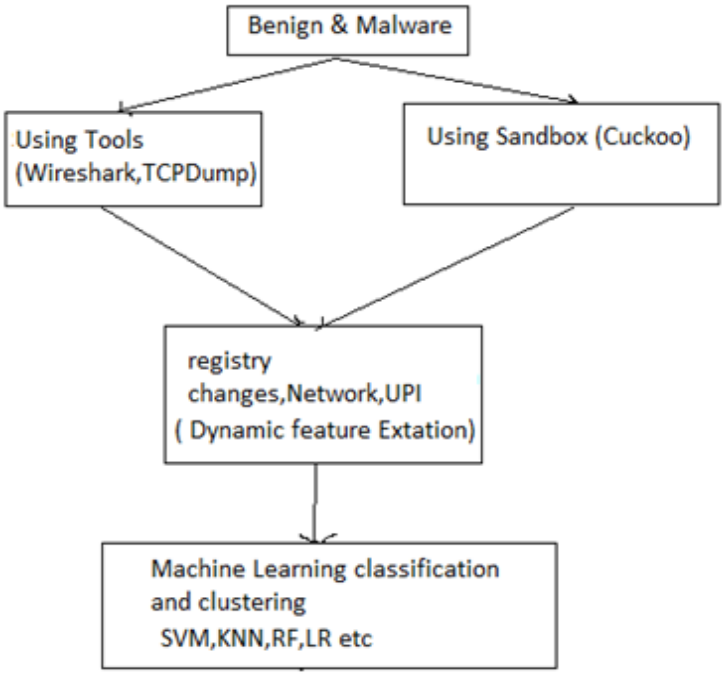

Fig 4. Dynamic malware detection

In this approach a malicious file is executed in a virtual environment and monitored its activity, during execution malware features are captured such as APIs, system event, system event, network events, browser events etc.

5.2.1 Work done in Dynamic feature (Behavioral based) malware detection techniques

M. Rhode et al. present a Malware prediction model, based on the behavioural (dynamic feature) data of an executable snapshot. This model use keras for RNN and scikitlearn to implement machine learning algorithm. Model present 94% accuracy for any executable. It takes 5 second of execution to predict whether a file is malicious or benign. Model can block windows process that is being used in OS [8].

E. Masabo et al. present Deep learning techniques. Dataset collection, Exploratory analysis, Preprocessing, Keras Deep learning and SVM algorithm is used. Deep learning algorithm achieved a better accuracy of 97% compared to 95% achieved by SVM Cannot classify malware and benign [9].

M. B. Bahador et al. present a Signature based approach using HLMD model for malware detection and classification. Solution: Hardware-level approach that uses behavioral signatures generated from performance counter traces to detect and classify malicious programs. An HLMD technique is compared with other machine Learning Techniques for malware classification, but HLMD gave best result. This techniques used only for LINUX [10].

6. Analysis of malware detection techniques and machine learning algorithm

In this section, a static and dynamic technique is presented in the tabular form, with their used machine learning techniques and malware feature for analysis. These techniques are used in windows environment for detection of malware. Table 1 show the detailed analysis of both techniques.

SA = Static Analysis
DA=Dynamic Analysis
RTF=Run Time Feature
RC=Register Contents
IM=binary file image representation
PEH=portable executable header
PSI=Printable String Information
MN=Mnemonic Frequency

7. Conclusion and future scope

In this literature review an analysis is made for a malicious file using static and dynamic tools and techniques which are used till now. A detail literature review is given in this paper how researcher used features of a malicious file for malware detection, Machine learning is one of main focused area in this paper and described many algorithm with static and dynamic techniques. In future hybrid technique can also be included.

Table 1: Malware Analysis on machine learning algorithm and malwares features.[3]

Authors	Types		ML Algorithm								Features						
	SA	DA	DT	SVM	KNN	NB	LR	ANN	RF	ADA	MN	PEH	AC	PSI	RC	IM	RTF
Ali et al. (2017)[11]		■		■						■			■				■
Ghafir et al.(2018)[12]		■		■	■				■	■							■
Shukla et al. (2019) [13]	■	■		■		■							■	■			
Huda et al. (2018) [14]		■						■			■						
Naz et al.(2019)[15]	■		■	■		■			■			■					
Alrzini et al. (2020) [16]		■			■							■					
Azeez et al.(2021)[17]		■	■			■		■	■			■					
Usman et al. (2021) [18]		■	■								■						■
Kumar et al. (2020) [19]		■							■					■			
Mao et al. (2017) [20]		■															■
Mohaisen et.al.(2015) [21]		■	■	■	■		■						■				■
Nagano and Uda (2017) [22]	■			■	■						■		■				
Narra et al.(2016) [23]	■				■						■						
Nauman et al. (2016) [24]	■	■											■				
Nayaranan et al (2016) [25]		■		■	■			■								■	
Pan et al. (2016) [26]		■						■					■				
Pfeffer et al. (2017) [27]	■	■											■				
Raff et al. (2016) [28]	■						■				■						
Searles et al. (2017) [29]	■			■													
Srndic et al. (2016) [30]	■			■									■	■			
Stiborek et al.(2018) [31]		■		■					■								■
Wagner et al.(2017) [32]		■			■								■				
iain j et al (2020)[33]		■			■							■					
S. Euh et al(2020) [34]		■		■					■				■				
H. Xue et al (2019)[35]	■	■	■		■						■						

References

1. D. Du, Y. Sun, Y. Ma , F. Xiao," A Novel Approach to Detect Malware Variants Based on Classified Behaviors", IEEE, 81770 – 81782 ,2019.
2. Uikey, R., Gyanchandani, M., "Survey on Classification Techniques Applied to Intrusion Detection System and its Comparative Analysis". International Conference on Communication and Electronics Systems,2019.
3. Singh, J., & Singh, J.,"A survey on machine learning-based malware detection in executable files". Journal of Systems Architecture, vol 112, 2020.
4. J. Zhang, "machine learning with feature selection using principal component analysis for malware detection: A case study", January 2019
5. J. Saxe, K. Berlin,"Deep Neural Network Based Malware Detection Using Two Dimensional Binary Program Features", International Conference on Malicious and Unwanted Software,IEEE,2015
6. A. Malhotra, K. Bajaj,"A hybrid pattern based text mining approach for malware detection Using DBScan", CSI Transactions on ICT, 4, 141–149, 2016
7. X. Meng, Z. Shan, F. Liu, B. Zhao, J. Han, H. Wang, J. Wang,"MCSMGS: Malware Classification Model Based on Deep Learning", International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC),IEEE,2017
8. M Rhode , P. Burnapb , K. Jones,"Early-stage malware prediction using recurrent neural networks", Computer Security, 77,578-594,2018.
9. E. Masabo, K. S. Kaawaase, J. Sansa-Otim," Big Data: Deep Learning for detecting Malware", Symposium on Software Engineering in Africa (SEiA), IEEE,2018.
10. M. B. Bahador,M. Abadi, A. Tajoddin," HLMD: a signature based approach to hardware level behavioral malware detection and classification", The Journal of Supercomputing 75,5551– 5582, 2019.
11. Mirza, Q. K. A., Awan, I., & Younas, M. , "CloudIntell: An intelligent malware detection system", Future Generation Computer Systems, 86, 1042-1053, 2018.
12. Ghafir, I., Hammoudeh, M., Prenosil, V., Han, L., Hegarty, R., Rabie, K. and Aparicio-Navarro, F.J., "Detection of advanced persistent threat using machine-learning correlation analysis", Future Generation Computer Systems, 89, pp.349-359, 2018.
13. Shukla, S., Kolhe, G., PD, S.M. and Rafatirad, S., 2019, November. Stealthy malware detection using rnn-based automated localized feature extraction and classifier. In 2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI) (pp. 590-597). IEEE.
14. Huda, S., Islam, R., Abawajy, J., Yearwood, J., Hassan, M.M. and Fortino, G., "A hybrid-multi filter-wrapper framework to identify run-time behaviour for fast malware detection", Future Generation Computer Systems, 83, pp.193-207,2018
15. Naz, S. and Singh, D.K., 2019, July. Review of machine learning methods for windows malware detection. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.
16. Alrzini, J. and Pennington, D., 2020, July. A review of polymorphic malware detection techniques. In International Conference on Interdisciplinary Computer Science and Engineering (ICICSE2020).
17. Azeez, N.A., Odufuwa, O.E., Misra, S., Oluranti, J. and Damaševičius, R., March. Windows PE Malware Detection Using Ensemble Learning. In Informatics (Vol. 8, No. 1, p. 10). Multidisciplinary Digital Publishing Institute,2021.
18. Usman, N., Usman, S., Khan, F., Jan, M.A., Sajid, A., Alazab, M. and Watters, P., "Intelligent Dynamic Malware Detection using Machine Learning in IP Reputation for Forensics Data Analytics". Future Generation Computer Systems, 118, pp.124-141,2021.
19. Kumar, A., Abhishek, K., Shah, K., Patel, D., Jain, Y., Chheda, H. and Nerurkar, P., 2020, November. Malware Detection Using Machine Learning. In Iberoamerican Knowledge Graphs and Semantic Web Conference (pp. 61-71). Springer, 2020
20. Mao, W., Cai, Z., Towsley, D., Feng, Q. and Guan, X., "Security importance assessment for system objects and malware detection", Computers & Security, 68, pp.47-68,2017

21. Mohaisen, A., Alrawi, O. and Mohaisen, M.,"AMAL: high-fidelity, behavior-based automated malware analysis and classification", computers & security, 52, pp.251-266.2015
22. Nagano, Y. and Uda, R.," Static analysis with paragraph vector for malware detection", In Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication (pp. 1-7),2017
23. Kolosnjaji, B., Zarras, A., Lengyel, T., Webster, G. and Eckert, C.,"Adaptive semantics-aware malware classification" In International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (pp. 419-439). Springer, 2016
24. Nauman, M., Azam, N. and Yao, J.,"A three-way decision making approach to malware analysis using probabilistic rough sets", Information Sciences, 374, pp.193-209,2016
25. Narayanan, B.N., Djaneye-Boundjou, O. and Kebede, T.M.,"Performance analysis of machine learning and pattern recognition algorithms for malware classification" In 2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS) (pp. 338-342). IEEE,2016
26. Pan, Z.P., Feng, C. and Tang, C.J.,"Malware classification based on the behavior analysis and back propagation neural network", In ITM Web of Conferences (Vol. 7, p. 02001). EDP Sciences,2016
27. Pfeffer, A., Ruttenberg, B., Kellogg, L., Howard, M., Call, C., O'Connor, A., Takata, G., Reilly, S.N., Patten, T., Taylor, J. and Hall, R.," Artificial intelligence based malware analysis", arXiv preprint arXiv:1704.08716,2017
28. Raff, E., Zak, R., Cox, R., Sylvester, J., Yacci, P., Ward, R., Tracy, A., McLean, M. and Nicholas, C., "An investigation of byte n-gram features for malware classification", Journal of Computer Virology and Hacking Techniques, 14(1), pp.1-20,2018
29. R. Searles et al., "Parallelization of Machine Learning Applied to Call Graphs of Binaries for Malware Detection," pp 69-77, 2017.
30. Šrndić, N. and Laskov, P., ,"Hidost: a static machine-learning-based detector of malicious files", EURASIP Journal on Information Security, pp.1-20,2016
31. Stiborek, J., Pevný, T. and Rehák, M., "Multiple instance learning for malware classification", Expert Systems with Applications, 93, pp.346-357,2018
32. Wagner, M., Rind, A., Thür, N. and Aigner, W., "A knowledge-assisted visual malware analysis system: Design, validation, and reflection of KAMAS", Computers & Security, 67, pp.1-15.2017
33. I. j. Cruickshank and k. m. Carley," Analysis of Malware Communities Using Multi-Modal Features",vol 8, pp 77435 – 77448,2020.
34. S. Euh , H. Lee , D. Kim , D. Hwang," Comparative Analysis of Low-Dimensional Features and Tree-Based Ensembles for Malware Detection Systems",vol 8, pp 76796 – 76808,2020.
35. H. Xue , S. Sun, G. venkataramani, T. lan," Machine Learning-Based Analysis of Program Binaries: A Comprehensive Study", vol 7,pp 2169-3536,2019.

AUTHORS PROFILE

Vivekanand kuriyal is a M.Tech student in Graphic era deemed to be University, Dehradun , India. His research interests include Cyber security using Machine learning. Now, he majors in detection and classification of malware.

Mr. Dibyahash Bordoloi obtained the M.Tech. Degree in the Computer Science from IIT Kharagpur,India. He is now working in the Graphic era Deemed to be university as an Associate professor. His research interests include Computer networking, Machine Learning, Data Mining and NLP.

Dr. Devesh Pratap Singh received the M.Tech. degree in computer science and engineering and the Ph.D. degree from Uttarakhand Technical University, Dehradun, India, in 2009 and 2015, respectively. He is currently a Professor and the Head of the Computer Science and Engineering Department, Graphic Era Deemed to be University, Dehradun. His research interests include information security, wireless sensor networks, the Internet of Things, and soft computing.

Dr. Vikas Tripathi has done BE in information technology from Technocrats institute of technology, Bhopal, M. Tech in Software engineering from Indian institute of information technology Gwalior and PhD from Uttarakhand technical university, Dehradun. He is actively involved in research related to Software engineering, Computer Vision, Machine learning and Video Analytics