Review On Spyware - A Malware Detection Using Datamining

[1]Mrs. Pushpa, [2]S.Santhiya,
[1]*HOD,Dept of MCA & SS , VLB Janakiammal College of Arts and Science ,Coimbatore.Tamil Nadu.India.*
[2]*Student of IV MSc (SS) VLB Janakiammal College of Arts and Science ,Coimbatore.Tamil Nadu.India.*

Abstract:

Nowadays computers are being threatened to unknown malware activities when it is connected to network. (i.e.) in system, malicious software is downloaded and installed without users knowledge. Malicious software is the one which have an intention to harm the system such as slow downing the network performance, increasing the electricity bill or to get confidential information including bank credit card number, login password details in order to misuse, to have a control over users by screen monitoring and key logger to track key or content typed by user. After gaining or processing, all such information is send to hacker by means of this software. One major type of malware is spyware. Spyware is one kind of malicious software act as spy and help hackers to steal personal information to misbehave. This paper presents the high definition , description of malware , its types and data mining concept which pave the way to detect and remove the malware mainly spyware with its tools and techniques.

I. INTRODUCTION

Malware is a concept used in malicious software. The term malware is a program which disrupts or distorts the system or computer. It's types include virus , Trojans , grayware , worm , rootkit , ransomware , keyloggers and botnet.

A. Virus:

It ruins the system through invading into it along with operating system in RAM in between the activity of boot up and shutdown of system. It is a file infector. It incredibly spread and corrupts the files during sharing of files or software among the computers. Normally it attaches to the files and exploit or delete the functionality. It appears like executable files at the time of run in software.

Macro virus, a type of virus hides into an application and can be removed by installing antivirus software. It avoids the virus getting downloaded from internet, e-mail attachment or USB drive. Even though antivirus software installed, regular update is needed in order to tackle upcoming or trending or unfamiliar virus.

B. Trojans:

It won't allow the user to predict the originality of undergoing action. It looks like legitimate software and it allows other malware to step in. It will not permit to access to required resources. It act as backdoor for cyber criminals to get user bank information. It is the most dreadful among all malwares.

C. Grayware:

It is an undesirable application or a program. Due to security faults it worsens the system performance. It is off two types:-

1. Adware:

It allows the advertisement to popup on the screen in web to perform malfunction,

When user moves the mouse over the screen unknowingly about this adware.(i.e.) after moving cursor slightly or nearer to the ad, the hacker gets permission to access resources of users system.

2. Spyware:

Spyware is a software can be downloaded by user not aware of it, when user in need to download other files. Spyware stick on to other files it cannot be seen by naked eye. At the time when needed file gets downloaded spyware also downloaded seamlessly. It monitors users internet browsing activities and send information to the person, who created the spyware. It gives space for adware.

Grayware can be detected by all antivirus software after processing each part as module and finally malware advertisement gets removed.

D. Worms:

Worm is a program it degrade the entire system (i.e.) it diminish the total files in operating system until , the system become free from file or empty. It increases network traffic. It propagates into many files easily if the system is distributed.

E. Rootkit:

It is a program or software tool unrevealed to users. It support malware and it let the way for hacker to perform admin role in users system such as clicking the mouse , opening another file making the user to be in a position of standstill. The only way to take part after attack of this is to reinstall the entire system (OS). It is automatically installed or installed by Intruder when he or she grabs the control of user as admin.

F. Ransomware:

Ransomware is software affects the system by underlying it until its motive gets fulfilled. Its motive can be in the form of fee. It also encrypt file, and make in accessible to resources and decrypts at end after high ransom is paid.

G. Key loggers:

It is software, monitors and registers the user type information in keyboard and sends it back to hackers. The information can be password, credit card details etc. But this is not applicable to virtual keyboard.

H. Botnet:

Botnet a word derived from "Robot and Network". It is a number of devices connected internally to a system. It's work is to carry out "Distributed Denial-Of-Service Attack (DDOS attack), steal data and send spam. We can remove this by installing antivirus software and making it to update automatically and being careful about what we are clicking, downloading and opening.

II. SPYWARE IN BRIEF

It gets installed to a system in 2 ways by hacker or by user.

By user, it is installed by accepting the End User License Agreement (EULA) (i.e.) when installing the needed application or file signing it to its terms and conditions in order to install application. This terms and condition can also be a spyware sometimes.

Hacker sends spyware, since hacker is a manufacturer of software or by through advertisement. When user clicks the advertisement, a massive changing gets happened in system (i.e.) a sudden violation can happen while clicking popup ad.

GUID (Global Unique Identifier) helps the hacker to know the browsing activities.

Using private network also, hackers can misuse user profile.

By activities of user, co-related activities are sending by hacker as an advertisement. If users view that ad, hacker gains some amount as a profit.

Hacker can send user activities to other corporate or small software companies for their business enhancement. By knowing user interest company can send more ad based on user interest so that the company can gain more profit if user view or select and purchase the product.

In Internet Explorer, hacker can track user with help of Browser Helper Object (HBO) interface or it depict as browser toolbar.

Majority of spyware penetrate into system by peer-to-peer client, downloaded with free games.

Scanner also plays an important role in detecting spyware, if this give a result as false-negative, this is very danger, since the vulnerabilities are not detected. If the result after scanning is false positive then it becomes over imagination of threat will be undergone.

The screen recorder spyware observe the screen changes and send information to hacker about the screen, viewed or modification done by user.

III. DATAMINING

Data mining is the activity of collecting and analyzing the data related data to make the data available in an understandable structure. In other words, we can say this as "Knowledge discovery in databases".

In simple term it is a process of KDD.

It involves six steps

A) ***Anomaly detection:*** Detecting error in data.
B) ***Association rule learning:*** Learning the dependency of variables.
C) ***Clustering:*** Similar data structures are grouped or clustered.

D) ***Classification:*** Differentiating the new format of data with old data structure or format.

E) ***Regression:*** Minute errors are to be finding out among relational data structure.

F) ***Summarization:*** Wrapping up of all related data to generate report.

Compared to other techniques, data mining is best suitable one to detect malware.

IV. SPYWARE DETECTION :

In later days, spyware detection done in 2 ways:

1) ***Signature based detection:***

This detection is done by analyzing the signature. If the pattern of the signature is predefined then it is detected, if it is new then spyware cannot be detected.

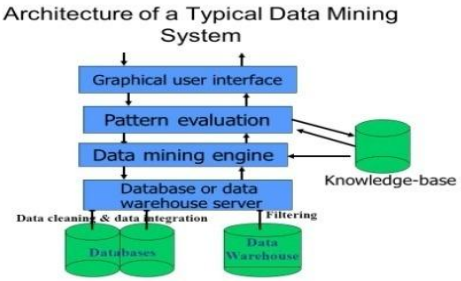

Fig1: Datamining process

2) ***Heuristic based detection:***

It is like a decision making process. With knowledge, the codes are applied to heuristic analyzer. Always it determines the file as false positive since the file is not spyware. This is the main drawback of it. Hence generic signature pattern arises. Even though it is proved to detect known spyware well, it is not applicable to detect unknown spyware. So further detection techniques are used like decision tree, depth search method, and Naïve bay's algorithm.

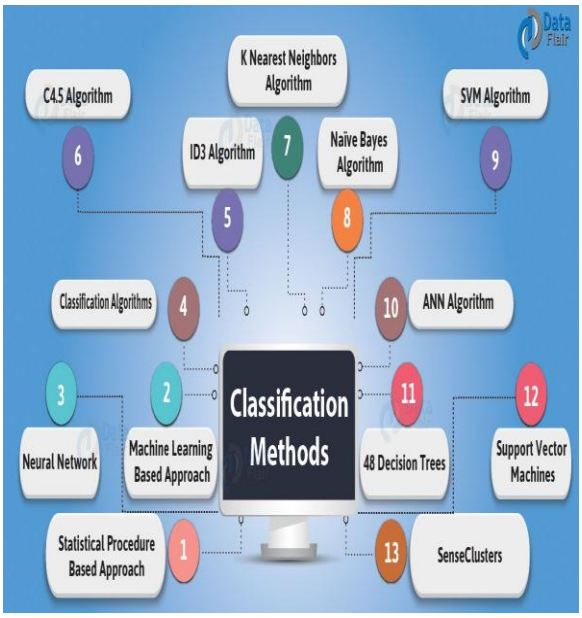

Fig2: Datamining Algorithm Classification to Detect Malware

V. CLASSIFIERS:

Classifiers are used to select attribute of data from known related dataset (i.e.) here data collection is made using attribute selection method. This helps to find spyware more accurately than signature based detection .These classifiers are to be made using datamining algorithm.

VI. CONCLUSION:

Data mining is the best suitable one to detect spyware, a malware and its types in contrast to other techniques.

In this paper we explained the malware , its types with new definition, description and detection of spyware , a malware in simple terms with more clarity in precise way. In future Spywares can be explained and implementation of detection of spyware may be done using different approach or techniques in datamining with help of a tool, not only detection removal method also be stated in distinct or in clear-cut way. It involves collection of a bigger dataset of Spyware and Benign files, implementation of various Classifier Learning Algorithms.

Motive: The motive of this paper is to make all people conscious about threats of computer and mobile. Since it is very important because computer and mobile are being a part of our daily life in emerging world. Making changes in these will change the entire human life in disguise manner. So it is life saving one to be aware of it.

REFERENCES

[1] Ravi, C & Manoharan, R. Malware Detection using Windows Api Sequence and Machine Learning . International Journal of Computer Application, Vol.43, No.17, 2012.

[2] Ravi, C & Chetia , G. Malware Threats And Mitigation Strategies: A Survey, Journal of Theoretical and Applied Information Technology, Vol. 29, No. 2, pp. 6973, 2011.

[3] Egele, M. S, A Survey on Automated DynamicMalwareAnalysis.ACM Computing Surveys, Vol. 44, No. 2, 2012.
M. Wu, Y. Huang, S. Kuo, "Examining Web-based Spyware invasion with stateful behavior monitoring," 13th Paci_c Rim International Symposium on Dependable Computing (PRDC '07), 17-19 Dec.
Piscataway, NJ, USA: IEEE, pp. 275-81.

[4] Detection of Spyware by Mining Executable Files, 2010 International Conference on Availability, Reliability and Security 978-0-7695-3965-2/10.

[5] "Review on Feature Selection of Gene Expression Data for Autism Classification" – IJIRSET- International Journal of Innovative Research in Science, Engineering and Technology Vol. 5, Issue 3, March 2016 ISSN (Online): 2319-8753 ISSN (Print) : 2347-6710.

[6] "A Survey on Autism Spectrum Disorder Classification" - IJISET - International Journal of Innovative Science, Engineering & Technology Vol. 3 Issue 3, March 2016.ISSN 2348 – 7968.

[7] "Classification Scheme For Detecting Autism Using Three Tier Feature Selection And Single Tier Gene Selection" International Journal of Pure and Applied Mathematics Volume 117 No. 15 2017, 683-697.

[8] Yamini.K , Sivapriyadarshini.S (2010), ―A Data Mining Approach For On Time Detection Of Spyware Threat‖, IEEE International Conference on Computational Intelligence and Computing Research

[9] Bozagac .C.D, ―Application of Data Mining based Malicious Code Detection Techniques for Detecting new Spyware‖, White paper, Bilkent University 2005.

[10] Shahzak, R.K,"Detection of Spyware by Mining Executable files" Availability, Reliability, and Security, 2010. ARES '10 International Conference on Date of Conference: 15-18 Feb. 2010

[11] Nwokedi Idika and Aditya P. Mathur. "A Survey of Malware Detection Techniques."Technical report, Software Engineering Research Center, 2007.

[12] Martin Boldt, Andreas Jacobsson, Niklas Lavesson, and Paul Davidsson. "Automated Spyware Detection Using End User License Agreements." isa, 0:445–452, 2008.

[13] N. Lavesson, M. Boldt, P. Davidsson and A. Jacobsson, "Learning to Detect Spyware using End User License Agreements," Knowledge and Information Systems , in .press.

[14] N. Idika and A. Mathur, "A Survey of Malware Detection Techniques.," Software Engineering Research Center, Technical report 2007.

[15] M. Siddiqui, M. C. Wang, and J. Lee, "Detecting Internet worms Using Data Mining Techniques," Journal of Systemics, Cybernetics and Informatics, vol. 6, no. 6, pp. 48-53, 2009..

[16] R. Veeramani and Rai Nitin, "Windows API based Malware Detection and Framework Analysis," International Journal of Scientific & Engineering Research, vol. 3, no. 3, March 2012.

[17] N. Ye, "The Handbook of Data Mining". Lawrence Erlbaum Associates, Inc, 2003.

[18] T. Wang et al., "A Surveillance Spyware Detection System Based on Data Mining Methods," 2006.

[19] D. Olson and D. Delen, Advanced data mining techniques. Springer-Verlag Berlin Heidelberg, 2008.

[20] A. Sami, H. Rahimi, B. Yadegari, and S. Hashemi, "Malware Detection Based on Mining API Calls," ACM Symposium on Applied Computing, April 2010.