Virus Scanning & Detection Using Cryptographic Hash Function

Prof. Himgouri O. Tapase, Ruturaj Jadhav, Rutuja More, Kajal Katkar, Bhavika Oswal.

Department of Computer Science & Engineering,

YSPM's Yashoda Technical Campus, Satara-415001, Affiliated to DBATU University Lonere, Maharashtra, India

ABSTRACT

This paper provides the abstract detail of computer virus scanning and detection. The computer virus is a more harmful for the computer. People use computers for all kinds of activities such as online games, E-marketing, entertainment, emails, Social media, study, research, etc. The risk of infection with malicious software on the computers is increasing day by day. Most of user in all over world are facing the problem related to the thread it has been very difficult to prevent computer from being infected. The virus spread through computer and network. Many engineers have developed antivirus software to prevent virus from attacking your computer system. This research paper shows the description of preventing of computer virus and the serve related research works in fields and to develop an advisory program to help ordinary users learn about computer viruses and antivirus, understand the various features of antivirus software products, and select appropriate anti-virus software to protect their computers.

Key words: Antivirus, Virus removal, Detection of virus

1.INTRODUCTION

This research paper discovers the factor which shows the virus attacks among computer users. Now a day there is various growth in the smart phones, laptops, computers and applications, the amount of the virus damage the user privacy, crash their data. In today's world mobiles, computers are used to shopping, online gaming, office purpose, social network etc. So there is need to protect our data & devices. Keeping devices secure is the important task, so secure devices from the virus Securing devices is important thing and the big challenge since the viruses are increasing day by day. Virus is computer program that damage, destroy, crash devices and the user's personal data. Also virus is a computer program that can copy itself and infect a computer without permission and knowledge of user.

There are two common method of anti-virus system scan and detect the virus. First is "virus signature definition" & second "heuristic algorithm". Today we are in digital era where most of things are depends on the network, computer, applications. All the infrastructure such as government, financial institutions, banking sector, healthcare section are connected to the internet, some of information is important such as financial data, government data, personal data, so there is most of the chances of unauthorized access & losses of data because of infected by virus so "virus scanning and detection system is developed". This is the system of virus scanning and detection. In this system first of all user can register. After registration user can login with the id and password. login page blocks the unauthorized user. In my profile page shows all information related to a user. So user can get his id and username of registration. User interact using user interface. This interface has many options. Using that options user will perform tasks. Scan drive module Using scan drive option user will scan specific drive. In that option all files will scan which drive user has selected. There are two modules one is shows scanned files and another shows infected files. there is one option that is Main Page. Using that option user can go back to user interface. In manual scan user can select the particular file or folder which he wants to scan. Which include all types of file formats. This manual scan option shows selected particular file. So user can access that file. If we select any file for the scanning, then this shows the particular files generates the unique value called as checksum value. After that scan the selected file and display the result. This shows result of scanned particular file. If newly generated MD5 checksum value match with stored value, then it shows MD5 checksum value is found in virus dictionary otherwise it shows MD5 checksum value is not found in virus dictionary. If viruses not detected using above all options, then user can select second filter option to scan the computer system. User can scan the all drive. This show scanning process of all drives using command prompt. MD5 value generation show that if user wants to generate MD5 check sum value then user select particular file and generate checksum value. After that the select MD5 generator This shows that user select the particular file that he wants to generate checksum value. Then MD5 value can be generated of selected file by the user. Also it generates checksum value in virus dictionary and show it. File is created successfully with the content. User can access that file easily and scan it. After scanning show the result Phpmyadmin Database shows all the data stored in database.

I. LITERATURE REVIEW

According to the researchers, an everyday computer virus which demolish a whole computer system within a seconds. The beat thing that result an inflection is no longer have a data occupy the computer. Computer virus attacks have become major issue and can spread rapidly through the Internet and give rise to more damages.

Implementation of a Malware Scanner using Signature-Based Approach for Android Applications (Praful R. Pardhi, Jitendra Kumar Rou 2021):

In this paper, they implemented the signature-based scheme for scanning the applications in the mobile devices that are installed through legal market store or from the third party market and rates such applications as high, medium, and safe applications by analyzing the permissions used by such applications to detect known malware using the signature data set.

A survey of Static Android Malware Detection Techniques (Aiman Ahmad Abu Samra 2019 IEEE):

In this paper, they present a survey of the two major approaches of the static Android malware detection the permission based Detection technique and the signature based detection technique. It is a comparative study, that should be helpful for researchers in this topic.

Malware Detection in Mobile Through Analysis of Application Network Behaviour By Web Application (Himgouri P. Barge 2016):

In this paper they have taken information to detects the mobile malware by analysing suspicious network activities through the traffic analysis. In their system, the detection algorithms which they are using are works as modules inside the Open Flow controller, and the security rules can be imposed in real time. Here, they are using new behaviour-based anomaly detection system which is used for identifying meaningful deviations in a mobile application’s network behaviour. Here, they are trying to detect a new type of mobile malware with self-updating capabilities. This kind of malware neither identified by using the standard signatures approach nor applying static or dynamic analysis methods. The detection is completely based on the security rules can be imposed in real time. Here, they are using new behaviour-based anomaly detection system which is used for identifying meaningful deviations in a mobile application’s network behaviour.

Mobile Malware Detection through Analysis of Web Application Network Behaviour (Himgouri P. Barge 2016):

In this paper they have taken mobile malware by identifying suspicious network activities through real-time traffic analysis, which only requires connection establishment packets. Specifically, their detection algorithms are implemented as modules inside the Open Flow controller, and the security rules can be imposed in real time. They have present a new behaviour-based anomaly detection system for detecting meaningful deviations in a mobile application's network behaviour.

A machine learning approach for Linux malware detection. (K. Asmita & P. Vinod 2014):

They Introducing a novel approach using machine learning for identifying malicious Executable Link Able Files. The system calls are extracted dynamically using system call tracer Trace. In this approach identified best feature set of benign and malware specimens to build classification model that can classify malware and benign efficiently.

Mining specifications of malicious behaviour", Foundations of Software Engineering, (S. Jha and C. Kruegel 2007):

In this paper they used algorithm. Using that algorithm, they detect the malware. That algorithm provides a succinct description of malicious behaviour present in a malware, it can also be used by security analysts for understanding the malware.

Dynamic analysis of malicious code (U. Bayer, A .Moser, C. Kruegel and E. Kirda 2006):

In this paper they show Malware analysis is the process of determining the purpose and functionality of a given malware sample (such as a virus, worm, or Trojan horse). This process is a necessary step to be able to develop effective detection techniques for malicious code.

An automated virus classification system (M. Gheorghescu 2005):

In this paper, they introduce an innovative classification system that uses an average desktop machine. The classification system compares new and unknown samples with all existing malware, and within a few minutes, returns matches for that sample based on evolutionary behaviour of existing malware. Compared to previous methods, their method is independent of the malware class and language.

III, Methodologies

In this application we are using an algorithm to find virus and threats in computer system. This will help users to find threats and virus easily.

ALGORITHM USED:

1.MD5:

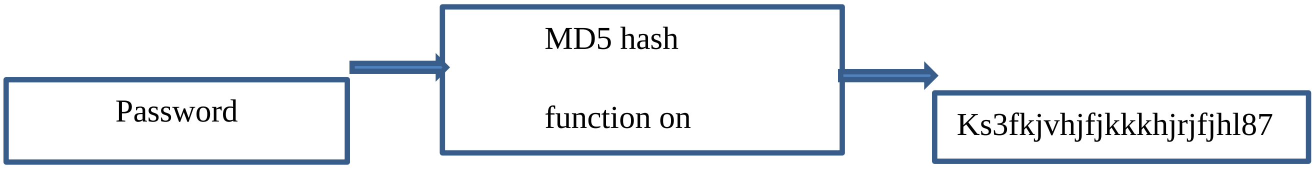
It stands for **Messages Digest Algorithm** which is a cryptography hash function algorithm in
this algorithm message taken as input of any variable length that means it doesn't have fixed length and it produce output into fixed length i.e. 16 bytes' length. It is advanced security purpose. This algorithm is improvement in MD4 algorithm. The digest size of MD5 is 128 bits.

There are two modules one is shows scanned files and another shows infected files. there is one option that is Main Page. Using that option user can go back to user interface. In manual scan user can select the particular file or folder which he wants to scan. Which include all types
of file formats. This manual scan option shows selected particular file. So user can access that
file. If we select any file for the scanning, then this shows the particular files generates the unique
value called as checksum value. After that scan the selected file and display the result. This
shows result of scanned particular file.

If newly generated MD5 checksum value match with stored value, then it shows MD5 checksum value is found in virus dictionary otherwise it shows MD5 checksum value is not found in virus dictionary. If viruses not detected using above all options, then user can select second filter option to scan the computer system. User can scan the all drive. This show scanning process of all drives using command prompt. MD5 value generation show that if user wants to generate MD5 check sum value then user select particular file and generate checksum value. After that the select MD5 generator.

This shows that user select the particular file that he wants to generate checksum value. Then MD5 value can be generated of selected file by the user. Also it generates checksum value in virus dictionary and show it. File is created successfully with the content. User can access that file easily and scan it.

After scanning show the result. Phpmyadmin Database shows all the data stored in database.



Working of MD5:

STEP:

1. Append Padding Bits

2. Append Length Bits

3. Initialize MD buffer.

4. Process Each 512 Bits.

1. Append Padding Bits:

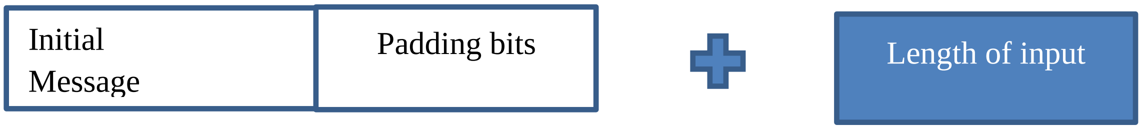
When the user receives the string which user has to input the user have to make sure that the size is 64 bits short of a multiple of 512. By using padding the extras bits are added.

Initial Message
Padding bits

Total length should be of 64 bits less but less than multiple of 512

2. Append Length Bits:

In the Step first we have to add more bits to make final string multiple of 512.



3. Initialize MD buffer.

In the initialize MD buffer there are 4 buffer such as N, O, P and Q. Each buffer which are use are of 32 bits.

N = 12 34 67 89

O = 76 ac gd jg

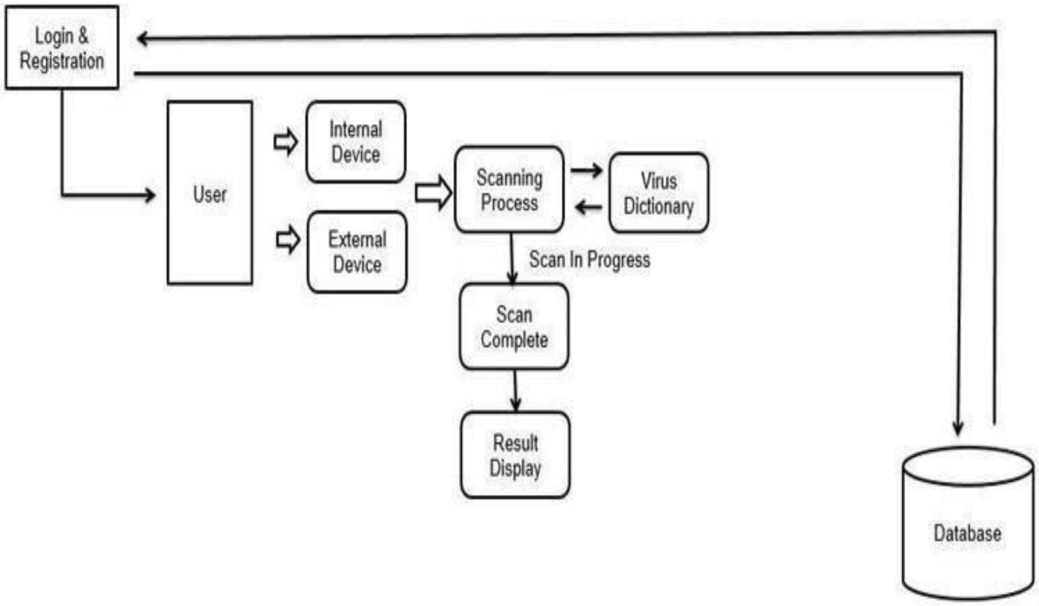
P = fg ki hg mn

Q = 78 56 45 34

3. Process Each 512 Bits:

This is a main process of MD5 algorithm. In this process there are 64 operations in the 4 Round. Each round there are 16 operations such as 1 round consist of 16 operations, 2 round consist of 16 operations, 3 round consist of 16 operations, 4 round consist of 16 operations. Total there are 64 operations.

IV. System Architecture:



V. RESULT

This shows result of scanned particular file. If newly generated MD5 checksum value match with stored value, then it shows MD5 checksum value is found in virus dictionary otherwise it shows MD5 checksum value is not found in virus dictionary.

VI.CONCLUSION

Viruses s are very harmful that can be devastating to many sector such as government sector, banking sector etc. Viruses are well known attacker and they slow down operation. Basically by using this system we can protect data from viruses. Main conclusion of this project is to secure and save the important data, files, folders, and external devices and internal devices

VII. REFERENCES

1. M. Gheorghescu, "An automated virus classification system", Virus Bulletin Conference, pp294-300, October 2005.

2. U. Bayer, A. Moser, C. Kruegel and E. Kirda, "Dynamic analysis of malicious code", Journal in Computer Virology, vol. 2, pp. 67-77, 2006.

3. M. Christodorescu, S. Jha and C. Kruegel, "Mining specifications of malicious behavior", Foundations of Software Engineering, pp. 110, 2007.

4. Moser, C. Kruegel and E. Kirda, "Exploring multiple execution paths for malware analysis", Proceedings of 2007 IEEE Symposium on Security and Privacy, 2007.

5. Josse, "Secure and advanced unpacking using computer emulation", Journal in Computer Virology, vol. 3, no. 3, pp. 221-236, 2007.

6. Michael Bailey, Jon Oberheide, Jon AndersenZ, Morley Mao, Farnam Jahanian and Jose Nazario, "Automated classification analysis of internet malware" in Recent Advances in Intrusion Detection. RAID 2007 Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, vol. 4637, pp. 178-197, 2007.

7. Asmitha and P. Vinod," A machine learning approach for Linux malware detection", International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), pp. 825-830, 2014.

8. Aiman Ahmad Abu Samra, "A survey of Static Android Malware Detection Techniques",26-March-2019 IEEE 7th Palestinian International Conference on Electrical and Computer Engineering (PICECE).

9. Praful R. Pardhi, Jitendra Kumar Rout, "Implementation of a Malware Scanner using Signature-Based Approach for Android Applications", 2021 19th OITS International Conference on Information Technology (OCIT), 1-March-2022.

10. Saurabh Pranjale, Soumya Mudgal, "Impact of Cyber-Attacks on Economy of Smart Grid and their Prevention", 28-April-2022, Vol. 8 No. 2(2022).