Analyzing Online Un-authorize Activity (Cyber Crime) Using Data Mining

Suraj B. Kokne[1], Ashwin G. Sonone[2], Ankush R. Ade[3], Bhagyashri P. Kedari[4], A.A. Bamanikar[5]
[1,2,3,4,5] *Student, P.D.E.A.'s COE Manjari (Bk.), Pune, Maharashtra*

Abstract- **It is very important to maintain security of data and protect it from unauthorized or illegal access. Many intruders try to attack the systems by using different types of cyber attacks like Denial of Service (DoS), Brute force attack, Man-In-the-Middle attack etc. Our focus of research in this paper is Denial of Service (DoS) attacks with the help of pattern recognition techniques in data mining. Through which the Denial of Service attack is identified. When the client and server exchange messages among each other, there is an activity that can be observed in log files. Log files give a detailed description of the activities that occur in a network that shows the IP address, login and logout durations, the user's behavior etc. In denial of service attack, IT resources of any organization are overloaded with imitation messages or multiple requests from unauthorized users. This system will try to apply pattern recognition technique on the data of user's or organization's log file.**

Index Terms- **Denial of Service, Log File, Cyber Crimes, Data mining, outliers, Association rules.**

I. INTRODUCTION

Security is the most important thing in every computer network. We have to improve security in a computer network. So in our paper we proposed a system which help us to detect the attacker on the particular computer from where he/she attacking the server. We are avoiding DoS(Denial of Service) Attack and SQL Injection attack on the system. If we want to block any application for particular user then admin only block the application for that user only.

We are detecting attacker with the help of log file. When the client and server exchange messages among each other, there is an activity that can be observed in log files. Log files give a detailed description of the activities that occur in a network that shows the IP address, login and logout durations, the user's behavior etc. Our focus of research in this paper is Denial of Service (DoS)attacks with the help of pattern recognition techniques in data mining.

Pattern recognition techniques is used to analyze the log file.using this technique the system can detect the attacker from the log file. Association rules are used to set the threshold value.if user sends the requests more than threshold value to the server then it will be considered as a DOS attack and system informs admin that DOS attack happen on particular computer in the network.Through which the Denial of Service attack is identified. Denial of service is a very dangerous attack that can damage the IT resources of an organization by overloading with imitation messages or multiple requests from unauthorized users. Cyber security is concerned with protecting IT resources like server; network etc. from performing illegal activities or fradulent acts. Data mining is also applicable to problem solving or network intrusions. Therefore in this paper we focus the applications of data mining for cyber security applications.

II. PROBLEM STATEMENT

To detect illegal activity of intruder and detect attack done by him like (DoS) by applying data mining technique and monitoring data of log files of legal users.

III. MODULES

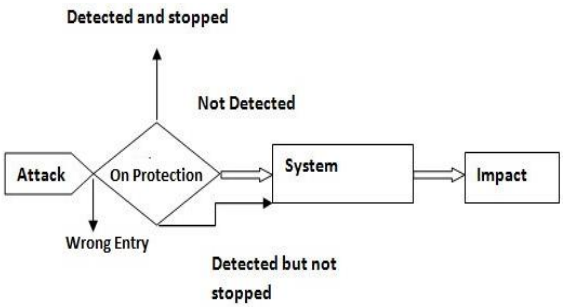
There are two main modules in propose system. First is admin module in that users are registered by admin. Admin controls all the registered users. All the user activity is stored in the log file and admin can read the log file and using pattern recognition system admin can detects the attacker. Admin has also right to block the specific user as well as particular application.

Second module is user module, in this module user has to only register into the system. If any user detected from the network it will considered as the attacker and Admin module can block the particular user.

III. DENIAL OF SERVICE ATTACK

In network architecture, the networking strategies should be made prone to identify intruders that attack the system (or cause a denial of service attack). DoS attacks [6] are some of the oldest Internet threats and continue to be the top risk to networks around the world. DoS events making it difficult to detect them. DoS attack remains a serious problem that increasingly affects company resources, in most cases a DoS attack caused the services to be completely unavailable impacting their business directly which is a potential financial loss to businesses.

IV. SYSTEM FLOW DIAGRAM



In this system we have applied the data mining techniques for identifying the Denial of Service attack. This type of attack is very dangerous. It makes the server busy by imitation messages and repeated queries. The server is congested by trace packets, in order to mitigate the server performance. In this project, we have discussed about Cyber security, cyber-crimes their types, clustering, outliers and pattern recognition. We have applied the famous data mining technique called as pattern recognition on the log file. We set a threshold value. If the number of similar requests are received at the server, which is greater than the threshold value, we assume this as an attack and the administrator is been informed. By this approach we can identify the denial of service attack easily as in DoS attack, the attacker or the hacker sends same multiple requests in order to mitigate the server performance. As well as we implement the SQLS injenction and U2R attack also , this attacks is also dangerous. This types of attacks makes system slower. We are preventing this types of attacks in our System.

V. METHODOLOGY

In our research paper, we have shown the concept of data mining techniques to identify cyber-attacks. Our focus of attention would be on “finding patterns” in a log file (records that occur in the system) which shows the sequence of events. From this log file we identify patterns. To start with, we use the clustering technique to discover the type of cyber-crime, Denial of service (DoS) attacks. As we know that clustering is grouping of data that has similar features. So this grouping helps to discover similar patterns of data that occur constantly in the log file.

Step 1: Evaluate the log file.
Step 2: Mine the date with time
Step 3: Scan the data
Step 4: Add the found data in the main file.

When the above procedure is carried out, we will record that data which contains normal patterns and also abnormal patterns (malicious). By using the clustering technique we identify the data that occur repeatedly. System Configuration: In order to run our obtained data, we use the Windows Server to maintain the database. Initially we run the data that contains zero attacks and then add them to the master file or log file. The ICMP (Internet Control Message Protocol) will make the system inactive by sending voluminous amount of “ping” command. Now the data that contains the normal activities and the data that contains attacks are passed through the technique that we have proposed. If the observations of the log file show normal behavior then they will be ignored. If the observations show multiple requests of the same transaction, then this data will be directed through our algorithm “Apriori” and will be shown in the attack logs. This algorithm will detect if similar patterns of requests exist in the normal records prior to consider it as attack. If the algorithm finds out the pattern and or finds the number of request for the same transaction more than the threshold value it is

considered as an attack and it sends signal or message to the administrator about the suspected attack.

RAM INfo -free memory: 80,128
allocated memory: 126,720
max memory: 253,440
total free memory: 206,848

IP Address- 192.168.1.119

MAC Address-54-8C-A0-C5-0A-87

Running Process-System Idle
Process,System,smss.exe,csrss.exe,wininit.exe,services.exe,lsass
.exe,svchost.exe,svchost.exe,svchost.exe,
svchost.exe,svchost.exe,svchost.exe,svchost.exe,WUDFHost.exe,igf
xCUIService.exe,RtkAudioService64.exe,svchost.exe,spoolsv.exe,wl
anext.exe,
conhost.exe,svchost.exe,svchost.exe,svchost.exe,svchost.exe,MsMp
Eng.exe,mysqld-nt.exe,esif_uf.exe,dasHost.exe,NisSrv.exe,
SearchIndexer.exe,sppsvc.exe,SppExtComObj.Exe,GoogleCrashHandler
.exe,GoogleCrashHandler64.exe,IntelCpHeciSvc.exe,csrss.exe,winlo
gon.exe,dwm.exe,sihost.exe,
taskhostw.exe,explorer.exe,igfxEM.exe,igfxHK.exe,RuntimeBroker.e
xe,RAVBg64.exe,esif_assist_
64.exe,ShellExperienceHost.exe,SearchUI.exe,RtkNGUI64.exe,
AdobeARM.exe,jusched.exe,MpCmdRun.exe,svchost.exe,OneDrive.exe,W
indows10UpgraderApp.exe,taskeng.exe,GoogleUpdate.exe,GoogleUpdat
e.exe,GoogleUpdate.exe,
eclipse.exe,fontdrvhost.exe,ApplicationFrameHost.exe,Tomcat7w.ex
e,dllhost.exe,svchost.exe,PartAssist.exe,chrome.exe,chrome.exe,c

We could see the DoS attack that has been made by the anonymous user (intruder) initially by gaining the access to the system (server) by posing as a authenticated user. In denial of service attack, the attacker gains the access through the vulnerabilities present in the system and copies the message sent by an authenticated user and makes multiple copies of the same request or query and sends it to the server. So, the server will process the same query or the request sent by a user for multiple times. In this way, the server is kept busy by processing the same request multiple times. This is called as denial of service attack. Another example is the "ping" attack where multiple ping requests will be sent from one user or multiple users and the server is again overloaded with processing the same request. This type of attack is severe. We apply data mining techniques to identify these types of attacks by finding similar patterns or request from the users. In our approach, we define a threshold of minimum support (5). If the same request is received to the server more than the threshold value, it assumes it as an attack and notifies the administrator. In some cases, based on the working environment, the threshold value could be set accordingly.

VI. OUTLIER

The term outlier refers to those readings in a data that are comparatively different from the rest of the data. Data mining techniques for detecting outliers will be extensively used for spotting abnormal behavior. When the outliers are discovered, their existence can signify important information. This information has helped many crime agencies, banks with unpredicted results. Outlier detection has been extensively studied in the recent years. Outlier is that concept that its existence is comparatively different from the remaining set of data. According to , outliers have been classified into two categories:

A. Classic Outlier approach In this approach, whatever outliers are observed in a dataset on the basis of transactions. Nevertheless, problems persist to exist to apply mining techniques on data such as repetition of large data sets, availability of vague information etc.

B. Spatial Outlier approach The term "spatial" means or refers to the objects that are present in space or have a geographical existence. Spatial Outlier refers to spatially referred object whose non spatial values are comparatively different in the same region.

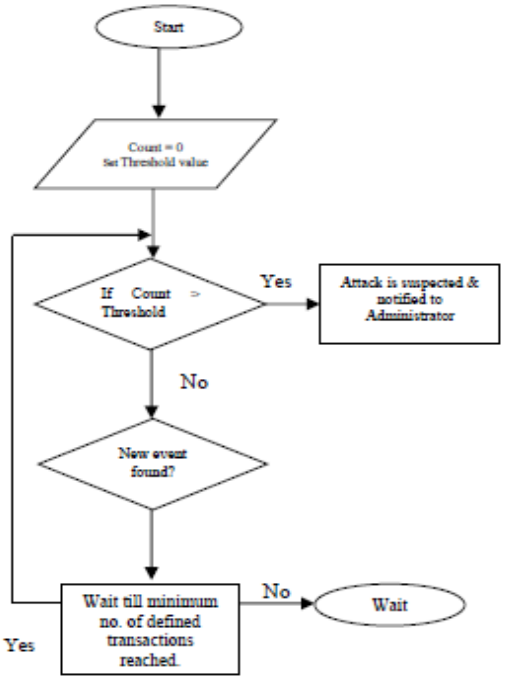


Fig. Flow Chart for detecting DoS Attack

VII. CONCLUSION

In this paper we have applied the data mining techniques for identifying the Denial of Service attack. This type of attack is very dangerous as it jeopardizes the IT resources. It makes the server busy by imitation messages and repeated queries. The server is congested by traffic packets, in order to mitigate the server performance. In this research paper, we have discussed about Cyber security, cyber-crimes their types, clustering, outliers and pattern recognition. We have applied the famous data mining technique called as pattern recognition on the log file. We set a threshold value. If the number of similar requests are received at the server, which is greater than the threshold value, we assume this as an attack and the administrator is been informed. By this approach we can identify the denial of service attack easily as in DoS attack, the attacker or the hacker sends same multiple requests in order to mitigate the server performance.

REFERENCES

[1] Know Your Enemy: Learning About Security Threats, 2nd Edition.ISBN: 0321166469. The Honeypot Project 2004.

[2] M.Khan , S.K.Pradhan, M.A.Khaleel, “Outlier Detection for Business Intelligence using data mining techniques”, International journal of Computer Applications (0975 -8887), Volume 106- No. 2, November 2014.

[3] Masud, M.M, Gao,J.Khan, ”Peer to Peer Botnet Detection for Cyber Security: A Data Mining Approach”. In proceedings: Cyber-security and information Intelligence research workshop. Oakridge national Laboratory, Oakridge May 2008.

[4] Internet Security Threat Report, Volume 21, April 2016, Symantec Crime Report.

[5] Ibrahim Salim, T.A.Razzack,”A study on IDS for Preventing denial of service attack using outliers techniques”, 2nd IEEE international conference on Engineering and technology, March 2016.

[6] S.S Rao, SANS Institute Infosec Reading Room.,”Denial of service Attack and mitigation techniques: Real time implementation with detailed analysis”, 2011.

[7] Data Mining:Concepts and Techniques, Third Edition, Jiawei Han and MichelineKamber, ISBN-13, 9780123814791.

[8] Mining of Massive Data Sets, AnandRajaraman, Jure Leskovec, Jeffrey D. Ullman,2014

[9] A. Klein, F. Ishikawa, and S. Honiden. Efficient heuristic approach with improved time complexity for qos-aware service composition. In ICWS, pages 436–443. IEEE, 2011.

[10] Tripathy, M.Khan, M.R.Patra, H.Fatima, P.Swain, “ Dynamic web service composition with QoS clustering” IEEE , International Conference on Web services, 2014.