Dos Attack and Suspicious Activity Detection for a Book Application

Shruthi C. V[1], Mrs. Asha[2]

[1]PG Student, Computer Science ,Dr Ambedkar Institute of Technology, Bangalore, Karnataka, India
[2]Associate Professor, Computer Science, Dr.Embedkar Institute of Technology, Bangalore, Karnataka, India

ABSTRACT

Globally the internet is been accessed by several people within their restricted domains. When the client and server exchange message among each other, there is an activity that can be observed and tracked in detail of the activities that occur in a network that shows the, login and logout durations, the user's behaviour etc. there are several types of attacks occurring from the internet. In this work the first focus is to provide product recommendations based on the collaborative filtering and content based recommendations. In personal setting the user in order to filter the products based on the transaction log maintained in terms of order information and order details. In Content based recommendations the recommendations related to the product are made based on the transactions performed by the end user. The Session tracking is performed for every click action and every navigation of the user and then behaviour based habitat file is generated. Two kinds of intrusion are detected one is behaviour change using Least Common Sub Square algorithm and then Dos Attack which is repeated action performed by the user within the limited time frame.

Keywords: Denial of Service, Log File, Cyber Crimes, Data mining, Association rules.

I. INTRODUCTION

Cyber Security is that branch of Computer Technology that deals with security in cyberspace. Cyberspace refers to the description of policies regarding the networks and computer system. The policies laid out in the Cyber security are for the reason of avoiding the malicious activity or unauthorized access to secured information. Since the emergence of high structured networks, there arises a concern about how intelligently these networks are secured.

Cyber-crime is one of the violence activities that can be conducted through internet. This is a large term reports everything from electronic cracking or denial of service attack that causes trading sites to lose money. Cyber is totally related to internet of things. Crime refers to something that is done illegally or without authorization. Cyber-crime is an integration of crime and computer. Any offence or crime in which a computer is used a cyber-crime. Date mining is the process of discovering patterns in huge data sets involving methods at the intersection of machine learning, statistics and database systems.

When a particular transaction/activity is performed repeated by a user in order to slow down the system is a suspicious activity. This may lead to system coming down for duration of period and then finally causing huge amount of revenue loss. This kind of attack is DOS attack. In the previous approach only system is monitored by a set of people/monitoring tools and if more traffic occurs then an additional server is added to manage the traffic which is time

CSEIT183632 | Received : 08 July 2018 | Accepted : 22 July 2018 | July-August-2018 [3 (6) : 175-180]

consuming and the entire set up has to be done this is called as downtime.

The previous approach does not take into consideration the set of actions which the user performs and track them where as the proposed approach will track each user action and navigation patterns to track the user behaviour over a period of time. The previous approach does not take session specific user behaviour over a period of time whether the proposed approach does that. The previous approach has only product buying and does not provide recommendations like content based recommendations based on user transactions or collaborative based recommendations based on ratings.

II. RELATEDWORK

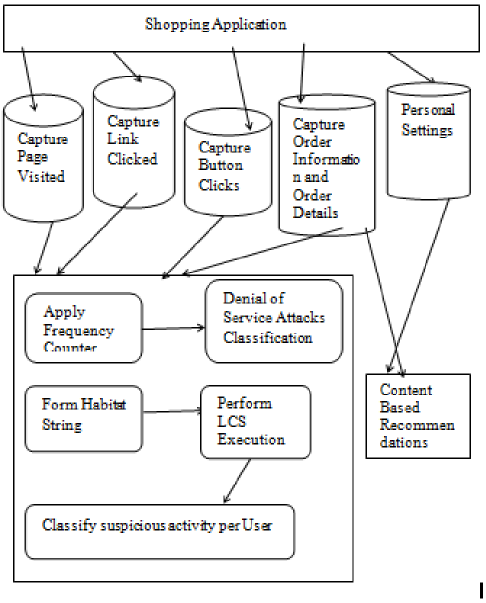
Real time implementation with detailed analysis of amongst various online attacks hampering IT security [6]; Denial of Service has the most devastating effects. It has also put tremendous pressure over the security experts lately, in bringing out effective defence solutions. These attacks could be implemented diversely with a variety of tools and codes. Since there is not a single solution for Dos this attack has managed to prevail on internet for nearly a decade. Hence, it becomes indispensable to carry out these attacks in small test bed environments in order to understand them better. These real time attacks are measured and analysed using network traffic monitors. In addition to that, this project also details various defence strategies that could be enabled on Cisco routers in order to mitigate these attacks. The detection and mitigation mechanism designed here are effective for small network topologies and can also be extended to analogous large domains.

Outliner Detection for Business Intelligence using data mining describe that a review of various outliner detection techniques from data mining perspective [2].Existing studies in data mining focus generally on finding patterns from large datasets and using it for organizational decision making. However, finding exceptions and outliers did not receive much attention in the data mining fields as other topic received.

III. PROPOSED METHOD

In the proposed approach the data mining techniques namely association rule mining and clustering algorithms are applied on the log file of the web application and then a grouping is made of the user who is trying to do the denial of service attacks. Denial of service attack is an attack which brings the system down after huge numbers of same repetitive requests are made to the application.

The following fig.1 shows the System Architecture of the project. As shown in the fig first an application is needed in order to implement the Audit and in this case we consider the access to application with internal functions as well as external functions. Event Computation is a process in which each action of the user in the application is captured based on action name and user id. The action name can be button click, URL click or some other actions. The navigation patterns are sequence of actions captured per session.



I

Figure 1. System Architecture diagram

A. Registration

This Module is responsible for allowing any external customer to perform the registration by proving the details like First Name, Last Name, User Id, Password, Email, City, State and Country. If the user id already exists then user is not allowed to register.

B. Login

Login Module is responsible for allowing the user to access the user with valid credentials and deny the access for user with invalid credentials. The Users are of two kinds one is Admin and other is Customer. If it is Admin then he/she can see the habitat file for each session of the users using the application which has the session tracking, Find Suspicious pattern using LCS and Detect Dos Attack. If it is customer then he/she can purchase product and then get the recommendation.

C. Product Buying

Product Buying is responsible for purchasing the products by providing the valid IPIN and Account No. If the credentials are valid and also user has sufficient balance then product is purchased and then two important information's are tracked namely Order Information and Order Details.

The Order information can be described as below

Table 1

Name	Description
ORDERID	Unique ID representing the Order and acts like the primary key
LOGINID	Login ID of the user
ORDERDATE	The Date of purchase
TOTALAMOUNT	Total Transaction Amount
Email	Email Id of the Logged in User

And Order information can be described as below

Table 2

Name	Description
ORDERID	Unique Order Id and acts like a primary key. The ORDERID of Order Details and Order Information are maintained in sync
PRODUCTID	The id of the product which is being purchased
QUANTITY	The quantity of the product purchased in a single transaction

D. Personal Setting

This is a setting set by the user in order to filter the products based on the transaction log maintained in terms of order information and order details. This is used by content based recommendation algorithm.

E. Content Based Recommendation

Content Based Recommendations is user specific data in order to provide recommendations. In this module the user will select a product and enter the credit card details and then completes the transactions. Behind the scenes the merchant maintains the transactions and then finds the best products suited for the user. Each user will have recommendations based on the personal settings set by the user.

F. Session Tracking And Habitat File

In this module whenever clicks on the link or clicks on the button or user navigates from one page to another page each time independent request is made and tracked based on the user id and session id. The habitat file is set of records which are set of actions performed by the user in each session.

G. Intrusion Detection using LCS

This module is responsible for taking a set of patterns and finds the LCS for each of the pattern. The pattern refers to one habitat of the user for specific session. If the LCS is new as compared to previous activities then the pattern is regarded as intrusion. Give two sequences, find the length of longest subsequence present in both of them. A subsequence is a sequence that appears in the same relative order, but not necessarily contiguous. For example, “abc”, “abg”, “bdf”, “aeg”, ‘”acefg”, .etc are subsequence of “aabcdefg”. So a string of length n has 2^n different possible subsequence.

H. Dos Attack Detection

This Module is responsible for finding the Dos Attacks over a period of window by measuring the frequency of the actions performed by the user and then ranking based on the highest frequency of steps.

I. Blocking Users of Dos

This module is responsible for blocking the user who is responsible for Dos Attack

Least Common Sub Square Algorithm

1) Consider the sequences namely $S1$ *and* $S2$
2) The length of S1 and length of S2 is computed
3) Find the maximum length of S1 and S2
4) Construct a matrix initial with zeros one 1st row and 1st column
5) If the value of the sequence alphabet is not there then maximum on top and left is taken
6) If the value matches then diagonal value is increment by a value of 1
7) If the Value does not match diagonal take a maximum between right top and current value
8) The point where the value is maximum is defined as the Length of LCS
9) The point from bottom till top is taken to get the string of LCS

How DOS Attack Threshold is computed

1. Obtain the List of Users from Habitat file
2. For each of the User Obtain the Latest Session APP ID
3. Add the Session APP ID to the List
4. Obtain List of Unique Action Names across the Session APP ID List
5. Compute the Frequency for each Action Name
6. Obtain the Maximum Frequency
7. The Maximum Frequency corresponds to DOS threshold

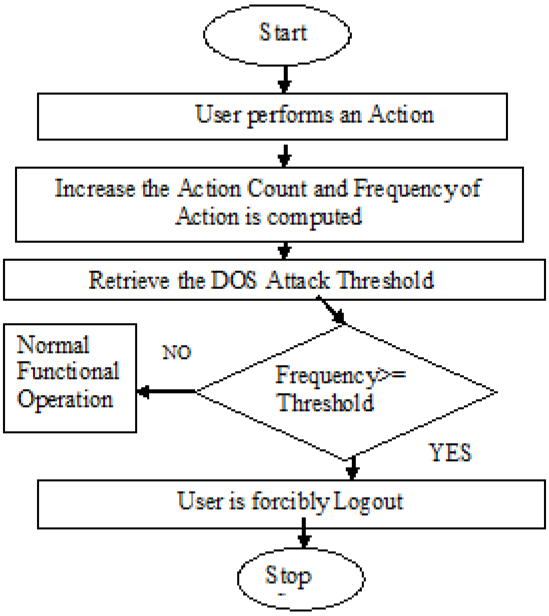
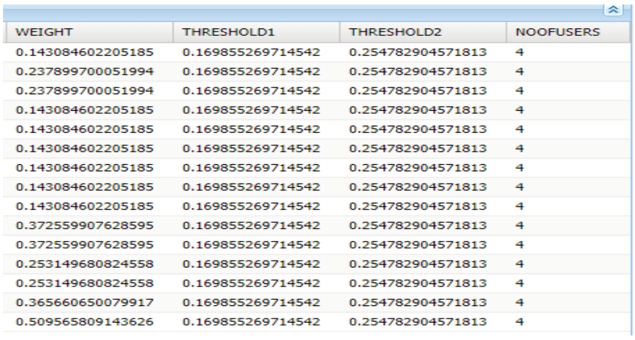

Start
User performs an Action
Increase the Action Count and Frequency of Action is computed
Retrieve the DOS Attack Threshold
Normal Functional Operation
NO
Frequency>= Threshold
YES
User is forcibly Logout
Stop

Figure 2. Dos Attack Diagram

IV. EXPERIMENTAL RESULT

In this project work user registration successfully has now will login valid user name and password on login nothing is there because no licenses are given by the particular user now log out .Admin give the licenses to the new user now what I will do logout than login in that particular user should able to see all six categories of book. User purchases different book in sufficient funds for valid account. View the rank books details in no book satisfies the content based ranking because no give the personal sitting than user is set the threshold.

Setting has been stored successfully now view the rank books seeing the book because it exceeds equal to the personal sitting other book does not satisfying the personal setting that is called content based recommendation. Budgets is warring if exceeds the amount that show the warring based on you decision you can continue with the transaction or ignore. In habitat file view the list of session name entire thing is captured .Dos attack per session, whereas LCS is across the session. LCS detects the suspicious output see the weight in between threshold1 and threshold2 is called type two suspicious. Type three suspicious is means no suspicious our weight exceeds the threshold2, if the session are type three suspicious not sent by the user only type two suspicious.



WEIGHT	THRESHOLD1	THRESHOLD2	NOOFUSERS
0.143084602205185	0.169855269714542	0.254782904571813	4
0.237899700051994	0.169855269714542	0.254782904571813	4
0.237899700051994	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.143084602205185	0.169855269714542	0.254782904571813	4
0.372559907628595	0.169855269714542	0.254782904571813	4
0.372559907628595	0.169855269714542	0.254782904571813	4
0.253149680824558	0.169855269714542	0.254782904571813	4
0.253149680824558	0.169855269714542	0.254782904571813	4
0.365660650079917	0.169855269714542	0.254782904571813	4
0.509565809143626	0.169855269714542	0.254782904571813	4

Figure 3

User is there always taking product buying coming back it not purchasing now load on sever increases.we need to detect Dos attack based on frequency and we need block the particular user.In graph show there are 11user,10 user are nondosattack and 1user is Dosattack.

Classification Graph

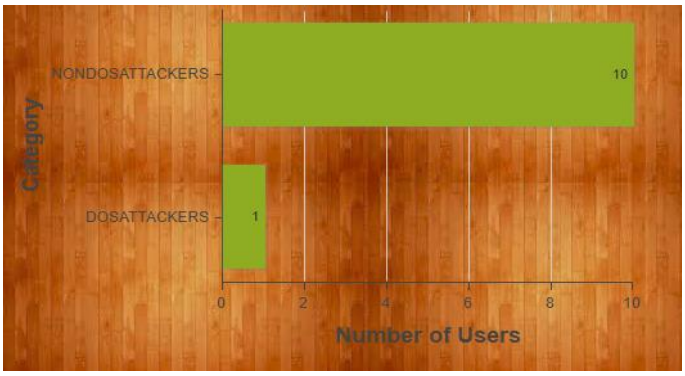


Figure 4

V. CONCLUSIONS

In this project the customer will be able to register into the application and after that the customer will be able to purchase different categories of books. The customer will be able to set the personal settings and get the content based recommendations based on the personal settings and transaction history. The customer will also be able to set the budget and then will be able to get the warning if sufficient amount is available and after that will be able to make a decision whether to proceed with transaction of the user. Each and every page visit and button click will be captured in the form of habitat file. The habitat file will have the tracking based on action name, action type, time and date of session along with user id as well as session id. The admin will be able to execute the LCS algorithm and then determine the LCS for each of the session. The admin will be able to run the Intrusion Detection in order to detect whether the sessions are low, medium or highly suspicious. If the sessions falls under low or medium suspicious then the user will get notified otherwise the user will not be notified for no suspicious. The Admin will be able to dynamically determine the DOS attack threshold and then the user will be logout if the repeated actions are performed. The Admin will be able to see the classification graphs as well as the DOS log in terms of which action has caused the dos attack and what is the reason i.e. page. The User who is responsible for DOS attack will be shown a message while login. Your account is blocked because of DOS attack. The admin will be able to unblock the user.

VI. 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 Micheline Kamber, ISBN-13, 9780123814791.

[8]. Mining of Massive Data Sets, Anand Rajaraman, 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.