Network Intrusion Detection System

DR.M.G.R.Educational and Research Institute

Chennai,Tamil Nadu-600095

R.Ajaykumar M.Subash Chandra bose S. Sriramkumar

[1]UG-PG student, Research Scholar, Asstt / Asso. Professor, Professor, Dr. (10,Times ,Italic)
Name of the Institute, City, Country, Pin, email ID of Correspondence author
[2]UG-PG student, Research Scholar, Asstt / Asso. Professor, Professor , Dr.
Name of the Institute, City, Country, Pin, email ID of Co-authors
(You can add max. 04 authors)

Abstract *– This report discusses the research done on the chosen topic, which is Network Intrusion Detection System. This project shows that monitoring and detection of the network will reduce the down time of the network and reducing future attacks. In addition, a comprehensive and organised analysis is conducted to verify the causes of the attack. It has been found that most household internet user lacks the means to strengthen their internet connection or networking system. The problem of this project is an unauthorised access into a home networking system that may cause harm by stealing private and confidential information as firewall and anti-virus won't be sufficient against a determine attacker. The scope for this project is to develop an intrusion detection system that will improve the security of home network as that is the potential user of this system. The objective of this project is to investigate the methods needed to detect any unauthorised access into a home networking system. The detection system will use an open source system that are readily available but will be tuned for the usage of home user and based on Windows operating system. In methodology section, it will discuss regarding the usage of Iteration Development Model as the methodology used in developing this project. In the results and discussions section, the preliminary findings consist of the findings from literature review research, own research and the use case diagrams of the system. Then, the prototype development process and results together with the testing results will be discussed in detail.*

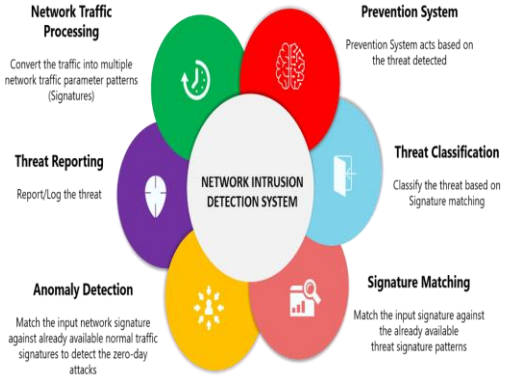
INTRODUCTION

Almost all of the people use the internet to carry out essential activities such as bill payment, bank transfer and etc. But attacks towards home network are not uncommon nowadays as everybody is connected to each other through the internet and the attack has been growing more frequent and severe. When an attack do occur, it is essential that a comprehensive and organised analysis is conducted to verify the causes of the attack and the damages of the attack. A thorough and timely investigation and response can serve to minimize network downtime and ensure that critical business systems are maintained in full operation.The determined hacker can find a way into your network either by establishing some type of connection and entering your virtual "front door" or by using social engineering tactics to obtain user ID and password information. Whatever the method used, the fact is that an intruder can get into your network and harm your business.

II. LITERATURE REVIEW

Types of Network Intrusion Detection System

Fig; 1 Network intrusion detection sysetem



An Intrusion Detection System is a defence mechanism, which detects hostile activities in a network. System will be compromise if the intrusion is not detected and possible prevented. One of the major benefits of intrusion detection system is it provides an overview of any unusual unscrupulous activities. According to (Amoroso, 1999), intrusion detection is "aprocess of identifying and responding to malicious activity targeted at computing and networking resources". Even though there are firewall and antivirus programs installed to protect their computer from any unwanted access, it can still be vulnerable to any unauthorised user. With the inclusion of network intrusion detection and prevention system, there will be another protection layer against potential hackers. Intrusion detection and prevention systems are much more secure than common firewall technology. Although considered to be an expansion of the original intrusion detection system, they are actually more a way of controlling who has access to a computer network. They not only control access, but also detect entry to the network, so the two systems are closely linked. There are 4 types of detection system. One of the systems is network-based detection system

where it is mostly used on virtual private servers, remote access servers, and routers by analysing various network protocols Wireless intrusion detection system works much like network-based system only that it applies on wireless networks (Adams). Access point misuse is one of the illegal activities that are monitored by the system. In hostbased system, works on an individual computers.

Some of the data that are commonly logged by network-based Intrusion Prevention System are:

· Timestamp

· Packet ID

· Event or action type

· Rating (e.g., priority, severity, impact, confidence)

· Network, transport, and application layer protocols

· Source and destination IP addresses

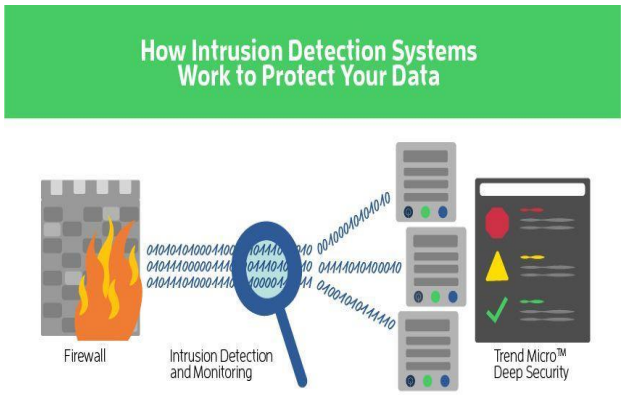
· Source and destination TCP or UDP ports, or ICMP types and codes

· Number of bytes transmitted over the connection

· Decoded payload data, such as application requests and responses

· State-related information

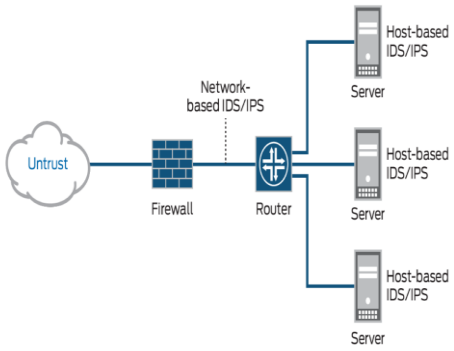
Fig ; 2 Working model



III. METHODOLOGY

All phases in System Development Life Cycle are included in iterative development. Phases in iterative development model are: · Planning – To plan what is needed to be done to make sure this system can be implemented on time. Analysis & Design – To determine the problem and solution. Implementation – To take the solution and implement it. Building the system. Deployment - Install the system and provide user manual, training and maintenance. Testing - Testing is conducted to make sure that each unit meets the user's requirement. Machine learning algorithms are then applied to the audit records that are processed according to the feature definitions to generate intrusion detection rules.

Figure ; 4 Principle of IDS/IPS

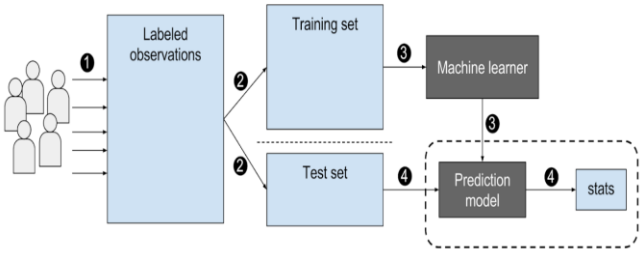


An Anomaly-Based Intrusion Detection System is a system for detecting computer intrusions and misuse by monitoring system activity and classifying it as either normal or anomalous. This encouraged us to study some Data Mining based frameworks for Intrusion Detection.

IV. MODELING

The final step is to build our model as the data is full prepared and ready. We'll one again leverage scikit-learn to split our data into train and test sets. We split our data into train and test in a 4:1 ratio 67% train data and 33% test data. We could opt 80-20 or 50-50 splits as well (train-test) Post that, we use the Decision Tree Classifier with entropy as our criterion and a max depth of 4. And then we train the model with our training just taking less than 1 second. Post-training, we do the prediction with the test data and that also takes less than a second. We can observe that the train and test accuracies are 91.27% and 90.66% respectively, which is pretty good.To see if the accuracy can be improved, we try to train and test using the XG Boost algorithm as well leveraging the XGBClassifier. After train and test, we can observe that the accuracy improved by quite a good margin. Training accuracy jumped to 98.72% and the test accuracy was 98.51%. The XG boost algorithm basically makes use of gradient boosting decision tree algorithm.

Fig; 3 Stats Prediction



V. Algorithms Used

DECISION TREE

The decision tree Algorithm belongs to the family of supervised machine learning algorithms

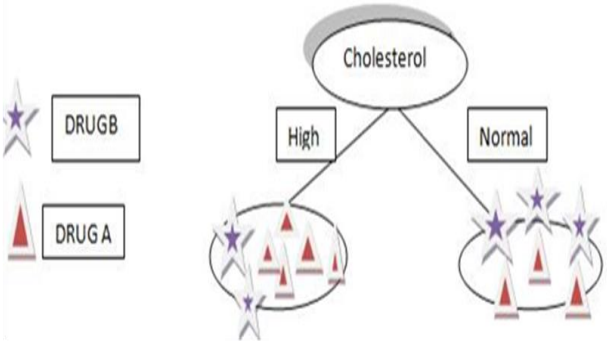
It can be used for both a classification problem as well as for a regression problem. The goal of this algorithm is to create a model that predicts the value of a target variable, for which the decision tree uses the tree representation to solve the problem in which the leaf node corresponds to a class label and attributes

are represented on the internal node of the tree. Let's take a sample data set to move further ….

Patient ID	Age	Sex	BP	Cholesterol	Drug
p1	Young	F	High	Normal	Drug A
p2	Young	F	High	High	Drug A
p3	Middle-age	F	Hiigh	Normal	Drug B
p4	Senior	F	Normal	Normal	Drug B
p5	Senior	M	Low	Normal	Drug B
p6	Senior	M	Low	High	Drug A
p7	Middle-age	M	Low	High	Drug B
p8	Young	F	Normal	Normal	Drug A
p9	Young	M	Low	Normal	Drug B
p10	Senior	M	Normal	Normal	Drug B
p11	Young	M	Normal	High	Drug B
p12	Middle-age	F	Normal	High	Drug B
p13	Middle-age	M	High	Normal	Drug B
p14	Senior	F	Normal	High	Drug A
p15	Middle-age	F	Low	Normal	?

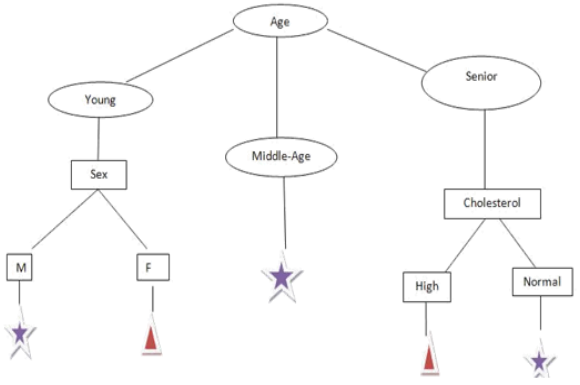
Suppose we have a sample of 14 patient data set and we have to predict which drug to suggest to the patient A or B.

Let's say we pick cholesterol as the first attribute to split data



Fig; 5 Example of decision tree

It will split our data into two branches High and Normal based on cholesterol, as you can see in the above figure.Let's suppose our new patient has high cholesterol by the above split of our data we cannot say whether Drug B or Drug A will be suitable for the patient.Also, If the patient cholesterol is normal we still do not have an idea or information to determine that either Drug A or Drug B is Suitable for the patient.Let us take Another Attribute Age, as we can see age has three categories in it Young, middle age and senior let's try to split



From the above figure, Now we can say that we can easily predict which Drug to give to a patient based on his or her reports.Assumptions that we make while using the Decision tree:

– In the beginning, we consider the whole training set as the root.

-Feature values are preferred to be categorical, if the values continue then they are converted to discrete before building the model.

XGBoost

The beauty of this powerful algorithm lies in its scalability, which drives fast learning through parallel and distributed computing and offers efficient memory usage.It's no wonder then that CERN recognized it as the best approach to classify signals from the Large Hadron Collider. This particular challenge posed by CERN required a solution that would be scalable to process data being generated at the rate of 3 petabytes per year and effectively distinguish an extremely rare signal from background noises in a complex physical process. XGBoost emerged as the most useful, straightforward and robust solution.

RANDOM FOREST

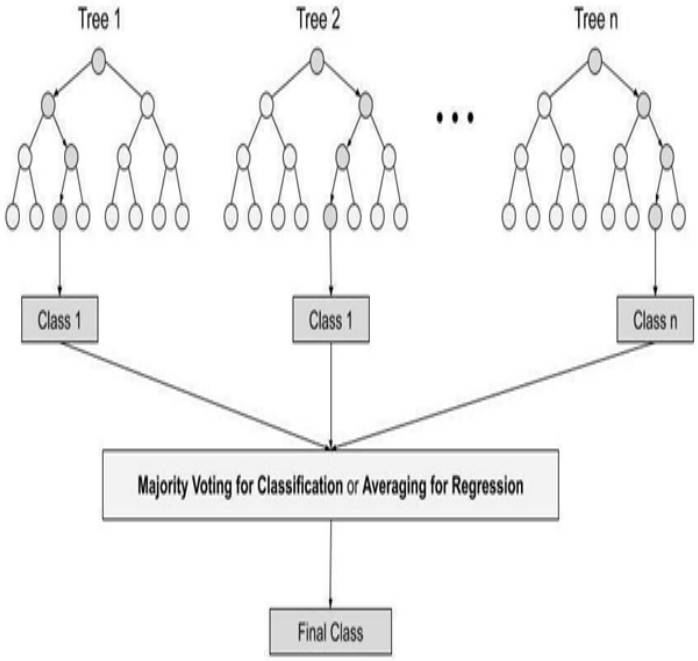
Step 1: In Random forest n number of random records are taken from the data set having k number of records.

Step 2: Individual decision trees are constructed for each sample.

Step 3: Each decision tree will generate an output.

Step 4: Final output is considered based on *Majority Voting or Averaging* for Classification and regression respectively.

Fig ; 6 Example of Random forest



VI. . CONCLUSION

This paper was written as an approach to detect any abnormal network activity in a home networking system. It has been shown that Intrusion Detection System is very useful for improving network security by adding another layer of security on top of the firewall and anti-virus programs. However in reality hackers may take advantage of any possible resources to attack a computer network, and these resources may be unstable and very difficult to categorise. These difficulties can reduce the effectiveness of this approach. Thus, a more secured way of using the internet is needed for the user by only surfing the internet with an updated anti-virus and understanding the risk of surfing the "wrong" sites.

VII. REFERENCES

1. Adams, K. Types of Intrusion Prevention Systems. Retrieved from eHow: http://www.ehow.co.uk/info_8039841_types-intrusion-prevention-systems.html

2. Amoroso, E. (1999). Wykrywanieintruzów. Warszawa 1999: Wydawnictwo RM.

3. Bo, J. (30 August, 2010). Phishing Methods and Prevention. Retrieved from Yahoo Voices: http://voices.yahoo.com/phishing-methods-prevention-6664318.html

4. Kaspian, P. (23 July, 2013). Network Security in 2013: Is Your Intrusion Prevention System Ready? Retrieved from Security Intelligence Blog: http://securityintelligence.com/network-security-in-2013-is-your-intrusionprevention-system-ready/#

5. Kazienko, P., &Dorosz, P. (3 April, 2003). Intrusion Detection Systems (IDS) Part I - (network intrusions; attack symptoms; IDS tasks; and IDS architecture). Retrieved from WindowSecurity.com: http://www.windowsecurity.com/articlestutorials/intrusion_detection/Intrusion_Detection_Systems_IDS_Part_I__network_intrusions_attack_symptoms_IDS_tasks_and_IDS_architecture.html

6. Liniger, R., & Vines, R. D. (2005). Phishing: Cutting the Identity Theft Line. Indianapolis: Wiley Publishing Inc.

7. Magalhaes, R. M. (10 July, 2003). Host-Based IDS vs Network-Based IDS (Part 1). RetrievedfromWindowSecurity.com: http://www.windowsecurity.com/articlestutorials/intrusion_detection/Hids_vs_Nids_Part1.html

8. Parno, B., Kuo, C., &Perrig, A. (n.d.). Phoolproof Phishing Prevention. Carnegie Mellon University.

9. Reid, C. E. (12 February, 2009). History of Phishing. Retrieved from All Spammed Up: http://www.allspammedup.com/2009/02/history-of-phishing/

10. Rouse, M. (September, 2005). Snort. Retrieved from SearchMidmarketSecurity.

11. Rouse, M. (May, 2007). Phishing. Retrieved from SearchSecurity: http://searchsecurity.techtarget.com/definition/phishing

12. Rozenblum, D. (2001). Understanding Intrusion Detection Systems**.** SANS Institute, 9.