International Journal of Advance Engineering and Research Development

Volume 3, Issue 4, April -2016

Monitoring And Controlling Of LAN Using IDS Through Honeypot And Send Alerts On Android Application

M Mohita Raol[1], Akshay Mandke[2], Somesh Padman[3], Prabodhan Kadepurkar[4]

[1,2,3,4]*Dept. Of Information Technology, PES Modern College Of Engineering,Pune*

Abstract —*In the technology driven world we live in, the information that is transferred over the network is prone to various kinds of attacks. Hence to maintain the integrity and secrecy of the data, Intrusion Detection Systems are introduced. Intrusion detection systems provide the ability to identify security drawbacks in a system. With networks getting faster and with the increasing dependency on the Internet both at the personal and commercial level, intrusion detection becomes a challenging process. Honeypot is a virtual concept that works by fooling attackers into believing it is a legitimate system; they attack the system without knowing that they are being observed covertly. When an attacker attempts to compromise a honeypot, attack related information, such as the IP address of the attacker, are collected. IDS captures the packets coming into the network and analyzes them using the probability of the packets to distinguish them as normal or intruded.*

Keywords- *Intrusion Detection System, ID3 algorithm, honeypot, decision tree, Attack patterns*

I. INTRODUCTION

Considering the current scenario of increasing technology, the protection of the large amount of crucial data stored in repositories of any organization has become a major issue. The organizations thus contact various security agencies to intervene and solve the problem of data security. The main concern which arises here is the continuous development of network attacks by the intruders who want to gain the crucial information about an organization by attacking its Local Area Network. An intruder can be defined as somebody attempting to break into an existing computer. This identity is popularly termed as a hacker, black hat or cracker. We need to solve this problem by developing such a system which tries to detect unknown attacks by analyzing the attack patterns of known attacks. This system which detects the unknown as well as known attacks is termed as intrusion detection system. The virtual concept of honeypot is used to make the system vulnerable to various kinds of attacks so that the analysis can be done on those attacks. A honeypot is a non-production system, design to interact with cyber-attackers to collect intelligence on attack techniques and behaviors. The attackers or the intruders are unaware that their activity is being monitored. In this way the system is open for all type of attacks coming through various kinds of packets, viz. TCP packet attack, UDP packet attack etc. The proposed IDS has multiple clients connected in a LAN which are controlled and monitored by a single administrator the server. The packets coming into the LAN are captured and analyzed by packet scanning module running on each client. The information about the attacks are send to the admin server. The security agencies can use this IDS to update their attack database continuously which can help them to protect the LAN of any organization against various security threats.

Purpose of an Intrusion Detection System:

The main purpose of the IDSis to provide security for an organization's local area network by detecting the attacks on the network. It also ensures the integrity, privacy, data confidentiality of the data being transmitted over the network.

The functionalities of the developedIntrusion Detection System includes:

- Multiple clients connected in a LAN
- Monitoring and controlling of the clients by a single administrator via the server
- Start Process, Kill process, Create file, Delete file, Read file options available to the administrator for controlling the clients
- Capturing of network packets coming into the LAN in client PC's
- Scanning and analysis of the captured packet's features.
- Finding probability of packet features for analysis
- Helps detecting unknown as well as known attacks.
- Attack details sent to server for administrator to take further action

II. PROPOSED APPROACH

2.1. Iterative Dichotomizer3 algorithm

ID3 builds a decision tree from a fixed set of examples. The resulting tree is used to classify future samples. The example has several attributes and belongs to a class (like yes or no). The leaf nodes of the decision tree contain the class name whereas a non-leaf node is a decision node. The decision node is an attribute test with each branch (to another decision tree) being a possible value of the attribute. ID3 uses information gain to help it decide which attribute goes into a decision node. The advantage of learning a decision tree is that a program, rather than a knowledge engineer, elicits knowledge from an expert.

2.1.1.Attribute Selection - Information Gain. a statistical property, called information gain is used to select the best attribute. Gain measures how well a given attribute separates training examples into targeted classes. The one with the highest information (information being the most useful for classification) is selected.
IG(A) is the measure of the difference in entropy from before to after the set S is split on an attribute A. In other words, how much uncertainty in S was reduced after splitting set S on attribute A.

$$IG(A,S) = H(S) - \sum_{t \in T} p(t)H(t) \qquad (1)$$

Information gain formula
Where,

- H(S)- Entropy of set S
- T- The subsets created from splitting set S by attribute A such that
- P(t)- The proportion of the number of elements in t to the number of elements in set S
- H(t)- Entropy of subset

Information Gain for attribute A on set S is defined by taking the entropy of S and subtracting from it the summation of the entropy of each subset of S, determined by values of A, multiplied by each subset's proportion of S.

2.1.2. Shannon Entropy. In order to define gain, we first borrow an idea from information theory called entropy. Entropy measures the amount of information in an attribute.ID3 Generates Decision Trees using Shannon Entropy.H(S) is a measure of the amount of uncertainty in the (data) set S (i.e. entropy characterizes the (data) set S).

$$H(S) = -\sum_{x \in X} p(x) \log_2 p(x) \qquad (2)$$

Entropy formula
Where,

- S- The current (data) set for which entropy is being calculated (changes every iteration of the ID3 algorithm)
- X- Set of classes in S
- P(x)- The proportion of the number of elements in class x to the number of elements in set S

When H(S) = 0, the set S is perfectly classified (i.e. all elements in S are of the same class).In ID3, entropy is calculated for each remaining attribute. The attribute with the smallest entropy is used to split the set S on this iteration. The higher the entropy,the higher the potential to improve the classification here.

2.1.3. Steps. Taking the example of the scenario of the Intrusion detection system, let us assume

- the training dataset contains 10 packets
- 6 packets are attacking packet and 4 packets are normal packets
- Each packet has attributes like source port, destination port, length, hop limit, acknowledgment flag, syn flag,rst flag, fin flag etc

1) Establish the target Classification/classification attribute
Is the packet intruded?
Yes = 6/10 No = 4/10

2) Compute Classification Entropy.
IE = -(6/10)log2(6/10) -(4/10)log2(4/10)
3) For each attribute of the packet, calculate Information Gain using classification attribute to determine the root node.
Taking first feature suppose as length
6 high lengths
2 medium lengths
2 small lengths

3 values for attribute length, so we need 3 entropy calculations

small: 0 yes, 2 no	I-small = $-(2/10)\log_2(2/10)-0$
medium: 0 yes, 2 no	I-medium = $-(2/10)\log_2(2/10)-0$
large: 0 no, 6 yes	I-large = $-(6/10)\log_2(6/10)-0$

Table 1. Entropy for each matching attribute

Now calculate Information gain for length
$IG_{Length} = IE(S_{Intrusion}) - [(2/10)* I_{small} (2/10)*I_{medium} + (6/10)*I_{large}]$
Calculate the same information gain for rest all attribute
4) Select Attribute with the highest gain to be the next Node in the tree (starting from the Root node).

Let us assume length to be the attribute with highest weight
Therefore, root node is length and then and we can immediately predict the packet is not fast when length is small or medium.

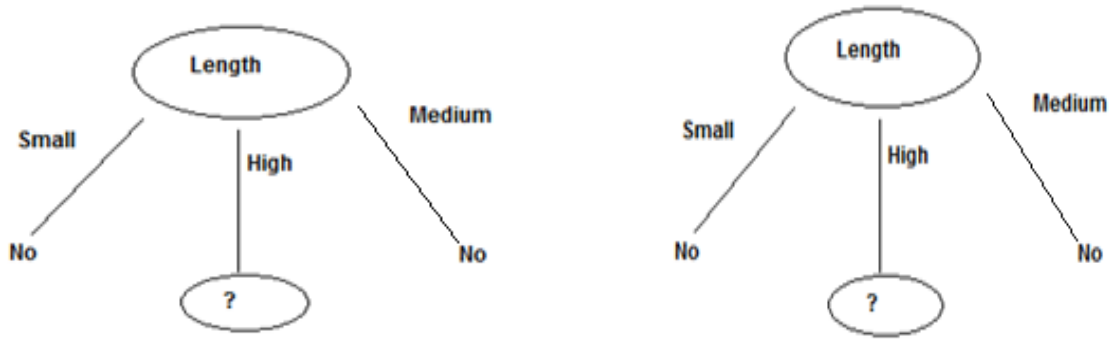

Figure 1.Root node selection

5) Remove Node Attribute, creating reduced table
Since we selected the length attribute for our Root Node, it is removed from the table for future calculations.
6) Repeat steps 3-5 until all attributes have been used, or the same classification value remains for all rows in the reduced table.

Final Decision Tree for classification:

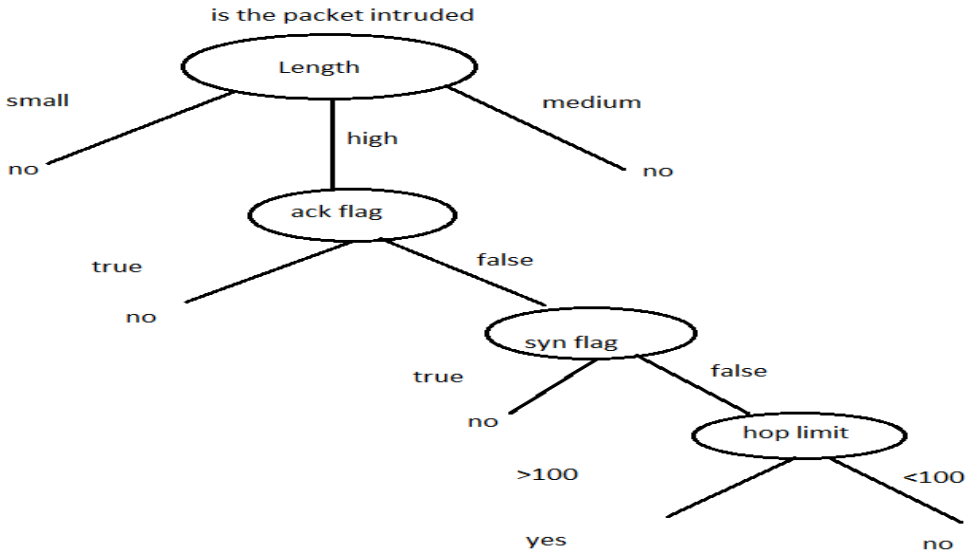

Figure 2.Final decision tree

Advantages of ID3 algorithm

- The ID3 can decide the best attribute by using the statistical property information gain. The gain measures how the attributes separates the training examples into target classes.
- ID3 builds the shortest and fastest tree.
- Easy prediction rules can be generated from training data.

III. FLOW CHART

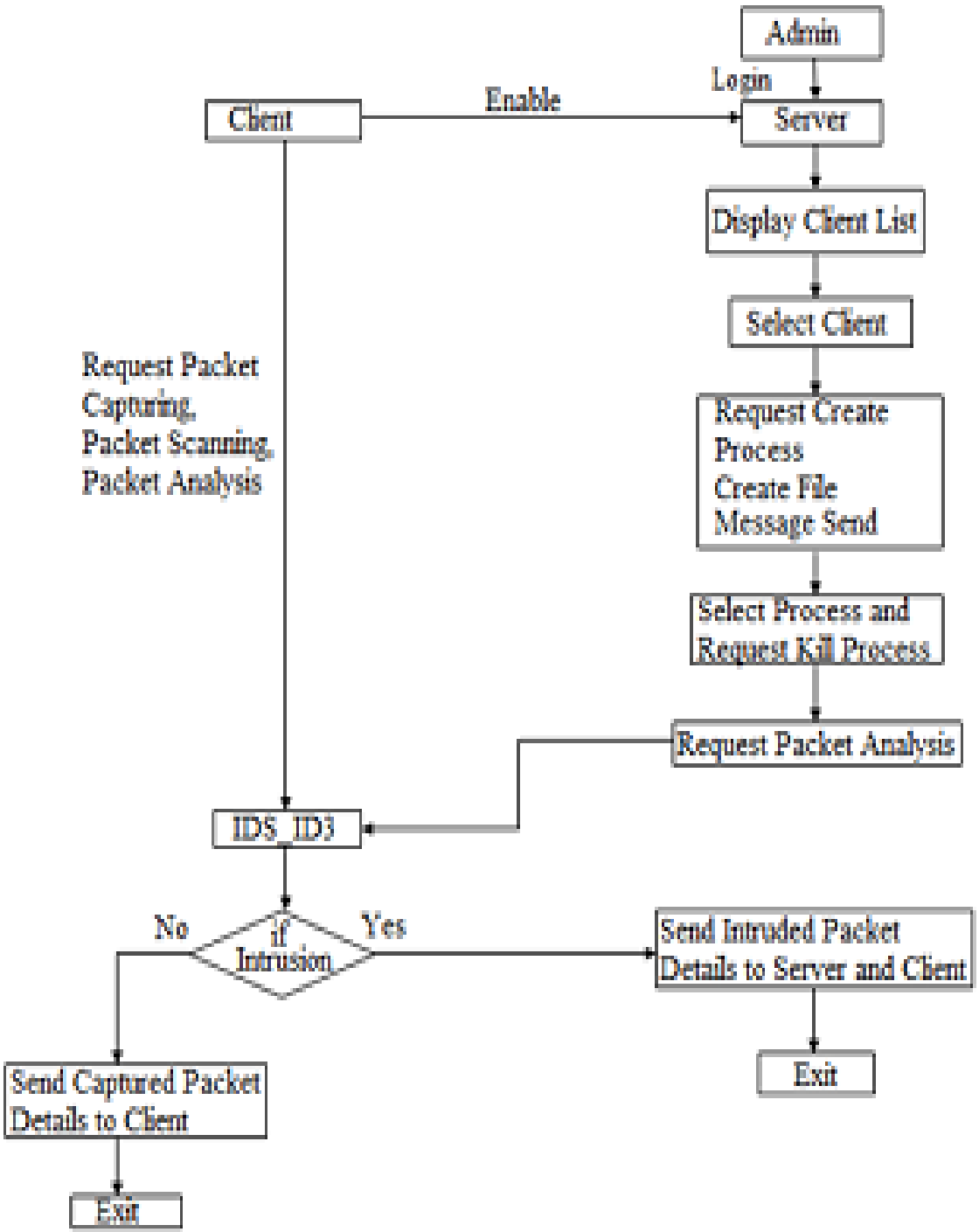

Figure 3.Flow diagram of the IDS model

IV. CONCLUSION AND FUTURE ENHANCEMENT

Thus, the idea behind this intrusion detection system is important to ensure data security of the confidential data that is transferred via a network. The security agencies can use the system to provide data security to the organizations.

- Provision can be done to include the Intruded packet's source IP address in a blacklisted IP list and further block those IPs
- MAC address of the intruded packets can be detected because even if the IP address changes the MAC address remains the same and it can be permanently blocked

REFERENCES

[1] Anant R. More, Vikas N. Nandgaonkar, Dr. Manoj Nagmode, Pramod P. Patil, "ID3 Algorithm for Intrusion Detection."International Conference on Recent Trends in engineering & Technology- 2013(ICRTET'2013).
[2] Sandeep Kumar, Prof Satbir Jain, "Intrusion Detection and Classification Using Improved ID3 Algorithm of Data Mining."International Journal of Advanced Research in Computer Engineering & Technology Volume 1, Issue 5, July 2012.
[3] Barbara, Daniel, Couto, Julia, Jajodia, Sushil, Popyack, Leonard, Wu, and Ningning, "ADAM: Detecting intrusion by data mining," IEEE Workshop on Information Assurance and Security, West Point, New York, June 5-6, 2001.
[4] MohammadrezaEktefa, Sara Memar, Fatimah sidi , Lilly Suriani Affendey" Intrusion Detection using Data Mining Technique"IEEE 2010.
[5] Dewan Md. Farid, Nouria Harbi, Emna Bahri, Mohammad ZahidurRahman, Chowdhury Mofizur Rahman "Attacks Classification in Adaptive Intrusion Detection using Decision Tree" 2010.

