IDS with Hybrid ID3 Algorithm

Suman Singh
Computer Science Department,
SEC, Sikar
RTU,Kota, Rajasthan, India

ABSTRACT

An intrusion detection system (IDS) is a device or software application that monitors network and/or system behavior for malicious activities or policy violations and produces reports to a Management Station. In this project we design and implement an IDS that is a software application and monitor network through a client server approach and to detect users activity we can use a process monitor for intruder detection. Here client application is gather system process and update information over server. Server contains the database and an intelligent algorithm to classify the process patterns. If classified data belongs to the previously detected attack then generate alarm and if the process pattern is not classified as previously then that means it is a new kind of attack and update the current database.

General Terms

Pattern Recognition, C4.5, ID3.

Keywords

IDS, process monitor, patterns, alarm, database. engines.

1. INTRODUCTION

An intrusion detection system (IDS) inspects all inbound and outbound network activity and identifies suspicious patterns that may indicate a network or system attack from someone attempting to break into or compromise a system. An IDS differs from a firewall in that a firewall looks out for intrusions in order to stop them from happening. The firewall limits the access between networks in order to prevent intrusion and does not signal an attack from inside the network. An IDS evaluates a suspected intrusion once it has taken place and signals an alarm. Working with all kind of system and is much complex, time consuming and expensive work for us, thus in this project we work with first kind of system and modify it to make the IDS more powerful and achieve better performance. Moreover it in our project we include the new pattern detection for newly found attack using a decision tree.

Problem Definition

In the field of computer science the network has its own importance and it is also area of research interest. Thus design and implementation of IDS and such kind of system is a traditional work when the computer networks are being introduced. There are various tools, technique and methodologies are proposed and implemented. But most of them are based on the database matching methodologies. Required to design a new kind of IDS using the process mining of different network computers And identify the intruders using an intelligent decision making and self-learning technique by which system automatically train itself and capable to classify legitimate and anomaly detections.

Solution Domain

To resolve the need of new system we propose a new system and its internal architecture to demonstrate the working and IDS system we involve the following work under design and development of system.

1. design a new system architecture for intelligent IDS
2. find the better effective and efficient nominal data set classifier
3. Implement with the IDS to classify the current or real time data.
4. Performance study of our designed system.

System Architecture

As we can see in the below given figure 1 there are two different and individual units are required to work. First client application resides over the client machine and gathers client machine running process and sends it to the server end. Server is another program designed to collect all information send by the client and update it over the database to generate the data for evaluation. Moreover it server contains the algorithm which train over the updated database and produce the messages for current found records. With their IP address.

2. LITERATURE SURVEY

Objective Formulation: After study of different research papers related to pattern recognition, data mining and IDS we decide our goals and guidelines to develop and deploy our complete system to protect the network computers from unwanted intrusion attacks.

1. Find the transparent and efficient algorithm to classify data and capable to make decisions.
2. Performance study of algorithms which is best fit for our application
3. Design a server with multi-threaded architecture and able to communicate all connected clients
4. Design a client application which is easily connect to the server and send the current activity related to local machine
5. Design an intelligent data base management system by which algorithm automatically train itself to test real time data
6. Classify newly appeared attacks and generate alarms when attack found

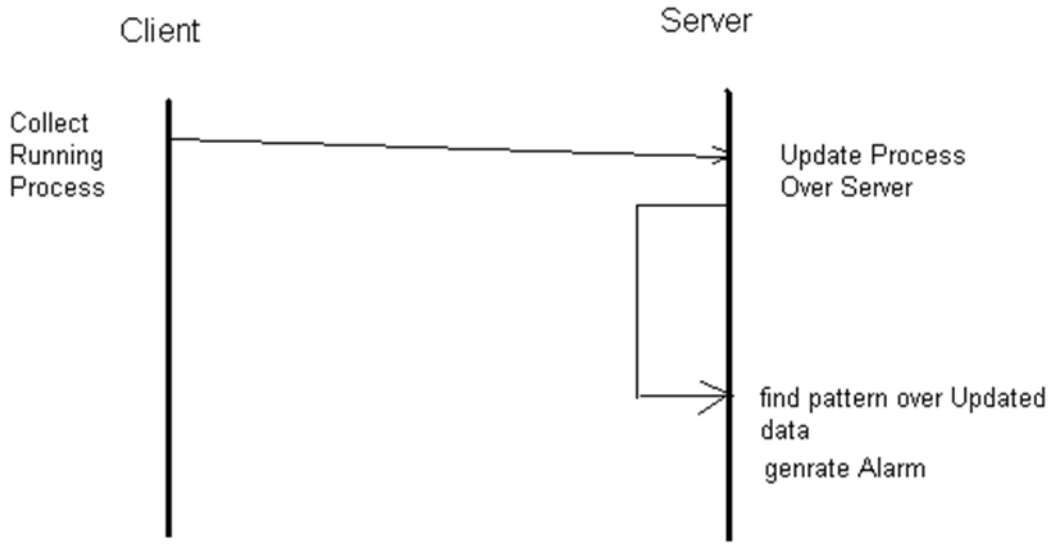

Figure 1: shows the basic working of system

3. RESULTS

In this project we include first performance comparison of two data mining algorithms C 4.5 and ID3 by which we select the appropriate algorithm for our application. The used data in our application is nominal for that purpose we evaluate results on the basis memory used, accuracy, and others.

3.1 Performance Comparison:

3.1.1 Accuracy

Accuracy of the system over nominal data is given the below table.

Table 1. Accuracy of the system over nominal data

Data set Size	C4.5	ID3
100	79.49	77.90
300	72.32	78.83
500	83.29	80.92
700	74.17	79.76
1000	88.27	81.22

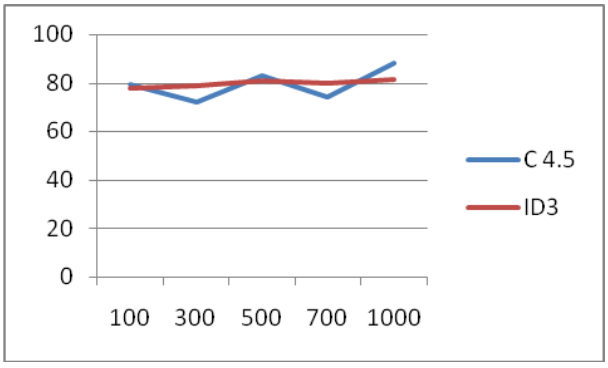

Diagram 1: Graph of Table1

As we can see in the above given results the performance of C4.5 is much fluctuating mean to say according to the data it varies to much thus it is not suitable for of our application and on other hand ID3 gives the smooth results and most of the time performance is constant thus it is much suitable for our application.

3.1.2 Memory used

The memory uses of the designed system is given below

Table 2. memory uses of the designed system

Data set Size	C4.5	ID3
100	23836	24032
300	24776	25168
500	26464	26464
700	27108	27536
1000	29348	29348

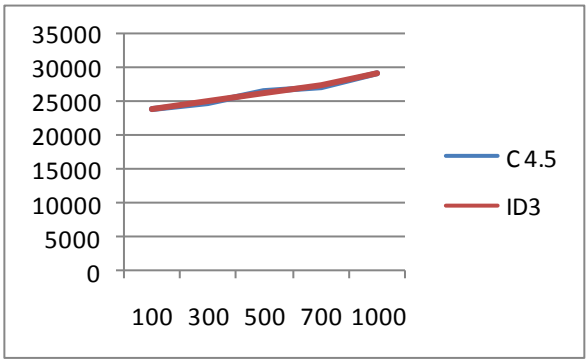

Diagram 2: Graph of Table2

As we can see in the above given results both the algorithm simulate the same behavior but the performance of ID3 is little bit poorer then the C4.5

3.1.3 Build time of system

Table 3. Build time of system over nominal data

Data set size	C 4.5(seconds)	ID3 (seconds)
1000	3.273	4.155
700	1.629	2.582
500	1.284	2.114
300	1.732	1.25
100	1.539	1.84

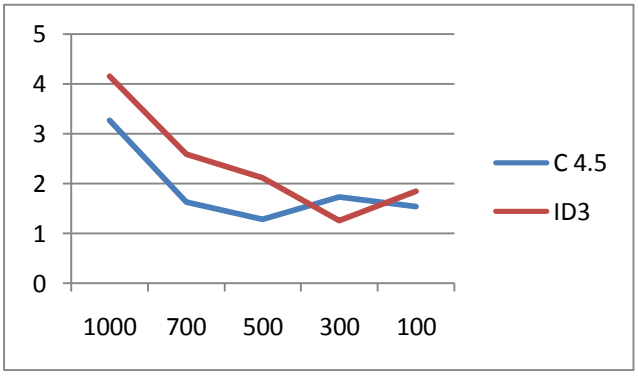

Diagram 3: Graph of Table3

In this part we simulate the build time of the model, by the archived below given results we can say C4.5 consumes less time to build model and Id3 take more time to form this model.

3.1.4 Search time

Here we provide the search time for the model

Table 4. Search time for model over data

Data set size	C4.5 (seconds)	ID3(seconds)
1000	.2642	.261
700	.1662	.527
500	.1642	.527
300	.142	.229
100	.242	.103

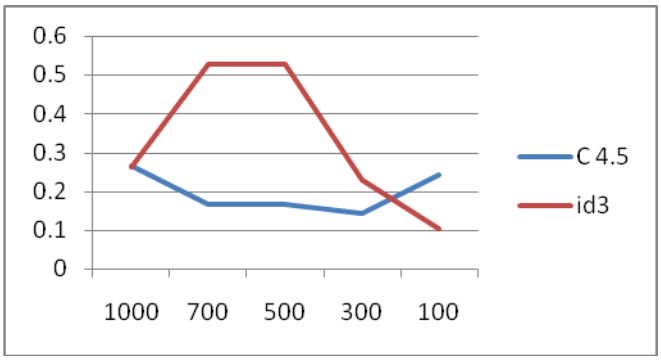

Diagram 4: Graph of Table4

The above given results shows the search time of the designed system the search time of Id3 is more than C4.5

4. CONCLUSION

In our result section we are include the performance of both system C4.5 and ID3, but in this era of computer science the accuracy of classification is main need thus we use ID3 algorithm for implementing the IDS.

Thus we can say we complete our desired to design IDS using Process mining and use of data mining algorithm ID3. And we are able to detect the anomaly detection using this. The designed system is a semi-automatic. And help required connecting the system in network

Advantage of the System

1. The performance of system at the client side and server side is distributed equally. Thus not complete load is depends on server side.

2. Use of ID3 algorithm is performing better to classify the data which is arrived at real time.

Disadvantage of System

1. The intrusion detection is based on process mining and monitoring that is a system dependent factor

2. At the client side system also consumes memory to collect information and send it serve

5. REFERENCES

[1] An Artificial Immune Model for Network Intrusion Detection, Jungwon Kim and Peter Bentley Department of Computer Science, University Collge London Gower Street, London, WC1E 6BT, U. K. Phone: +44-171-380-7329, Fax: +44-171-387-1397 email:{J.Kim, P.Bentley}@cs.ucl.ac.uk

[2] Testing Network-based Intrusion Detection Signatures Using Mutant Exploits, Giovanni Vigna vigna@cs.ucsb.edu William Robertson wkr@cs.ucsb.edu Davide Balzarotti balzarot@cs.ucsb.edu Reliable Software Group University of California, Santa Barbara Santa Barbara, CA 93106

[3] Training a Neural-Network Based Intrusion Detector to Recognize Novel Attacks Susan C. Lee and David V. Heinbuch, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 31, NO. 4, JULY 2001

[4] A Neural Network Based System for Intrusion Detection and Classification of Attacks Mehdi MORADI and Mohammad ZULKERNINE

[5] The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors Daejoon Jooa,*, Taeho Hongb,1, Ingoo Hanc,2

[6] Anomaly-based network intrusion detection: Techniques, systems and challenges P. Garcı´a-Teodoro a ,*,J.Dı´az-Verdejo a, G. Macia´-Ferna´ndez a,E.Va´zquez b a. Department of Signal Theory, Telematics and Communications – Computer Science and Telecommunications Faculty, University of Granada, Granada, Spain b. Department of Telematic Engineering - Universidad Polite´cnica de Madrid, Madrid, Spain

[7] Neural Network Based Intrusion Detection System for Critical Infrastructures, Ondrej Linda, Todd Vollmer, Milos Manic, Member, IEEE

[8] A hybrid intrusion detection system design for computer network security, M. Ali Aydın *, A. Halim Zaim, K. Gökhan Ceylan Department of Computer Engineering, Faculty of Engineering, Istanbul University, 34320 Avcilar, Istanbul, Turke

[8] A Simple and Efficient Hidden Markov Model Scheme for Host-Based Anomaly Intrusion Detection, Jiankun Hu and Xinghuo Yu, RMIT University D. Qiu, Sai Global Limited Hsiao-Hwa Chen, National Cheng Kung University

[9] Agent-based Intrusion Detection for Network-based Application Jianping Zeng1 and Donghui Guo2 (Corresponding author: Donghui Guo) Department of Physics, Xiamen University, Fujian 361005, China1 Department of Electronic Engineering, Xiamen University, Fujian 361005, China2 (Email: dhguo@xmu.edu.cn)

[10] Artificial Neural Network based Intrusion Detection System: A Survey Bhavin Shah Associate Professor, MCA Programme L. J. Institute of Management Studies Ahmedabad, India. Bhushan H Trivedi, PhD Director GLS Institute of Computer Technology Ahmedabad, India. International Journal of Computer Applications (0975 – 8887) Volume 39– No.6, February 2012

[11] Intrusion Detection Technique in Mobile Adhoc Network based on Quantitative Approach Saroj Hirnwal Professor Balaji College of Engineering Technology, Jaipur (Rajasthan) Kirti Chauhan M.Tech Scholar Balaji College of Engineering Technology, Jaipur (Rajasthan) Amit Gupta Asst. Professor Balaji College of Engineering Technology, Jaipur (Rajasthan), International Journal of Computer Applications (0975 – 8887) Volume 37– No.8, January 2012

[12] IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.8, August 2011 Manuscript received August 5, 2011 Manuscript revised August 20, 2011 **Implementing Rule based Genetic Algorithm as a Solution for Intrusion Detection System** Shaik Akbar† Dr.K.Nageswara Rao†† Dr.J.A.Chandulal†††, † Assoc. Prof, Dept. of C.S.E SVIET, Nandamuru, Krishna Dist, Andhra Pradesh, India †† Prof & H.O.D, Dept. of C.S.E P.V.P.S.I.T, Vijayawada, Krishna Dist, Andhra Pradesh, India ††† Prof, Dept. of C.S.E GITAM University, Visakhapatnam, Andhra Pradesh, India