Intrusion Detection System using Bayesian Approach for Wireless Network

Manoj Sharma
Assistant Professor
Royal Institute Of Management &Techonology
Gohana,Haryana(India)

Keshav Jindal
Assistant Professor
Royal Institute Of Management & Techonology
Gohana,Haryana(India)

Ashish Kumar
Assistant Professor
Royal Institute Of Management & Techonology
Gohana,Haryana(India)

ABSTRACT
An Intrusion Detection System (IDS) is a software or hardware tool used to detect unauthorized access of a computer system or network. A wireless IDS performs this task exclusively for the wireless network. These systems monitor traffic on your network looking for and logging threats and alerting personnel to respond. An IDS usually performs this task in two ways, with either signature-based or anomaly based detection. Almost every IDS today is at least in part signature-based. This means that known attacks can be detected by looking for these signatures. The other approach is anomaly-based systems. These are not often implemented, mostly because of the high amount of false alarms. It detects traffic which deviates from what it considers normal an alert is generated. The traditional IDS system is not work well for wireless network, but the wireless network is more vulnerable than a wired network. A major problem with current IDS that employs Bayesian network is they give a series of false alarms in system environment modification. There are two types of false alarms in determining the any deviations from normal pattern: false positive and false negative. The main goal is to keep these alarms as low as possible. So a BN is used to build automatic intrusion detection system based on signature recognition. The goal is to recognize signatures of known attacks, match the observed behavior with those known signatures, and signal intrusion when there is a match. IDS must be able to adapt to these changes. The goal is to provide a framework for an adaptive intrusion detection system that uses Bayesian network

Keywords
Intrusion Detection System, Bayesian Network, Directed Acyclic Graph, K2.

1. INTRODUCTION
A computer system should provide confidentiality, integrity and assurance against denial of service. However, due to increased connectivity (especially on the Internet), and the vast spectrum of financial possibilities that are opening up, more and more systems are subject to attack by intruders. These subversion attempts try to exploit flaws in the operating system as well as in application programs and have resulted in spectacular incidents like the Internet Worm incident of 1988 [1]. There are two ways to handle subversion attempts. One way is to prevent subversion itself by building a completely secure system. We could, for example, require all users to identify and authenticate themselves; we could protect data by various cryptographic methods and very tight access control mechanisms. However this is not really feasible because:

1. It is difficult to build a completely secure system. Miller gives a compelling report on bugs in developing fault free software is quiet difficult. Designing and implementing a totally secure system is thus an extremely difficult task.

2. The vast installed base of systems worldwide guarantees that any transition to a secure system, (if it is ever developed) will be long in coming.

3. Cryptographic technique have their own problems, password and secret information can be hacked by unauthorized users.

4. Even a truly secure system is vulnerable to abuse by insiders who abuse their privileges.

5. It has been seen that that the relationship between the level of access control and user efficiency is an inverse one, which means that the stricter the mechanisms, the lower the efficiency becomes.[2]

The history of security research has taught us a valuable lesson – no matter how many intrusion prevention measures are inserted in a network, there are always some weak links that one could exploit to break in. If there are attacks on a system, we would like to detect them as soon as possible and take appropriate action. This is essentially what an Intrusion Detection System (IDS) does. An IDS is a reactive rather than pro-active agent..

2. TYPES OF IDS
There are several types of IDSs available today, characterized by different monitoring and analysis approaches. Each approach has distinct advantages and disadvantages. Furthermore, all approaches can be described in terms of a generic process model for IDSs.Here we describe some techniques of intrusion detection

2.1 Anomaly Detection
Anomaly detection techniques establish a "normal activity profile" for a system; we could, in theory, flag all system states varying from the established profile by statistically significant amounts as intrusion attempts [3]. However, if we consider that the set of intrusive activities only intersects the set of anomalous activities instead of being exactly the same, we find a couple of interesting possibilities:

(1) Anomalous activities that are not intrusive are flagged as intrusive.

(2) Intrusive activities that are not anomalous result in false negatives (events are not flagged intrusive, though they actually are).

2.2 Misuse Detection
It uses a pre known signature or pattern to compare with incoming traffic. In the signature detection there are several methods to detect the intrusion patterns. The detection

approaches, such as expert system [4], pattern recognition [5], are grouped on the misuse. The concept behind misuse detection is that these systems are not unlike virus detection systems -- they can detect many or all known attack patterns, but they are of little use for as yet unknown attack methods. An interesting point to note is that anomaly detection systems try to detect the complement of "bad" behavior. Misuse detection systems try to recognize known "bad" behavior. The main issues in misuse detection systems are how to write a signature that encompasses all possible variations of the pertinent attack, and how to write signatures that do not also match non-intrusive activity

2.3 Network Based Intrusion Detection

The most obvious location for an intrusion detection system is right on the segment being monitored. Network-based intrusion detectors insert themselves in the network just like any other device, except they promiscuously examine every packet.

2.4 Host Based Intrusion Detection

Host based IDS exploit vulnerabilities particular to specific operating systems and application suites. Only host-based intrusion detection systems (the ones running as an application on a network-connected host) can correlate the complex array of system-specific parameters that make up the signature of a well-orchestrated attack.

3. METHODOLOGY

- We are performing the intrusion detection based on Bayesian network theorem. It is the probabilistic approach to detect the intrusion.
- We combine three algorithms k2, Bayesian and junction tree inference.
- K2 will perform the State Space for learning process, Bayesian will take probabilistic decisions and junction tree is the analogy to generate the efficient tree after intrusion detection.

Bayesian techniques are to create a plan of goal-directed actions. An event classification scheme is proposed based on Bayesian networks. Bayesian networks improve the aggregation of different model outputs and allow one to seamlessly incorporate additional information. [6]

3.1 Modules:

1. K2 Algorithm
2. Bayesian Recognition
3. Junction Tree Inference

Modules Descriptions:

A. K2 Algorithm:

In this we are dealing with incomplete records in the database so we opted for the Bayesian approach and particularly for the K2 algorithm [7]. K2 learning algorithm showed high performance in many research works.

Algorithm K2 used in learning step needs:

1. A given order between variables
2. The number of parents, u of the node.

K2 algorithm [8] proceeds by starting with a single node (the first variable in the defined order) and then incrementally adds connection with other nodes which can increase the whole probability of network structure, calculated using the g function. A requested new parent which does not increase node probability cannot be added to the node parent set.

B. Bayesian Recognition:

Bayesian methods utilize a search-and-score procedure to search the space of DAGs, and use the posterior density as a scoring function. There are many variations on Bayesian application [9] like greedy heuristic, combined with techniques to avoid local maxima in the posterior density (e.g., greedy search with random restarts or best first searches).Bayesian approaches are capable of dealing with incomplete records in the database [10]. The most serious drawback to the Bayesian approaches is the fact that they are relatively slow.

C. Junction Tree Inference

The idea of this procedure is to construct a data structure called a junction tree which can be used to calculate any query through message passing on the tree [11]. The first step of JT algorithm creates an undirected graph from an input DAG through a procedure called moralization. Moralization keeps the same edges, but drops the direction, and then connects the parents of every child.

3.2 Proposed Intrusion Detection System Architeture

We propose a framework for an intrusion detection system using Bayesian network which combines k2 learning process, Bayesian Recognition and Junction Tree. These all are present in our framework, it starts from k2 learning process and completes at Junction Tree. This intrusion detection system use testing and training dataset to detect intrusion in a network[12]. The training dataset can be updated by adding new intrusions signatures. Training dataset contains signatures of normal connections and signatures of several types of known attacks

Now we describe how framework starts and perform intrusion detection over a intrusion detection dataset. Working is described in six steps that are given below:

- Training data set is uploaded in our Bayesian IDS framework which contains normal signatures or connections and signature of known attacks. It is basically used to train the framework so that it can work on a testing data set.
- Testing data set is uploaded in our Bayesian IDS framework which contains thousands of computer connections. We have to detect the intrusion in this dataset.
- K2 learning process lists the values of different features of computer connections that are present in a testing dataset.
- Bayesian recognition uses these values for classifying the chances of attacks. It lists the output according to two classes: normal/anomaly and also with their corresponding parameter values for different features of all the computer connections present in a testing data set.
- Junction Tree produces the final result in the form of records. It mainly uses the result of Bayesian Recognition and produce output according to the actual and predicted classes for each host with respective ID numbers.

ARCHITECTURE OF PROPOSED INTRUSION DETECTION SYSTEM

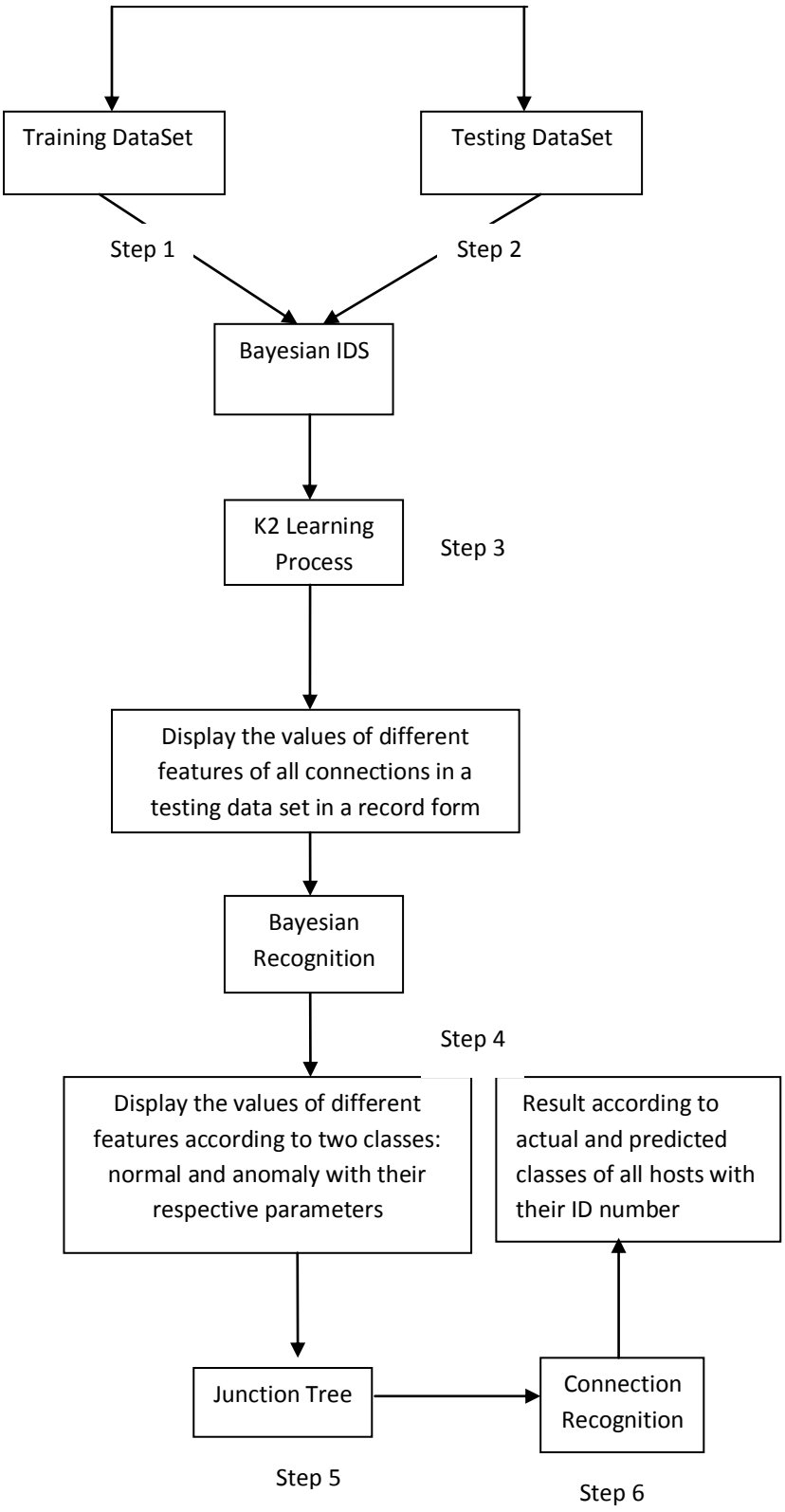


Fig 1: Architecture of Proposed Intrusion Detection System

3.3 .INTRUSION DETECTION DATASET

Dataset represent the data as rows of TCP\IP dumps where each row consist of computer connection .Packet information in TCP dump file is summarized into connections. A connection is a sequence of TCP packet starting and ending at some defined time and data flows between source IP address and target IP address under well define protocol [Kayacik, G et al (2005)]. or normal connection).

Each computer connection has 41 features and these features are grouped in to four categories:

1. Basic Features Basic features can be derived from packet header without inspecting payload
2. Content Features These features are used to access the payload of original TCP packet
3. Time Based Traffic Features These features are designed to capture the properties that mature over a 2 second temporal window.
4. Host Based Traffic Features These features utilizes historical windows estimated over a period of time.

4. RESULTS AND DISCUSSION

4.1 K2 Process Results

Once we start running proposed system first window will appear on screen .The Process start with first selecting training and testing datasets, they are uploaded in to system. Once Training Dataset Contain Normal connection signature and database of pre

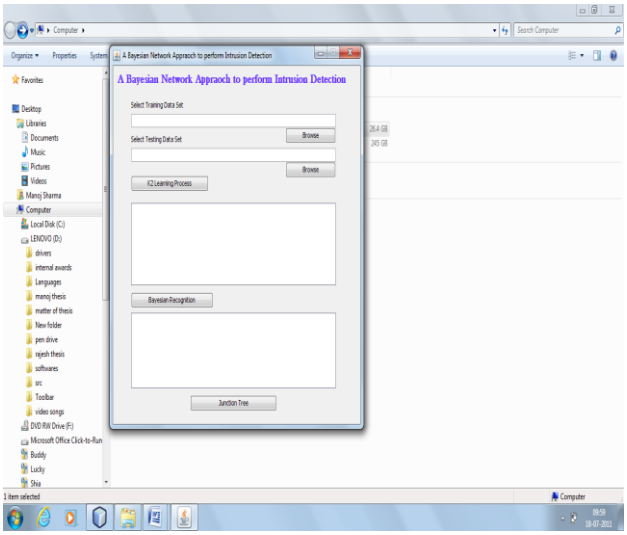
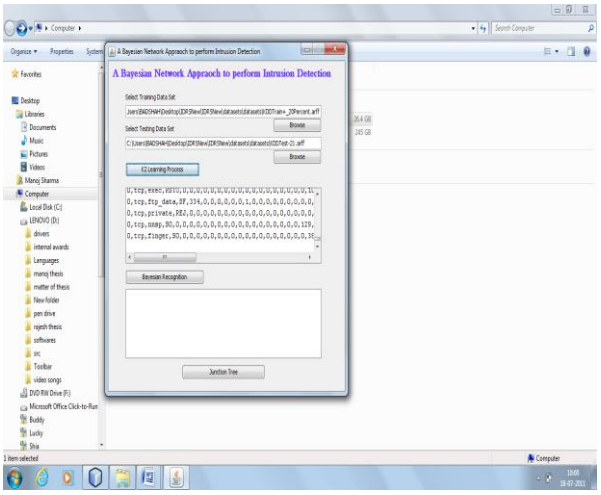


Figure 2: User Interface of Proposed Intrusion Detection System



known signature that can be updated if a new intrusion occurs, Testing Dataset contain the thousands of TCP/IP dumps or connection which we have to determine whether they are protected from any kind of threat or not. Now K2 learning process starts which gives the values of different features of a all connections present in a Testing Dataset as it is clearly visible on screen.

4.2 Bayesian Recognition and Junction Tree

Bayesian recognition classify the systems behavior in to two classes

- Normal
- Anomaly

Normal indicates that system is protected from any kind of attacks, and anomaly means that something happen wrong with system it means some kind of attack is made by intruders. It also provide some parametric value for all features that are present in intrusion dataset, these values are very helpful in determining to which extent attack is made. The result of Bayesian recognition is a input to construct a junction tree.

Junction tree produce a final outcome which contain the information about the all connection or nodes respective with their Id no and it list the difference between actual and predicted class help us in ensuring whether the prediction was made correct or incorrect. By this means we can able to determine that attack is made or not. It is the last module of our system which starts working by constructing a network from DAG i.e. created by K2 learning process and also get the results of Bayesian recognition to determine the unauthorized access or applications.

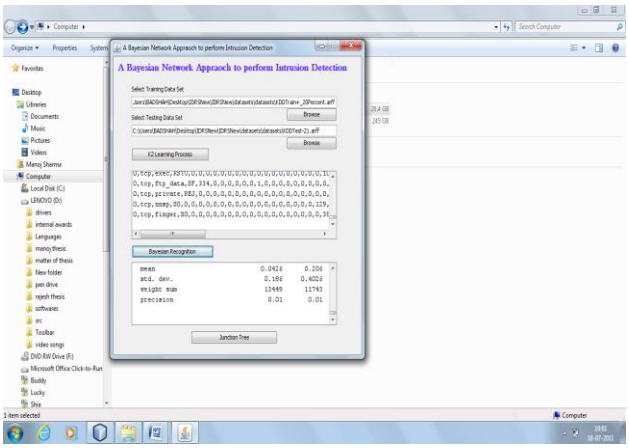


Figure 3 : Bayesian Recognition Process Results

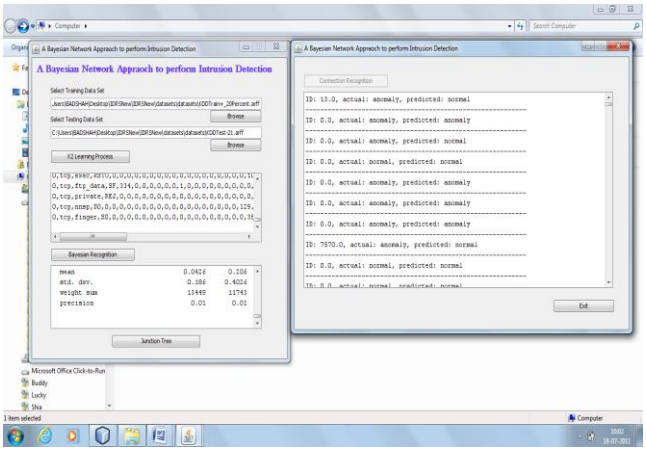


Figure 4 : Junction Tree Process Results

5. CONCLUSION

We outlined a framework for an adaptive intrusion detection system using Bayesian network. Bayesian networks provide automatic detection capabilities; they learn from audit data and can detect both normal and abnormal connections. Our system demonstrated a high performance when detecting Intrusions. This system can be improved by integrating an expert system which is able to provide recommendations based on attack types.

6. REFERENCES

[1] W Arbaugh., N. Shankar, Wan Y.C.J., "Your 802.11 Wireless Network Has No Clothes", University of Maryland, Mar. 2001.

[2]. R. Kumar, Isukapalli, V. Karunya, V. Raju, "Security in Mobile Computing Systems." INTRUSION DETECTION IN WIRELESS NETWORKS.

[3]. J. Krister and L. Stephen. "Network Security: Bayesian Network."

[4]. T. F. Lunt, R. Jagannathan, "IDES: The Enhanced Prototype C a Realtime Intrusion-Detection Expert System". Technical Report SRI-CSL-88-12, SRI International, Menlo Park, CA, 1988. Intrusion Detection (BNIDS) May. 2003.

[5]. M. Esposito, C. Mazzariello, "Evaluating Pattern Recognition Techniques in Intrusion Detection Systems". The 7th International Workshop on Pattern Recognition in Information Systems, pp. 144-153, 2005.

[6]. R. Goldman, "A Stochastic Model for Intrusions." In Symposium on Recent Advances in Intrusion Detection (RAID), 2002.

[7]. D.M. Chickering, "Learning Equivalence Classes of Bayesian Network Structure", Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, Reed College, Portland, Oregon, USA, pp. 150-157, 1996.

[8]. G.F. Cooper, "An overview of the representation and discovery of causal relationships using Bayesian networks", AAAI Press and MIT Press, pp. 3-62, 1999.

[9]. J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference". Morgan Kaufmann, 1997.

[10]. P. Spirtes, C. Glymour, R. Scheines, "Causation, Prediction and Search (Second Edition)", MIT Press, Cambridge, MA, USA, 2000.

[11]. N. Friedman, D. Koller, "Being Bayesian About Network Structure: A Bayesian Approach to Structure Discovery in Bayesian Networks", Machine Learning 50 (1-2), pp. 95-125, 2000.

[12]. F .Jemili, M. Zagdoud, "A Framework for an Adaptive Intrusion Detection System using Bayesian Network"

Monuba University Tunisia, 2010