Causal Discovery and Reasoning for Intrusion Detection using Bayesian Network

Yit Yin Wee, Wooi Ping Cheah, Shing Chiang Tan and KuokKwee Wee

***Abstract*—Computer security is essential in information technology world today; confidentiality, availability and integrity of data are the aspects concerned. Firewall has been widely deployed as a protection but it is no longer adequate to against the intelligent intrusions and attacks which keep changing and transforming. A network intrusion detection and analysis system has been introduced in this paper to resolve the problems of data confidentiality, availability and integrity. The challenge of the study is; first, to model the network intrusion detection domain and second, to perform causal reasoning for intrusion detection and analysis based on the domain model constructed earlier. In this paper, a methodology has been proposed to resolve the two problems mentioned above. Both problems will be addressed under causal knowledge driven approach where intrusion detection is viewed as fault diagnosis and prognosis processes. We have proposed Bayesian network for the modeling of network intrusion domain. Also, powerful reasoning capabilities of Bayesian network have been applied to discover intrusion attacks. Since the capabilities of causal reasoning using Bayesian network have not been fully discovered in the domain of intrusion detection by most of the researchers before, this research work is to bridge the gap. From the results of the experiment, we have concluded that the capability of Bayesian learning is reasonably accurate and efficient.**

***Index Terms*— Bayesian network, causal discovery, causal reasoning, intrusion detection, soft computing.**

I. Introduction

A network intrusion detection and analysis system tries to detect and analyze the impacts of malicious activity such as denial of service attacks, port scans or even attempts to crack into computers by monitoring network traffic [1]. Construct a model of network intrusion and to use it for subsequent detection and analysis is the major problem in this domain. To fulfill this aim, many attempts and methods have been proposed. Although, some of the modern methods could help in some way, they are still being improved and updated; no methodology has ever claimed to provide a general purpose solution for intrusion detection and analysis.

Manuscript received March 29, 2011. This research is supported by the GRA Grant funded by Multimedia University, Malaysia (Project ID IP20110105018).

Y.Y Wee is with the Multimedia University, Jalan Ayer Keroh Lama 75450 Bukit Beruang Melaka Malaysia. (e-mail: yywee@mmu.edu.my).

W.P Cheah is with the Multimedia University, Jalan Ayer Keroh Lama 75450 Bukit Beruang Melaka Malaysia. (e-mail: wpcheah@mmu.edu.my).

S.C Tan is with the Multimedia University, Jalan Ayer Keroh Lama 75450 Bukit Beruang Melaka Malaysia. (e-mail: sctan@mmu.edu.my).

K.K Wee is with the Multimedia University, Jalan Ayer Keroh Lama 75450 Bukit Beruang Melaka Malaysia. (e-mail: wee.kuok.kwee@mmu.edu.my).

The difficulties faced in network intrusion are to model the domain of intrusion and the reasoning on the model. There are two ways of constructing a domain model. In the knowledge engineering approach, domain experts in collaboration with a knowledge engineer identify the relationships between domain variables manually. This approach is optimal when the number of variables is reasonably small and controllable manually. It is problematic when we are facing with a large number of variables, and very often, it is completely impractical as the complexity of the problem grows exponentially with the number of variables. In the data mining approach, the domain model is derived automatically by using an algorithm that will learn it from the network intrusion data. This approach will reduce human effort in the construction of the model. However, the success of data mining approach heavily relies on the availability of a huge set of data. Data mining approach is very appropriate to intrusion detection and analysis as it is a data rich domain. One example is the KDD Cup 1999 (computer network intrusion detection data set), which contains 41 variables and up to a staggering 4 million records. Causal reasoning is about diagnosing the root cause(s) and predicting the effect of the intrusion. In this paper, a causal knowledge-driven approach is adopted. Although this approach is widely used in other domains such as medical and mechanical diagnosis, there is limited application in the domain of intrusion detection and analysis. This paper is to bridge the gap.

Bayesian network has been proposed to solve both the modeling and reasoning problems. Supported by powerful learning algorithm, Bayesian network serves as a good modeling tool for a data rich domain like intrusion detection [2]. Besides, it also provides an efficient evidence propagation mechanism and powerful reasoning capability. The availability of several powerful commercial level tools for learning Bayesian network from data has made it a practical modeling framework for causal reasoning. In a data rich domain like intrusion detection and analysis the availability of learning tools will greatly reduce the cost and effort in constructing Bayesian causal models. Moreover, many of these tools have been commercialized into today's market, such as Hugin[3] and Netica[[4], making it a more mature framework for causal knowledge modeling and reasoning.

In the following, Section II discusses background of intrusion detection and analysis. Section III discusses the fundamental, inference mechanism and learning algorithm of Bayesian network. In Section IV, a methodology has been described and experimental results have been discussed. The conclusion of the paper has been included in Section V.

II. Intrusion Detection and Analysis

In order to cope with the growing trend of intrusion, Intrusion Detection System (IDS) was established for the purpose of malicious activities detection to strengthen the security, confidentiality, and integrity of critical information systems. IDS are very popular security tool in the last two decades, and today, IDS based on computer intelligent are attracting a lot attention from current research community [5]. It has several advantages that can be summarized as follows [6]:

- It can detect external hackers as well as internal network-based attacks.
- It scales easily to provide protection for the entire network.
- It offers centralized management for correlation of distributed attacks.
- It provides defense in depth and gives the system administrators the ability to quantify attacks.
- It provides an additional layer of protection.

Many commercial products regarding intrusion detection are introduced in today's market to detect the known attack, and many researches today are still looking for new unknown signatures. However, there is still challenge to be faced in intrusion detection research. The main challenge that intended to be solved by the researches from past decades until now is the false alarm rate problem. False negative will definitely cause huge damage to the system.

- *False positive*: administrator believes there is an intrusion and raises the alarm but actually there is none.
- *False negative*: administrator did not raise an alarm when there was an intrusion.

Although there are many types of IDS, Host-based Intrusion Detection System (HIDS) and Network-based Intrusion Detection System (NIDS) are the two main types. HIDS is used to analyze the internal event such as process identifier while NIDS is to analyze the external event such as traffic volume, IP address, service port and others.

A. Intrusion Detection Method

Intrusion detection comprises of three main techniques, which are signature-based intrusion detection, anomaly-based intrusion detection, and hybrid intrusion detection.

Signature-based intrusion detection requires constant updates on their database because it heavily relies on predefined set of attack signatures. Signature-based intrusion detection works by monitoring packets of network and compares them against a database of signatures or attributes from known malicious threats. One of the advantages for this type of IDS is it produces a low rate of false positive alarm. Unfortunately, the signature-based intrusion detection needs a set of signatures for possible attacks to be defined in advance as it is not capable to detect new intrusion events and it can only detect previously known attacks. Therefore, they must be constantly updated with the signature of new attacks [7].

Anomaly-based intrusion detection differs from signature-based intrusion detection as it creates a base line profile of the normal system, thus providing it the capability to distinguish the incoming system activity to be either normal or anomalous. If the activities are found to be anomalous, an anomaly alarm will be generated by the detection system. Several benefits are offer by anomaly intrusion detection. The main benefit of Anomaly intrusion detection is able to detect the previously unknown attack. However, the drawback of anomaly intrusion detection is it generate high rate of false alarm. This can make the system unusable by flooding and eventually desensitizing the system administrator with large numbers of incorrect alerts [8].

A hybrid or combination of anomaly-based approach and signature-based approach are also used in the present IDS. The signature technique detects known attack while the anomaly technique aids in the detection of new and unknown attacks. Even so, the hybrid systems are not always better compared to stand alone detection method as the efficiency and effectiveness of the detection system also relies on other parameters or conditions [9].

B. Anomaly Intrusion Detection

Anomaly intrusion detection is categorized into three main techniques: feature selection, categorization and causal reasoning.

1) Feature selection

As the data to be processed for intrusion is very huge even for a small network, feature selection technique is used as it enables the user to identify the important features/input and at the same time eliminating insignificant input. With this, the processing time required and the storage space utilized for detection can be reduced significantly. This will yield a more efficient and effective result [10].

2) Categorization

In order to increase the effectiveness of anomaly intrusion detection, categorization of data is required. Generally the categorization can be divided into classification and clustering, the former being a process that categorize the types of attack using supervised data and the latter using unsupervised data. Statistical, knowledge-based and machine learning approaches are the method used to categorize the data. Although categorization technique is important for intrusion detection, however it is too restrictive as it does not provide comprehensive reasoning capability.

3) Causal reasoning

Causal reasoning is a process capable of identifying any cause(s) leading to certain effect(s) and the causal relationships among various events. Before causal reasoning can be initiated, the structure of the model must first be constructed. Knowledge engineering or data-mining approach can be used to model the structure required for causal reasoning process. Causal reasoning is a comprehensive process that includes the feature selection, classification and also diagnosis and prognosis. However, studies show that there are limited research and work done on the effectiveness of causal reasoning in intrusion detection. A proposal using causal mapping approach to establish a systematic procedure for constructing Bayesian network from domain knowledge of experts was tabled in [7]. However, it is not in the intrusion detection domain. Furthermore, due to huge datasets the knowledge engineering approach is deemed not a viable approach for

intrusion detection domain.

Causal knowledge reasoning using Fuzzy Cognitive Map (FCM) has been proposed [12] as an approach for anomaly intrusion detection. Packets with low causal relations to attacks are dropped and packets with high causal relations to attacks are highlighted in that experiment. By building a global matrix, FCM concept and causal relations are modeled. However, a powerful causal reasoning mechanism that supports forward and backward chaining is not available.

C. Intrusion Analysis

Intrusion analysis is another area of concern by researchers in the recent year. Intrusion analysis is to investigate reasons and methods of the attacks when the system has been attacked by the intruder. The information gained from the intrusion analysis will be very useful knowledge and reference to against the attack in the future. It can be done according to the information provided by system/network log or through the causal reasoning. Analysis of the computer intrusion is done after the causal model has been constructed. Assumption can be made by referring to the probability of the variable.

III. Bayesian Network

Bayesian networks or Belief networks are graphical models that represent the probabilistic relationships among a set of variables under uncertainty domain. Bayesian network model is represented in a directed acyclic graph and conditional probability tables (CPTs). Bayesian network has been used in various areas, such as machine learning, text mining, natural language processing, speech recognition, signal processing, bioinformatics, error-control codes, medical diagnosis, weather forecasting, and cellular networks [11].

A. Bayesian Network Fundamental

Bayesian reasoning uses Bayes' theorem, a formula to inverse conditional probabilities. Suppose X and Y are two events that may occur. Define $P(X)$ as the probability that event X occurs and define $P(Y)$ as the probability that event B occurs. Suppose further that the fact that one of the events did actually occur influences the possibility that the other event did also occur. Let $P(X, Y)$ be the probability that both events occurred and let $P(X|Y)$ be the conditional probability that event X occurred, given that event Y did actually occur. The probability that X and Y both occur is equal to the product of the probabilities that Y occurs and the conditional probability that X occurs: $P(X,Y) = P(X|Y)P(Y)$. We can interchange X and Y in the previous equation: $P(X,Y) = P(Y|X)P(X)$. Hence, combining the equations we get: $P(X|Y)P(Y) = P(Y|X)P(X)$ and thus: $P(X|Y) = P(Y|X)P(X)/P(Y)$. This is Bayes' theorem and from it, derives how a piece of evidence should modify one's believe in the occurrence of event X.

Figure 1 depicts an example of a BN consisting of five discrete variables: A, B, C, D, and E. The dependence relations are expressed in terms of conditional probability distributions for each variable in the network. Each variable has a set of possible values called its 'state space' that consists of mutually exclusive values of the variable. For example, each variable may have two possible states, '+' or '–'. If there is an arc pointing from X to Y, we say X is a parent of Y. For each variable we need to specify a CPT, one for each configuration of states of its parents. The CPTs given in a BN specify the prior joint distribution of the variables. Then the product of all CPTs gives the posterior joint distribution of the variables. Thus, the joint distribution of variables changes each time when new information is gathered about the variables. Figure 1 depicts these CPTs: $P(A)$; $P(B)$; $P(D)$; $P(C|A,B)$; and $P(E|C,D)$. Once all the CPTs have been completed, the BN can be compiled and used for analysis. This is performed by altering the states of some nodes while observing the effect of this on other nodes. The impact of changing any variable is transmitted through the network in accordance with the relationships expressed by the CPTs. Changes in any node simply arise from the combined effect of changes in all the nodes linked to it either directly or indirectly since the BN encodes a joint probability distribution over all the nodes. Every time the state of a node changes, the joint distribution is updated through the iterative application of Bayes' theorem. This refers to the process of computing the posterior marginal probability distributions of a set of variables after obtaining some observations of other variables in the model.

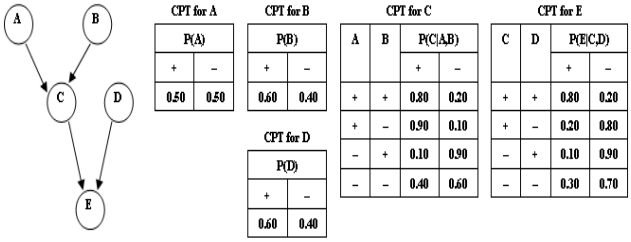

Figure 1: Illustrative BN Graphical Structure and the CPTs

A fundamental assumption of a BN is that when the conditionals for each variable are multiplied, the joint probability distribution is computed for all variables in the network. In reference to Figure 1, $P(A,B,C,D,E) = P(A) \otimes P(B) \otimes P(C|B,A) \otimes P(D) \otimes P(E|D,C)$ where $\otimes$ denotes point-wise multiplication of CPTs. However, according to the rule of total probability the equation would read as: $P(A,B,C,D,E) = P(A) \otimes P(B|A) \otimes P(C|B,A) \otimes P(D|C,B,A) \otimes P(E|D,C,B,A)$. Therefore, the assumptions made here are (comparing the two equations) $P(B) = P(B|A)$ (B is independent of A), $P(D) = P(D|C,B,A)$ (D is independent of C, B and A), and $P(E|D,C) = P(E|D,C,B,A)$ (E is conditionally independent of B and A given D and C). In theory, the posterior marginal probability of a variable, X, say $P(X)$, can be computed from the joint probability by summing out all other variables except X one by one. In practice, such an approach is not computationally feasible when a large number of variables are involved. The key to efficient inference lies in the concept of local computation where we compute the marginal of the joint without actually computing the joint distribution. Several commercial software tools are available such as Hugin [3] and Netica [4], which can automate the process of inference. Both of these tools allow the user to enter the BN structure graphically,

enter the numerical details, and then make inferences. The resulting inferences can then be shown graphically using bar charts.

B. Bayesian Network Inference Mechanism

The computation of the posterior probability distribution for a set of query nodes and given values for some evidence nodes is the basic task for probabilistic inference in Bayesian network [13]. In general, there are two major classes of inference mechanisms: exact and approximate inferences. One of the most popular algorithms in exact inference is message passing algorithm, known as Kim and Pearl's message passing algorithm [13]. The basic idea is that at an iteration of the algorithm, the belief function is updated locally using three types of parameters: the message it receives from its parent, its prior message and the conditional probability distribution. These parameters are used to update local belief in three steps: belief updating, bottom-up propagation and top-down propagation, which can be done in any order [13]. Another popular algorithm for exact inference is clustering algorithm proposed in [14]. The algorithm performs in two stages. First, the network is transformed into a polytree (junction tree), probability updating is performed subsequently. The clustering algorithm is default algorithm in GeNIe. Markov chain Monte Carlo (MCMC) algorithm is the one designed for approximate inference [13]. MCMC generates an event by making a random change to the previous event. It does this by randomly sampling a value for one of the non evidence variables, conditioned on the current values of the variables in the Markov blanket, which includes the parents, children and children's parent. This is implemented in CaMML software [15].

C. Bayesian Network Learning Algorithm

There are two stages of learning in Bayesian network, which are structure learning and parameter learning. In Bayesian network, the direct acyclic graph is called the structure and the values in the conditional probability distribution are called the parameters [16]. Learning the structure is considered a harder problem than learning the parameters. The parameter learning is to learn the strength of these dependencies, as encoded by the entries in the CPTs. Bayesian network structure learning algorithms are generally fallen into two groups, search-and-scoring based algorithms and dependency analysis based algorithms [10]. Dependency analysis approach takes the view that Bayesian networks depict conditional independence relations among the variables. Hence, the approach tries to construct a Bayesian networks using dependency information obtained from the data. In search-and-scoring approach, Bayesian networks encode joint probability distributions and a measure for assessing the goodness of the encoding can be derived [17]. A measure is used (Bayesian, Minimum Description Length (MDL) or Kull-back-Leibler (KL) entropy scoring function) as a criteria for finding out the best Bayesian structure, which maximizes the used measure and best fits the data. The comparison of the two approaches is in [18].

IV. Methodology

The methodology consists of several steps. The first step is to derive a causal knowledge model that captures the causal relationships between the domain variables from a public domain network intrusion related data. Bayesian network is used to represent the causal model capitalizing on its strength in uncertainty handling, efficient evidence propagation, good track records, and availability of powerful learners [2]. Bayesian network is again used in the next step, capitalizing on its strength in diagnostic, prognostic, and hybrid inferences. After the model has been constructed, causal reasoning related to intrusion detection and analysis is carried out. By observing some anomalous event, the source of intrusion can be detected and the impacts of such intrusion can be predicted. The correctness of the learned model from Bayesian network needs to be assured before using it for causal reasoning. The verification is done by capitalizing the feature selection capability of Bayesian learning. A particular variable is marked as a target/class variable, and it has been shown that the set of selected features/variables using Bayesian learning can be used to predict the value of the class variable almost equally well as compared to the prediction done by using the complete set of domain variables. The methodology is composed of four steps: data pre-processing, causal discovery, verification of Bayesian causal model learned, and causal reasoning for intrusion detection and analysis. The steps are elaborated in the following subsections.

A. Data Pre-processing

Public domain dataset named KDD Cup 1999 dataset [19] is used in this experiment. The dataset is based on 1998 DARPA Lincoln Lab network connection. KDD'99 intrusion detection dataset is a very famous dataset in the intrusion detection domain, and it has been used widely for the evaluation of various intrusion detection techniques. It is a huge dataset, which consists of approximately 4,900,000 single connections and 41 features per connection. This is too large for the experiment as it cannot be supported by the free software. Therefore, a set of randomly selected 27,933 records having 41 features from "10% KDD 1999" data subset has been used in the experiment. All the network connections are categorized into either normal or 24 other types of attack, which fall into four main categories as follows [20]**:**

- Denial of Service Attack (DoS): Attacker makes the system too busy to handle the legitimate request or legitimate user to use the machine/service.
- User to Root Attack (U2R): Attacker tries to get the access rights from a normal user account.
- Remote to Local Attack (R2L): Attacker tries to exploit the system vulnerabilities in order to control the remote machine through network as local user.
- Probing Attack: Attacker tries to gather useful information about the target host in order to look for exploit.

The 41 features in the dataset consist of all forms including continuous, discrete and symbolic data. In the data pre-processing stage, discretization of the dataset is needed. This is because most of the algorithms or software tools do

not accept mixed or continuous dataset. The label of the 41 features and their network data features are shown in Table 1.

TABLE 1: NETWORK DATA FEATURE LABELS

Label	Features	Type
X1	duration	Continuous
X2	protocol-type	Discrete
X3	service	Discrete
X4	flag	Discrete
X5	src_bytes	Continuous
X6	dst_bytes	Continuous
X7	land	Discrete
X8	wrong_fragment	Continuous
X9	urgent	Continuous
X10	hot	Continuous
X11	num_failed_logins	Continuous
X12	logged_in	Discrete
X13	num_compromised	Continuous
X14	root_shell	Discrete
X15	su_attempted	Discrete
X16	num_root	Continuous
X17	num_file_creations	Continuous
X18	num_shells	Continuous
X19	num_access_files	Continuous
X20	num_outbound_cmds	Continuous
X21	is_host_login	Discrete
X22	is_guess_login	Discrete
X23	count	Continuous
X24	srv_count	Continuous
X25	serror_rate	Continuous
X26	srv_serror_rate	Continuous
X27	rerror_rate	Continuous
X28	srv_rerror_rate	Continuous
X29	same_srv_rate	Continuous
X30	diff_srv_rate	Continuous
X31	srv_diff_host_rate	Continuous
X32	dst_host_count	Continuous
X33	dst_host_srv_count	Continuous
X34	dst_host_same_srv_rate	Continuous
X35	dst_host_diff_srv_rate	Continuous
X36	dst_host_same_src_port_rate	Continuous
X37	dst_host_srv_diff_host_rate	Continuous
X38	dst_host_serror_rate	Continuous
X39	dst_host_srv_serror_rate	Continuous
X40	dst_host_rerror_rate	Continuous
X41	dst_host_srv_rerror_rate	Continuous

B. Causal Discovery

Causal discovery aims to learn the structure and parameter from the data provided. The relationships among variables (data elements) will be discovered using appropriate Bayesian learning tools to construct a causal model at this stage. The process consists of two parts, which are structure learning and parameter learning.

As we have mentioned in the earlier stage, data-mining approach is used in the methodology. So, the structure of the Bayesian model is learnt automatically from the provided data without any human intervention. GeNie [21] is used for this purpose. For clarity purposes, the problem size has been reduced by focusing on the variables that are directly related to the class variable (i.e., X42). The algorithm used in Genie to build the model in our experiment is Greedy Thick Thinning. The relationship between variables can be identified after the structure has been constructed. Figure 2 illustrates the Bayesian network model after structure learning but before parameter learning is complete.

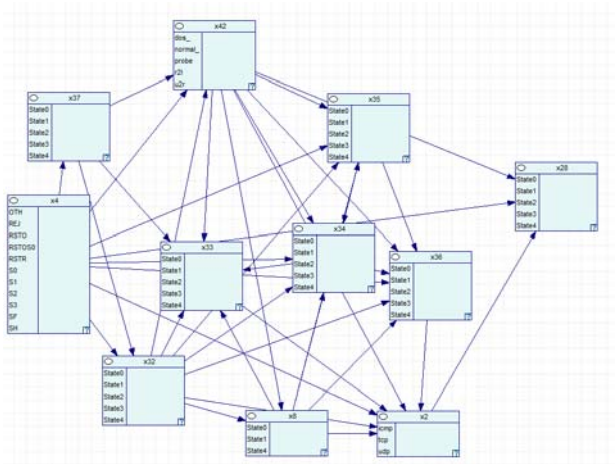

Figure 2: Bayesian Network Model after Structure Learning

After the relationships among the variables have been recognized, the next step is to complete the causal model by doing the parameter learning. Parameter learning in Bayesian network is to discover the probabilistic relationships between domain variables, which are captured in conditional probability table (CPT). Learning parameter is generally more straight-forward than learning the structure. Several algorithms can be used in parameter learning. One of the famous algorithms is Expectation-Maximization (EM) algorithm. An example of a CPT for a specific variable in the dataset called *dst_host_srv_diff_host_rate* is shown below:

TABLE 2 (A): CPTS FOR THE DST_HOST_SRV_DIFF_HOST_RATE1

x4	OTH	REJ	RSTO	RSTOS0	RSTR
State0	0.66666667	0.89135159	0.8974359	0.75	0.9954955
State1	0.0833333...	0.0756192...	0.0128205...	0.0625	0.0011261...
State2	0.0833333...	0.0043459...	0.0128205...	0.0625	0.0011261...
State3	0.0833333...	0.0013037...	0.0256410...	0.0625	0.0011261...
State4	0.0833333...	0.0273794	0.0512820...	0.0625	0.0011261...

TABLE 2 (B): CPTS FOR THE DST_HOST_SRV_DIFF_HOST_RATE2

S0	S1	S2	S3	SF	SH
0.97043011	0.72222222	0.76470588	0.5	0.93321846	0.96428571
0.0120967...	0.0555555...	0.0588235...	0.125	0.0147936...	0.0089285...
0.0026881...	0.11111111	0.0588235...	0.125	0.0402202...	0.0089285...
0.0013440...	0.0555555...	0.0588235...	0.125	0.0007144...	0.0089285...
0.01344086	0.0555555...	0.0588235...	0.125	0.0110532...	0.0089285...

C. Verification of Bayesian Causal Model Learned

A by-product of Bayesian network learning is that we can get a set of features that are on the Markov blanket of the class node. The Markov blanket of a node *N* is the union of *N*'s parents, *N*'s children, and the parents of *N*'s children. This subset of nodes can shield *N* from being affected by any node outside the blanket. When using a Bayesian network classifier on complete data, the Markov blanket of the class node forms a natural feature selection, as all features outside the Markov blanket can be safely detected from the Bayesian network. This can often produce a much smaller Bayesian network without compromising the classification accuracy. The verification is done by capitalizing the feature selection capability of Bayesian

learning. The feature selection algorithm – CEFS in Tetrad IV [22] and BN PowerConstructor [23] have been used to build the reduced structure. The relationship between the variables can be identified after the structure has been constructed and the feature variables that have relationship with the class variable are adopted for running the accuracy test using different classification algorithms. The experimental results are listed in the following table.

TABLE 3: PERCENTAGE OF CORRECTLY CLASSIFIED INSTANCES

Features Algorithm	41	25	30	7	10
J48	94.24	94.16	93.96	91.29	92.53
DecisionTable	92.98	93.12	92.4	90.96	91.85
VFI	76.14	73.37	70.32	64.52	65.22
JRIP	93.57	93.36	92.83	90.61	91.32
SimpleCart	94.28	94.19	93.96	91.34	92.61
MultilayerPerceptron	93.09	92.88	94.00	90.21	92.37
Classification ViaClustering	57.17	55.12	58.42	49.05	52.50
RBFNetwork	80.75	87.73	86.51	86.61	86.91

In the experiment, other features that do not have the direct relationship with the class node have been removed from the dataset. First of all, the original dataset that contains 41 features and class has been tested using different algorithms and the accuracy has been recorded. After that, 25 features that have the direct relationship with the class node have been used to do the same test. The 7 and 10 features shown in table below have been selected using BN PowerConstructor. The percentage of correctly classified instances for different algorithms and the different number of features are shown in Table 3 above. According to the table, the percentage of the correctly classified instances does not change much even though the number of features has been reduced. This has proven that Bayesian network learning has successfully figure out the correct relationship among the nodes.

D. Causal Reasoning for Intrusion Detection and Analysis

Bayesian network model is completed when the structure and parameter of the network have been learned. Causal reasoning with the ability to diagnose the root cause(s) and predict the outcome(s) can only take place after the model is successfully constructed. GeNie that supports the prognostic, diagnostic and hybrid reasoning is adopted in the experiment. However, GeNie consists of several limitations. It either can support large amount of connections with few features or little connections with many features. Unfortunately, intrusion detection domain is a data-rich domain, which comprises huge amount of connections with various features. Hence, we reduce the number of features in the KDD cup dataset to 10 features, which have been selected using feature selection from BN PowerConstructor.

1) Prognostic reasoning

Prognostic reasoning is the ability to predict the future outcome(s). Prognostic reasoning can be done due to the capability of the evidence propagation mechanism in the Bayesian network. The probability of the connected nodes is affected by altering the value of the certain node(s) in a network. The posterior probability for the states in each of the remaining nodes is automatically updated by the evidence propagation mechanism.

According to Figure 3, when there is an evidence of REJ in the flag node(x4), the percentage of the state4 in the connections that have "REJ" errors node(x28) is raised to 82%. The prediction is logical as when there is an evidence of the reject flag, the connections that have the reject error will increase accordingly. Moreover, the percentage of the probe attack type is raised to 66% by setting the evidence of REJ in flag node. From this point of view, we believe the prediction is correct as flag with REJ and probe attack are correlated.

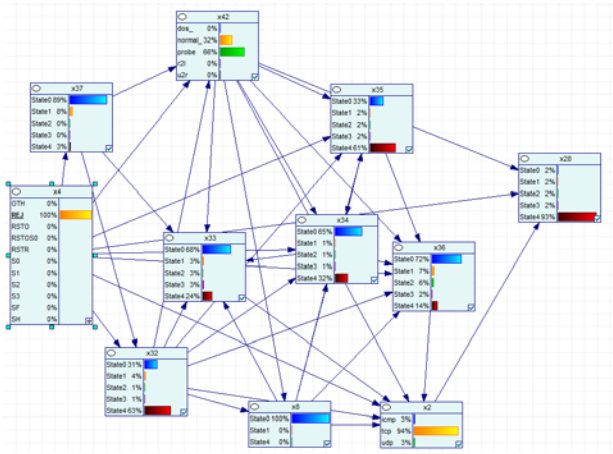

Figure 3: Predictive Bayesian Network Model

2) Diagnostic reasoning

The ability to diagnose the root cause(s) of certain event(s) is called diagnostic reasoning. When something happens, people would like to investigate the cause(s) that influence the target variable (i.e., variable of interest). One of the things we do is to use Bayesian network to diagnose the possible root cause(s) by changing the probability of a variable at hand. Figure 4 shows when there is an evidence of the state4 in the connections that have "REJ" errors node(x28), the probability of the three main causes, which are class node(x42), flag node(x4) and protocol node(x2) will change. The probability of probe in class node has increased to 59%, TCP in protocol node has increased up to 38% and REJ in flag node has increased up to 63%. It means that the high percentage of the connection that has reject error is mainly caused by the type of attack, the flag and the protocol of the connection.

From the result, we can make assumption that most of the probing attack will get the reject error connection and mostly happened in transfer connection protocol (TCP). This is reasonable because probe mostly happened in TCP port, which are connection-oriented and therefore give good feedback to the attacker. Furthermore, probing attack will get a lot of reject connections as probing attack will send packet to all ports to check which port is open and then looking for an exploit. So, most of the connections will get reject error. The assumption that we have made is logical and it shows that Bayesian network has correctly doing the diagnostic reasoning.

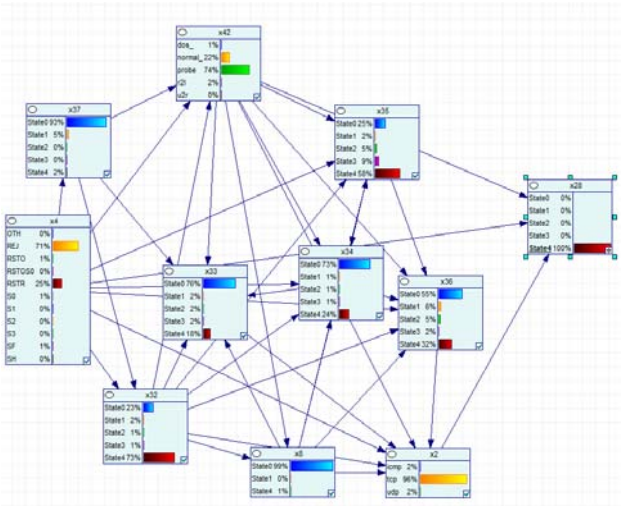

Figure 4: Diagnostic Bayesian Network Model

3) Hybrid reasoning

Hybrid reasoning combines the prognostic and diagnostic inferences. The purpose of hybrid reasoning is to allow us to do the diagnosis and prognosis mentioned earlier simultaneously. Values are preset for both target node and cause node in order to observe how it affects the posterior probability of other nodes. As shown in Figure 5, when there is an evidence of REJ and stage4 for node x4 and x28, the probability of the remaining nodes will change.

The amalgamation of both diagnosis and prognosis increase the probability of stage0 in node x33, x34 and x36 and raise the probability in TCP of protocol node (X2) to 99%. Based on the explanation in the earlier statement, the change of the probability in TCP is acceptable and reasonable.

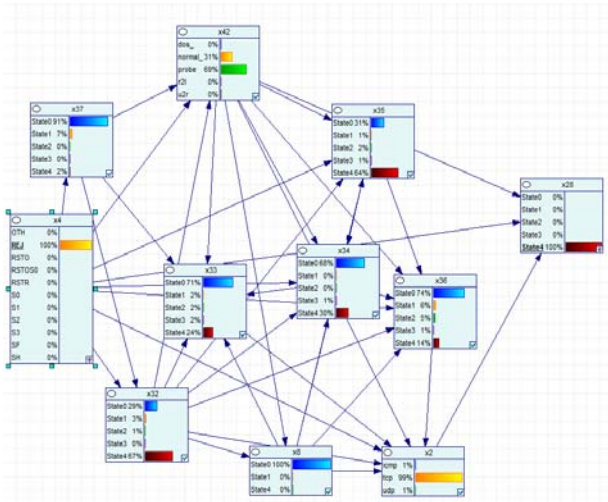

Figure 5: Combination of Prognostic and Diagnostic Bayesian Network Model

V. Conclusion and Future Work

Bayesian network is a popular causal model in last decades. It is a well established method for probabilistic causal reasoning. However, there are only few researchers who have applied Bayesian network for causal reasoning in the intrusion detection domain. A methodology to solve the two major problems found in network intrusion detection has been proposed in this paper. A causal knowledge driven approach using Bayesian network is adopted for the modeling and reasoning about the intrusion domain. Although some research work has been done in intrusion detection using Bayesian network over the years, the full capabilities of Bayesian network have not been fully utilized in this domain. At this stage, an experiment has been carried out to test the accuracy of the Bayesian network learning algorithms. A public domain dataset has been used in the experiment for benchmarking. As the results shown, the capability of Bayesian learning is reasonably accurate and efficient. Locally generated simulation data will be used in similar experiments, and more details on causal reasoning will be explored and analyzed in future work.

References

[1] Wu, J. & Z. Hu. 2008. Study of Intrusion Detection Systems (IDSs) in Network Security. In Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on, 1-4.

[2] W. P, Cheah. 2009. A Methodology for Constructiong Causal Knowledge Model from Fuzzy Cognitive Map to Bayesain Belif Networks. *In PhD Thesis*, Department of Computer Science, Chonnam National University, South Korea.

[3] http://www.hugin.com/

[4] http://www.norsys.com/

[5] Wu, S. X. & W. Banzhaf (2010) The use of computational intelligence in intrusion detection systems: A review. *Applied Soft Computing,* 10**,** 1-35.

[6] Endoft, C., E. Schultz & J. Mellander. 2004. *Intrusion Detection & Prevention*. McGraw-Hill Osborneo Media.

[7] Nadkarni, S. & P. P. Shenoy (2004) A causal mapping approach to constructing Bayesian networks. *Decision Support Systems,* 38**,** 259-281.

[8] Christopher, K., M. Darren, R. William & V. Fredrik. 2003. Bayesian Event Classification for Intrusion Detection.

[9] Patcha, A. & J.-M. Park (2007) An Overview ofAnomaly Detection Techniques: Existing Solutions and Latest Technological Trends. *Computer Networks,* 51**,** 3448-3470.

[10] Chebrolu, S., A. Abraham & J. P. Thomas (2005) Feature Deduction and Ensemble Design of Intrusion Detection Systems. *Computers & Security,* 24**,** 295-307.

[11] García-Teodoro, P., J. Díaz-Verdejo, G. Maciá-Fernández & E. Vázquez (2009) Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges. *Computers & Security,* 28**,** 18-28.

[12] Jazzar, M. & A. Jantan. 2008. An Approach for Anomaly Intrusion Detection Based on Causal Knowledge-Driven Diagnosis and Direction. In *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing,* ed. R. Lee, 39-48. Springer Berlin / Heidelberg.

[13] Korb, K. & A. Nicholson. 2004. Bayesian Artificial Intellegence. CRC Press.

[14] Lauritzen, S. & D. Spiegelhalter (1988) Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society, Series B,* 50**,** 157-224.

[15] Russell, S. J. & P. Norvig. 2003. Artificial Intelligence : A Modern Approach. Upper Saddle River, EUA : Prentice-Hall.

[16] Neapolitan, R. 2003. *Learning Bayesian Networks*. Prentice Hall.

[17] Man Leung, W. & L. Kwong Sak (2004) An Efficient Data Mining Method for Learning Bayesian Networks using an Evolutionary Algorithm-based Hybrid Approach. *Evolutionary Computation, IEEE Transactions on,* 8**,** 378-404.

[18] Cheng, J. (2002) Learning Bayesian Networks from Data: An Information-theory Based Approach. *Artificial Intelligence,* 137**,** 43-90.

[19] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[20] Ma, Y., D. Choi, S. Ata & H. Nguyen. 2008. Application of Data Mining to Network Intrusion Detection: Classifier Selection Model. In *Challenges for Next Generation Network Operations and Service Management*, 399-408. Springer Berlin / Heidelberg

[21] http://genie.sis.pitt.edu/about.html#genie
[22] http://www.phil.cmu.edu/projects/tetrad/tetrad4.html
[23] http://webdocs.cs.ualberta.ca/~jcheng/bnpc.html

Yit Yin Wee received her BSc in Computer Science from University Putra, Kuala Lumpur, Malaysia in 2009. She is currently a Masters student in Multimedia University, Melaka, Malaysia. Her research interests include artificial intelligence, data mining, networking, and computer security.

Wooi Ping Cheah received his BSc from Campbell University, US, in 1986 and his MSc in Software Engineering from the University of Science Malaysia, in 1993. He received his PhD in Computer Science from Chonnam National University, South Korea in 2009. He is currently a Lecturer at the Faculty of Information Science and Technology, Multimedia University, Malaysia. His research interests include artificial intelligence, software and knowledge engineering, decision support systems and data mining.

Shing Chiang Tan received the B. Tech. and M. Sc. (Eng.) degrees from University of Science Malaysia, and the PhD degree from Multimedia University in 1999, 2002, 2008 respectively. Currently, he is a senior lecturer with the Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia. His research interests include computational intelligence techniques (artificial neural networks, evolutionary algorithms, decision trees, etc) and their applications to pattern classification, condition monitoring, fault diagnosis and medical diagnosis.

Kuokkwee Wee received his BSc in Computer Science and MSc in Networking from University Putra, Kuala Lumpur, Malaysia in 2003 and 2005. He is currently a PhD student and working as a Lecturer at the Faculty of Information Science and Technology in Multimedia University, Melaka, Malaysia. His research interests include quality of service, broadband wireless access, networking and mobile communication.