Reasoning with Cause and Effect in Intrusion Detection

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

***Abstract*—Intrusion detection is an essential tool to protect hacking and unauthorized access in computer networks nowadays. Mechanisms used to attack keep evolving as the internet technology is improving. Hence, the task of differentiating authorized and unauthorized access has become more and more challenging. The modeling of network intrusion domain and causal reasoning for the intrusion detection has been proposed in this paper to address the security issues of a network. Bayesian network modeling with causal knowledge-driven approach has been selected for a network intrusion domain. Reasoning capabilities of Bayesian network have been adapted to perform detection and analysis in the domain.There are two main problems to be addressed in this paper: the first problem is to model the network intrusion domain and the second problem is to perform causal reasoning for intrusion detection and analysis. A methodology has been proposed to solve the two problems mentioned above. Intrusion detection is viewed as fault diagnosis in causal reasoning, and the analysis of the effect is viewed as fault prognosis. To address the first problem under causal knowledge-driven approach, we propose Bayesian network for the modeling of network intrusion domain. The second problem is addressed by applying the powerful reasoning capabilities of Bayesian network. The capabilities of causal reasoning using Bayesian network have not been fully discovered in the domain of intrusion detection. This research work is to bridge the gap.**

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

I. Introduction

Intrusion detection has been a major concern since last decade. Today, most research and discussions are focused on the tools and techniques used for protecting the computer networks. Thousands of papers could be obtained from the web about the improvement of the frameworks and techniques for intrusion detection system. However, there are only few researchers apply the concept of causal reasoning in this domain. The main problem here is to construct a model of network intrusion and to use it for subsequent detection and analysis. Although, some of the modern methods could help in some way, no methodology has ever claimed to provide a general purpose solution for intrusion detection and analysis. Knowledge engineering and data mining are the two approaches of constructing the domain model. Since intrusion detection is a data rich domain, knowledge engineering approach will be problematic in constructing the model. In the knowledge engineering approach, domain experts collaborate with a knowledge engineer to manually identify the relationships between domain variables. Very often, it is completely impractical as the complexity of the problem grows exponentially with the number of variables. On the other hand, data mining approach is more suitable in constructing the domain model. 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.

Causal reasoning, which is about diagnosing the root cause and predicting the effect of the intrusion, is done after constructing the model. 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 modelling and reasoning problems. The capabilities of Bayesian network are not fully discovered even though it is being used by many researchers in intrusion detection domain for classification and feature deduction. Supported by powerful learning algorithm, Bayesian network serves as a good modelling tool for a data rich domain like intrusion detection [1]. Besides, it also provides an efficient evidence propagation mechanism and powerful reasoning capability. It is a more mature framework as many Bayesian software tools have been commercialized into today's market, such as Hugin[2] and Netica[3].

Manuscript received September 12, 2012; revised October 23, 2012. This research is supported by the GRA Grant funded by Multimedia University, Malaysia (Project ID IP20110105018).

The authors are with the Multimedia University, Jalan Ayer Keroh Lama 75450 Bukit Beruang Melaka Malaysia. (e-mail: yywee@mmu.edu.my, wpcheah@mmu.edu.my, sctan@mmu.edu.my, wee.kuok.kwee@mmu.edu.my).

II. Intrusion Detection

Intrusion detection is the art of discovering and identifying any intrusive activities towards the networks system. In 1980, James P. Anderson introduced the ideas of intrusion detection and outlining ways to improve computer security auditing and surveillance [4]. Hence, intrusion detection system (IDS) has become very important to strengthen the security, confidentiality and integrity of critical information systems. Detection and analysis on the impacts of the malicious activities is the main objective for intrusion detection and analysis system [5].

A. Intrusion

Intrusion Detection System (IDS) is a security tool that is used to detect and analyse the incoming traffic activities and to raise the alarm if the activities are identified as anomalous. Basically, there are two types of intrusion detection systems: Host-based Intrusion Detection System (HIDS) and Network-based Intrusion Detection System (NIDS).

1) Network intrusion detection system (NIDS)

NIDS monitors and analyses external event such as traffic

volume, IP address, service port and content of individual packets flowing through a network to look for possible attacks. NIDS consists of the sensor and the console; these two components are placed in front of firewalls or at key network choke point. Normally the console acts as the central management which raises the alert when there is a possible intrusion happened. The sensor which is located on a segment of the network, normally in the demilitarized (DMZ) or at network borders, is used to capture all network traffic and analyze the contents. Below are the advantages and disadvantages of NIDS.

Advantages:

1) It can monitor the whole network from one location.
2) It can identify network traffic patterns and troubleshoot network problems.
3) It can protect whole network from intrusion.

Disadvantages:

(a) It generates false alarm easily.
(b) Unable to detect certain attacks (false negative).
(c) Unable to understand host specific processes or protect from unauthorized physical access.
(d) Unable to detect network in different network segments.

Example: Snort, Bro

2) Host-based intrusion detection system (HIDS)

HIDS analyses internal event such as process identifier, system calls, application logs, file-system modifications (binaries, password files, capability databases, access control lists) and other host activities and state. Every single network traffics needs to be analyzed before passing the host. Therefore, HIDS is useful for monitoring potentially dangerous user activities within the network. HIDS is installed locally on host machine such as servers, workstations, notebooks and computers. Therefore HIDS is a very versatile system compared to NIDS. Here are some of the advantages and disadvantages for HIDS.

Advantages:

1) Low false positive rate.
2) It has nearly real-time detection and response.
3) It suits well in encrypted environment.
4) It works well in switched network.
5) It monitors all system activities.

Disadvantages:

1) HIDS stops working once the host machine is compromised.
2) OS dependent.

Example: OSSEC

B. Intrusion Detection Model

Three popular detection techniques used for intrusion detection are anomaly-based intrusion detection, signature-based intrusion detection and hybrid intrusion detection.

Anomaly-based intrusion detection creates a base-line profile of the normal system and capable to distinguish the incoming system activity either normal or anomalous. An anomaly alarm will be generated if the system activities are found to be anomalous.

Signature-based intrusion detection relies on the predefined set of attack signatures, it monitors packets on the network and compares them against a database of signatures or attributes from known malicious threats. Unfortunately, it can only detect previously known attacks. Therefore, they must be constantly updated with the signature of new attacks [6].

Hybrid intrusion detection is the combination of anomaly-based and signature-based approaches. The anomaly technique aids in the detection of new and unknown attacks while the signature technique detects known attacks in this hybrid approach. Theoretically, although better intrusion detection is expected by combining both techniques, the resulting hybrid systems are not always better [7].

C. Anomaly Intrusion Detection

In this paper, we will focus on anomaly intrusion detection. Feature selection, categorization and causal reasoning are the three main techniques to be concerned.

Feature selection enables us to discover the important features and eliminate the insignificant or little contribution input. Since the data is very huge even for a small network, feature selection playing a significant role in real time intrusion detection. The processing time for detection and storage space will be reduced significantly. Besides, more efficient and effective results will be produced [8].

Categorization can be divided into classification and clustering. The main difference between classification and clustering is classification is a process to categorize the type of attack using supervised data while clustering is using unsupervised data. In the categorization process can be done in three ways: statistical, knowledge-based and machine learning approaches [9]. Categorization plays a significant role in intrusion detection but it is too restrictive to a target variable and it does not provide a comprehensive reasoning capability.

Causal reasoning is a process to identify any cause(s) leading to certain effect(s) and the causal relationships among various events. The structure of the model must be constructed before causal reasoning can take place. Modelling the structure can be done either by knowledge engineering or data-mining approach. Causal reasoning is a comprehensive process that includes the feature selection, classification and also diagnosis and prognosis. However, there is limited research work, which has applied causal reasoning in intrusion detection.

In [6], a systematic procedure for constructing Bayesian network from domain knowledge of experts using a causal mapping approach has been proposed. However, it is not in the intrusion detection domain. Furthermore, knowledge engineering approach is not the best for intrusion detection domain due to huge datasets. An approach based on causal knowledge reasoning for anomaly intrusion detection using Fuzzy Cognitive Map (FCM) has been proposed [10]. Packets with low causal relations to attacks are dropped and packets with high causal relations to attacks are highlighted in that experiment. The FCM concepts and causal relations are modelled by building a global matrix. However, a powerful causal reasoning mechanism that supports forward and backward chaining is not available.

III. Bayesian Network

Bayesian 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 (DAG) 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 [9].

A. 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 [11]. In general, there are two major classes of inference mechanisms: exact and approximate inferences.

1) Exact inference

One of the most popular algorithms in exact inference is message passing algorithm, known as Kim and Pearl's message passing algorithm [11]. 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 [11]. Another popular algorithm for exact inference is clustering algorithm proposed by Lauritzen [12]. The algorithm performs in two stages. First, the network is transformed into a polytree (junction tree), and then probability updating is performed on that polytree. The clustering algorithm is default algoritmn in GeNIe.

2) Approximate inference

Markov chain Monte Carlo (MCMC) algorithm is the one designed for approximate inference [11]. 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 [13].

B. 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 [14]. 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 [8]. 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 [15]. 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 [16].

IV. Methodology

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 [1]. Once the causal model has been constructed, reasoning with causes and effects is carried out to detect and identify the impact of the network intrusion. The source of intrusion can be detected by an observation of some abnormal events. After detecting the source, the impacts of such intrusion can be predicted. Before using the derived Bayesian model for causal reasoning, we need to be assured of the correctness of the learned model. 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, causal reasoning, and the verification of Bayesian causal model. The steps are elaborated below.

A. Pre-Processing of the Data

Knowledge Discovery and Data mining (KDD) cup 1999 [17] is the most used network intrusion dataset by the researchers to experiment on the computer network intrusion detection techniques. KDD'99 intrusion detection dataset is the public domain data which can be obtained from the web. It consists of approximately 4,900,000 single connections and 41 features per connection. The label of the 41 features and their network data features are shown in the table below.

TABLE I: Feature Labels for Networks Data in KDD Cup 1999

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

In the experiment, a set of randomly selected 27933 records having 41 features ("10% KDD 1999" data subset) has been used. All the network connections are categorized into either normal or 24 other types of attack, which fall into four main categories as follows [18]**:**

- 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.

B. Verification of Bayesian Causal Model

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 [20] and BN PowerConstructor [21] 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. 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 the table below. 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.

C. Causal Discovery

At this stage, the purpose is to discover the relationship between variables (data elements) using appropriate Bayesian learning tools and to construct a causal model. There are two parts of this process: structure learning and parameter learning.

The structure of the Bayesian model is learnt automatically without human intervention. GeNie has been adopted 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 Greedy Thick Thinning algorithm in GeNie[19] has been used to build the model. The relationship between variables can be identified after the structure has been constructed.

Parameter learning is done after the structure has been constructed. Parameter learning in Bayesian network is to discover the probabilistic relationships between domain variables, which are captured in conditional probability table (CPT) of each node. Learning parameter is generally more straight-forward than learning the structure. The parameter can be learnt using several algorithms. One of them is Expectation-Maximization (EM) algorithm. An example of a CPT for a specific variable called *dst_host_srv_diff_host_rate* shown below:

TABLE II: CPTS FOR THE DST_HOST_SRV_DIFF_HOST_RATE VARIABLE

x4	OTH	REJ	RSTO	RSTOS0	RSTR	S0	S1	S2	S3	SF	SH
State0	0.66666667	0.89135159	0.8974359	0.75	0.9964955	0.97043011	0.72222222	0.76470588	0.5	0.93321046	0.96428571
State1	0.0833333...	0.0756192...	0.0128205...	0.0625	0.0011261...	0.0120967...	0.0555555...	0.0588235...	0.125	0.0147936...	0.0089285...
State2	0.0833333...	0.0043459...	0.0128205...	0.0625	0.0011261...	0.0026881...	0.11111111	0.0588235...	0.125	0.0402202...	0.0089285...
State3	0.0833333...	0.0013037...	0.0256410...	0.0625	0.0011261...	0.0013440...	0.0555555...	0.0588235...	0.125	0.0007144...	0.0089285...
State4	0.0833333...	0.0273794	0.0512820...	0.0625	0.0011261...	0.01344086	0.0555555...	0.0588235...	0.125	0.0110532...	0.0089285...

TABLE III: PERCENTAGE OF CORRECTLY CLASSIFIED INSTANCES

Features / Algorithm	42	25	30	7	10
J48	94.2398	94.161	93.9605	91.2899	92.525
DecisionTable	92.9832	93.1157	92.4033	90.9605	91.8519
VFI	76.1429	73.372	70.3183	64.5151	65.2239
JRIP	93.5703	93.3627	92.8257	90.6133	91.3194
SimpleCart	94.2792	94.1923	93.957	91.3364	92.6073
MultilayerPerceptron	93.0942	92.8794	93.9963	90.2123	92.371
ClassificationViaClustering	57.1654	55.1212	58.4184	49.0459	52.5042
RBFNetwork	80.7468	87.7314	86.5106	86.6108	86.9187

D. Causal Reasoning

Causal reasoning is the ability to diagnose the root cause(s) and predict the outcome(s). There are many Bayesian networks tools and software in the market offer the causal discovery and causal reasoning functions. Free software called GeNie has been used in the experiment as it supports the prognostic, diagnostic and hybrid reasoning.

1) Reasoning with effects in the network intrusion

Reasoning with effects in the network intrusion is the ability to do the future outcome(s) prognosis and predict the effects for an intrusion. In Bayesian network, the probability of other nodes is affected by changing the value of certain node(s) in the network. Whenever the value of certain node is changed, the evidence propagation mechanism in the network will automatically update the posterior probability for the states in each of the remaining nodes. Fig. 1 shows 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) increased to 82%. This node is the most affected node and it is reasonable as REJ flag will raise the percentage of the connection that consist "REJ" errors.

2) Reasoning with causes in network intrusion

Reasoning with causes in network intrusion is able to diagnose the root cause(s) for an intrusion. The factor(s) that

will influence a target variable (i.e., variable of interest) and the root cause are the things that concern most people. In this case, Bayesian network is used to diagnose the possible root cause(s) by changing the probability of a variable of interest. Fig. 2 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 to 38% and REJ in flag node has increased to 63%. The increase of percentage in the three nodes is logical. Probe attack happened when intruder tries to attack the network and to gain information from the connection, and it is mostly happened in the transmission control protocol (TCP). Therefore the changes of the probability in those nodes are reasonable.

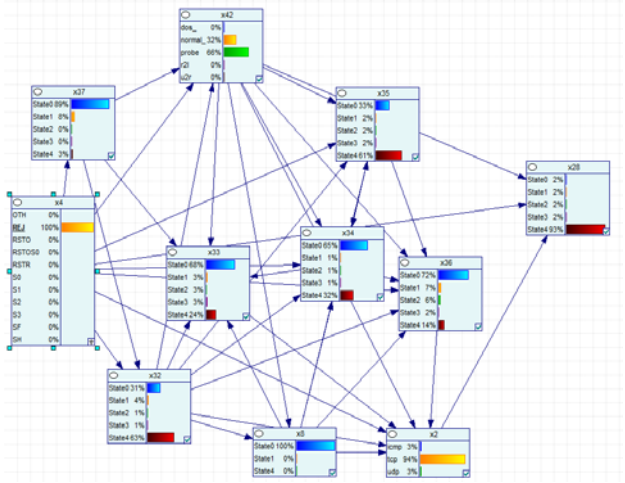

Fig. 1. Proposed bayesian network model to predict the effect

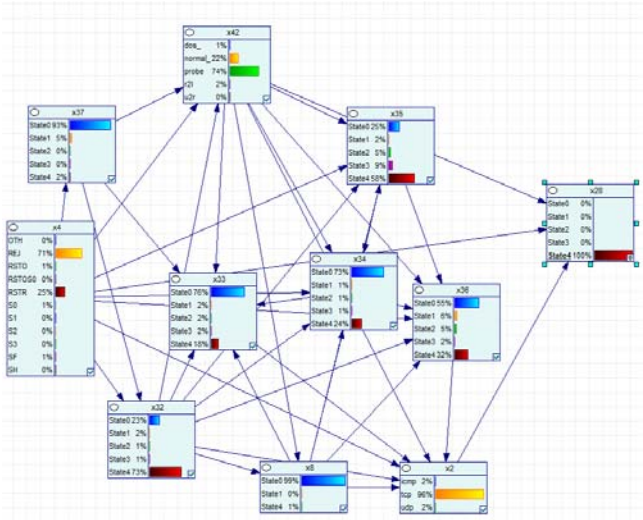

Fig. 2. Proposed bayesian network model to diagnose the root cause

3) *Hybrid reasoning with causes and effects*

Hybrid reasoning with causes and effects means the combination of predicting the effects and diagnosing the causes in the network intrusion. The diagnosis and prognosis as mentioned above will happen at the same time. The observation on how the posterior probabilities of other nodes are affected is done by setting the values for both target node and cause node. The changes of the posterior probabilities for the affected nodes are different if we set the value for both target node and cause node separately. As shown in Fig. 3, 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 will rise up the probability in TCP of protocol node (X2) to 99%.

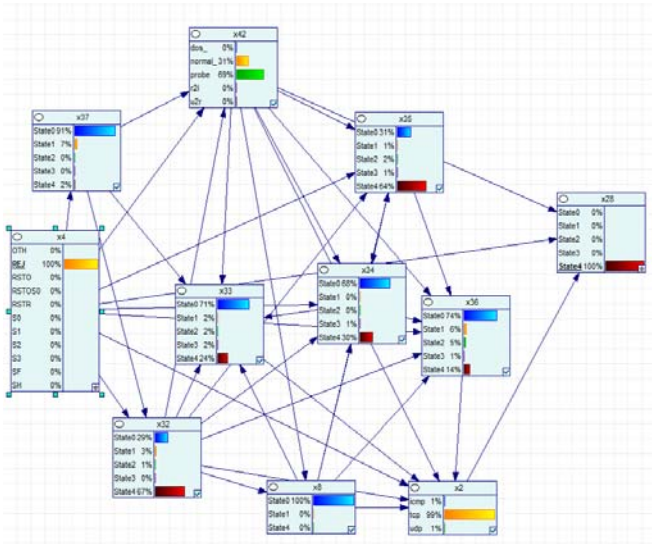

Fig. 3. Hybrid of prognostic and diagnostic bayesian network model

V. Conclusion and Future Work

Causal knowledge driven approach using Bayesian network is the methodology proposed to solve the two major problems in network intrusion. In this paper, the methodology mentioned above is adopted for the modeling and reasoning with causes and effects about the intrusion domain. Bayesian network is a mature framework and some research work has been done in intrusion detection using Bayesian network all over the years. However, the full capabilities of Bayesian network have not been fully utilized in this domain. The preliminary experiments have been carried out to test the accuracy of the Bayesian network learning algorithms. At this current stage, public domain datasets are used in the experiment for benchmarking. As the results shown, the capability of Bayesian learning is reasonably accurate and efficient. In the future work, locally generated network intrusion simulation data will be used in similar experiments, and more details on causal reasoning will be explored and analyzed in future work.

References

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

[2] Hugin Expert. [Online]. Available: http://www.hugin.com/

[3] Norsys Software Corp. [Online]. Available: http://www.norsys.com/

[4] J. P. Anderson, Computer Security Threat Monitoring and Surveillance, 1980.

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

[6] S. Nadkarni and P. P. Shenoy, "A causal mapping approach to constructing Bayesian networks," *Decision Support Systems*, vol. 38**,** pp. 259-281, 2004.

[7] A. Patcha and J. M. Park, "An Overview ofAnomaly Detection Techniques: Existing Solutions and Latest Technological Trends," *Computer Networks*, vol. 51, pp. 3448-3470, 2007.

[8] S. Chebrolu, A. Abraham, and J. P. Thomas, "Feature Deduction and Ensemble Design of Intrusion Detection Systems," *Computers and Security*, vol. 24, 295-307, 2005.
[9] P. G. Teodoro, J. D. Verdejo, G. M. Fernández, and E. Vázquez, "Anomaly-based Network Intrusion Detection: Techniques, Systems and Challengesm," *Computers and Security*, vol. 28, 18-28, 2009.
[10] M. Jazzar and A. Jantan, "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, pp. 39-48. 2008.
[11] K. Korb and A. Nicholson, "Bayesian Artificial Intellegence," *CRC Press*, 2004.
[12] S. Lauritzen and D. Spiegelhalter, "Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems," *Journal of the Royal Statistical Society, Series B*, vol. 50, pp. 157-224, 1988.
[13] S. J. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach. Upper Saddle River," *EUA: Prentice-Hall*, 2003.
[14] R. Neapolitan, *Learning Bayesian Networks*, Prentice Hall. 2003.
[15] W. M. Leung and L. K. Sak, "An Efficient Data Mining Method for Learning Bayesian Networks using an Evolutionary Algorithm-based Hybrid Approach," *Evolutionary Computation, IEEE Transactions on*, vol. 8, pp. 378-404, 2004.
[16] J. Cheng, "Learning Bayesian Networks from Data: An Information-theory Based Approach," *Artificial Intelligence*, vol. 137, pp. 43-90, 2002.
[17] KDD Cup 1999 Data. [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
[18] Y. Ma, D. Choi, S. Ata, and H. Nguyen, "Application of Data Mining to Network Intrusion Detection: Classifier Selection Model," In *Challenges for Next Generation Network Operations and Service Management*, pp. 399-408, 2008.
[19] About GeNIe and Smile. [Online]. Available: http://genie.sis.pitt.edu/about.html#genie
[20] Department Web Server. [Online]. Available: http://www.phil.cmu.edu/projects/tetrad/tetrad4.html
[21] Belief Network PowerConstructor. [Online]. Available: 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, Melaka campus, Multimedia University, 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.