A Study of Intrusion Detection in Data Mining

E.Kesavulu Reddy, Member IAENG, V.Naveen Reddy, P.Govinda Rajulu

Abstract: Network security technology has become crucial in protecting government and industry computing infrastructure. Modern intrusion detection applications facing complex problems. These applications has to be require reliable, extensible, easy to manage, and have low maintenance cost. In recent years, data mining-based intrusion detection systems (IDSs) have demonstrated high accuracy, good generalization to novel types of intrusion, and robust behavior in a changing environment. Still, significant challenges exist in the design and implementation of production quality IDSs. Instrumenting components such as of data transformations, model deployment, cooperative distributed detection and complex engineering endeavor.

Key Words: Data Mining, Intrusion Detection, Knowledge Discovery Database, Patterns

1. INTRODUCTION

Intrusion detection is the process of monitoring and analyzing the events occurring in a computer system in order to detect signs of security problems .The intrusion detection and other security technologies such as cryptography, authentication, and firer walls has gained in importance in last ten years. However, intrusion detection is not yet a perfect technology.

Intrusion detection is an area growing in relevance as more and more sensitive data are stored and processed in networked systems. An intrusion detection system (IDS) monitors networked devices and looks for anomalous or malicious behavior in the patterns of activity in the audit stream. A comprehensive IDS requires a significant amount of human expertise and time for development. Data mining-based IDSs require less expert knowledge yet provide good performance [35][5][29][13]. These systems are also capable of generalizing to new and unknown attacks. Data mining-based intrusion Building an IDS is a complex task of knowledge engineering that requires an elaborate infrastructure:

An effective contemporary production-quality IDS needs an array of diverse components and features, including

- Centralized view of the data
- Data transformation capabilities
- Analytic and data mining methods
- Flexible detector deployment, including scheduling that enables periodic model relation and distribution
- Real-time detection and alert infrastructure
- Reporting capabilities
- Distributed processing
- High system availability
- Scalability with system load

II. DATA MINING, KDD, AND RELATED FIELDS

The term data mining is frequently used to designate the process of extracting useful information from large databases. In this chapter, we adopt a slightly different view, which is identical to the one expressed by [14]. In this view, the term knowledge discovery in databases (KDD) is used to denote the process of extracting useful knowledge from large data sets. Data mining, by contrast, refers to one particular step in this process. Spherically, the data mining step applies so-called data mining techniques to extract patterns from the data. Additionally, it is preceded and followed by other KDD steps, which ensure that the extracted patterns actually correspond to useful knowledge. Indeed, without these additional KDD steps, there is a high risk of finding meaningless or uninteresting patterns [15]. In other words, the KDD process uses data mining techniques along with any required pre- and post-processing to extract high-level knowledge from low-level data. In practice, the KDD process is interactive and iterative, involving numerous steps with many decisions being made by the user [14]. Here, we broadly outline some of the most basic KDD steps:

1. Understanding the application domain: First is developing an understanding of the application domain, the relevant background knowledge, and the specific goals of the KDD endeavor.
2. Data integration and selection: Second is the integration of multiple data sources and the selection of the subset of data that is relevant to the analysis task.
3. Data mining: Third is the application of specific algorithms for extracting patterns from data.
4. Pattern evaluation: Fourth is the interpretation and validation of the discovered patterns. The goal of this step is to guarantee that actual knowledge is being discovered.
5. Knowledge representation: This step involves documenting and using the discovered knowledge.

Manuscript received Jan 27 2011, Revised Feb 24 2011

Asst. Prof. E. Kesavulu Reddy, Research Scholar, Computer Science, S.V.University College of CM & CS, Tirupati,Andhra Pradesh, India-517502. Mobile: +91 9866430097
Email: ekreddysvu2008@gmail.com

Prof. P. Govindarajulu, Vice-Principal, Computer Science, S.V.U.College of CM&CS, Tirupat, A.P., India-517502, Mobile: +919912349770
Email: pgovindarajulu@yahoo.com

In other words, data mining emphasizes the efficient discovery of simple, but understandable models that can be interpreted as interesting or useful knowledge. In fact, data mining is just a step in the KDD process. As such, it has to contribute to the overall goal of knowledge discovery.

Some Data Mining Techniques: Data mining techniques essentially are pattern discovery algorithms. Some techniques such as association rules [1] are unique to data mining, but most are drawn from related fields such as machine learning or pattern recognition. In this section, we introduce four well-known data mining techniques that have been widely used in intrusion detection. A broader and more detailed treatment of data mining techniques can be found elsewhere [19][27][3].

A potential source of confusion is that different data mining techniques assume different input data representations. For example, association rules have historically been discussed under the assumption that the input data is represented as a set of transactions [1][2]. Later, association rule mining over relational databases has been investigated [2]. Depending on the input data representations (sets of transactions versus relational databases), the association rule concept is presented differently. A related problem is that there are many different ways to represent the same data set in a relational database. In practice, the available input data does not necessarily follow this format. Then, it is the responsibility of the second KDD step to transform the available data into the format required by the data mining techniques.

- Association Rules
- Frequent Episode Rules
- Classification
- Clustering

III. DATA MINING MEETS INTRUSION DETECTION

The goal of intrusion detection is to detect security violations in information systems. Intrusion detection is a passive approach to security as it monitors information systems and raises alarms when security violations are detected. Examples of security violations include the abuse of privileges or the use of attacks to exploit software or protocol vulnerabilities. Traditionally, intrusion detection techniques are classified into two broad categories: misuse detection and anomaly detection [32].

Misuse detection works by searching for the traces or patterns of well- known attacks. Clearly, only known attacks that leave characteristic traces can be detected that way. Anomaly detection, on the other hand, uses a model of normal user or system behavior and ages significant deviations from this model as potentially malicious. This model of normal user or system behavior is commonly known as the user or system profile. A strength of anomaly detection is its ability to detect previously unknown attacks. Additionally, intrusion detection systems (IDSs) are categorized according to the kind of input information they analyze. This leads to the distinction between host-based and network-based IDSs. Host-based IDSs analyze host-bound audit sources such as operating system audit trails, system logs, or application logs. Network-based IDSs analyze network packets that are captured on a network. More information on intrusion detection in general can be found, for example, in a recent book by Bace (2000). In the past five years, a growing number of research projects have applied data mining to intrusion detection. Here, we survey a representative cross section of these projects. The intention of this survey is to give the reader a broad overview of the work that has been done at the intersection between intrusion detection and data mining.

IV. IDS TAXONOMY

The goal of an ID is to detect malicious traffic. In order to accomplish this, the IDS monitor all incoming and outgoing traffic. There are several approaches on the implementation of IDS. Among those, two are the most popular: Anomaly detection is based on the detection of traffic anomalies. The deviation of the monitored traffic from the normal profile is measured. Various different implementations of this technique have been proposed, based on the metrics used for measuring traffic profile deviation. Misuse/Signature detection: looks for patterns and signatures of already known attacks in the network traffic. A constantly updated database is usually used to store the signatures of known attacks. The way this technique deals with intrusion detection resembles the way that anti-virus software operates. Figure 1 shows taxonomy of Intrusion Detection Systems. More details and information on the various IDS systems and the way they work can be found in [31].

V. DRAWBACKS OF IDSS

Intrusion Detection Systems (IDS) have become a standard component in security infrastructures as they allow network administrators to detect policy violations. These policy violations range from external attackers trying to gain unauthorized access to insiders abusing their access. Current IDS have a number of significant drawbacks: • Current IDS are usually tuned to detect known service level network attacks. This leaves them vulnerable to original and novel malicious attacks.

• Data overload: Another aspect which does not relate directly to misuse detection but is extremely important is how much data an analyst can efficiently analyze. That amount of data he needs to look at seems to be growing rapidly. Depending on the intrusion detection tools employed by a company and its size there is the possibility for logs to reach millions of records per day.

• False positives: A common complaint is the amount of false positives an IDS will generate. A false positive occurs when normal attack is mistakenly classified as malicious and treated accordingly.

• False negatives: This is the case where an IDS does not generate an alert when an intrusion is actually taking place. (Classification of malicious traffic as normal) Data mining can help improve intrusion detection by addressing each and every one of the above mentioned problems. Remove normal activity from alarm data to allow analysts to focus on real attacks • Identify false alarm generators and ”bad” sensor signatures • Find anomalous activity that uncovers a real attack • Identify long, ongoing patterns (different IP address, same activity) To accomplish these tasks, data miners employ one or more of the following techniques:

- Data summarization with statistics, including finding outliers
- Visualization: presenting a graphical summary of the data
- Clustering of the data into natural categories
- Association rule discovery: defining normal activity and enabling the discovery of anomalies
- Classification: predicting the category to which a particular record belongs

VI. DATA MINING AND IDS

Data mining techniques can be differentiated by their different model functions and representation, preference criterion, and algorithms [17]. The main function of the model that we are interested in is classification, as normal, or malicious, or as a particular type of attack [18]. We are also interested in link and sequence analysis [12]. Additionally, data mining systems provide the means to easily perform data summarization and visualization, aiding the security analyst in identifying areas of concern [12]. The models must be represented in some form. Common representations for data mining techniques include rules, decision trees, linear and non-linear functions (including neural nets), instance-based examples, and probability models [17].

VII. SURVEY OF APPLIED TECHNIQUES

In this section we present a survey of data mining techniques that have been applied to IDSs by various research groups**.**

A. Feature Selection

”Feature selection, also known as subset selection or variable selection, is a process commonly used in machine learning, wherein a subset of the features available from the data is selected for application of a learning algorithm. Feature selection is necessary either because it is computationally infeasible to use all available features, or because of problems of estimation when limited data samples (but a large number of features) are present.”

B. Machine Learning

Machine Learning is the study of computer algorithms that improve automatically through experience. Applications range from data mining programs that discover general rules in large data sets, to information filtering systems that automatically learn users’ interests. In contrast to statistical techniques, machine learning techniques are well suited to learning patterns with no a priori knowledge of what those patterns may be. Clustering and Classification are probably the two most popular machine learning problems. Techniques that address both of these problems have been applied to IDSs.

VIII.CLASSIFICATION OF TECHNIQUES

In a classification task in machine learning, the task is to take each instance of a dataset and assign it to a particular class. A classification based IDS attempts to classify all traffic as either normal or malicious. The challenge in this is to minimize the number of false positives (classification of normal traffic as malicious) and false negatives (classification of malicious traffic as normal). Five general categories of techniques have been tried to perform classification for intrusion detection purposes:

A. Inductive Rule Generation

The RIPPER System is probably the most popular representative of this classification mechanism. RIPPER [7], is a rule learning program. RIPPER is fast and is known to generate concise rule sets. It is very stable and has shown to be consistently one of the best algorithms in past experiments [8]. The system is a set of association rules and frequent patterns than can be applied to the network traffic to classify it properly. One of the attractive features of this approach is that the generated rule set is easy to understand, hence a security analyst can verify it. Another attractive property of this process is that multiple rule sets may be generated and used with a meta-classifier [22] [24] [23] [25].

B.Genetic Algorithms

Genetic algorithms were originally introduced in the field of computational biology. Since then, they have been applied in various fields with promising results. Fairly recently, researchers have tried to integrate these algorithms with IDSs. • The REGAL System [33][34] is a concept learning system based on a distributed genetic algorithm that learns First Order Logic multi-modal concept descriptions. REGAL uses a relational database to handle the learning examples that are represented as relational tuples. Dasgupta and Gonzalez [10] used a genetic algorithm, however they were examining host-based, not network-based IDSs. Instead of running the algorithm directly on the feature set, they used it only for the meta-learning step, on labeled vectors of statistical classifiers**.** Each of the statistical classifiers was a 2-bit binary encoding of the abnormality of a particular feature, ranging from normal to dangerous.

• Chittur [6] applied a genetic algorithm and used a decision tree to represent the data. They used the ”Detection rate minus the false positive rate” as their preference criterion to distinguish among the data.

• Crosbie and Spafford [9] also used a genetic algorithm for sparse trees to detect anomalies. They attempted to minimize the occurrence of false positives by utilizing human input in a feedback loop.

C.Fuzzy Logic

Fuzzy logic is derived from fuzzy set theory dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic. It can be thought of as ”the application side of fuzzy set theory dealing with well thought out real world expert values for a complex problem” [21].

In Dickerson and Dickerson 2000 [11] the authors classify the data based on various statistical metrics. They then create and apply fuzzy logic rules to these portions of data to classify them as normal or malicious. They found that the approach is particularly effective against scans and probes. An enhancement of the fuzzy data mining approach has also been applied by Florez et al. The authors use fuzzy data mining techniques to extract patterns that represent normal behavior for intrusion detection. They describe a variety of modifications that they have made to the data mining algorithms in order to improve accuracy and efficiency. They use sets of fuzzy association rules that are mined from network audit data as models of ”normal behavior.” To detect anomalous behavior, they generate fuzzy association rules from new audit data and compute the similarity with sets mined from ”normal” data. If the similarity values are below a threshold value, an alarm is issued. They describe an algorithm for computing fuzzy association rules based on Borgelt’s prefix trees, modifications to the computation of support and confidence of fuzzy rules, a new method for computing the similarity of two fuzzy rule sets, and feature selection and optimization with genetic algorithms. Experiments showed that their approach not only reduces the number of rules, but also increases the accuracy of the system.

Luo also attempted classification of the data using Fuzzy logic rules. He demonstrated that the integration of fuzzy logic with association rules and frequency episodes generates more abstract and flexible patterns for anomaly detection. He also added a normalization step to the procedure for mining fuzzy association rules by Kuok, Fu, and Wong [35] in order to prevent one data instance from contributing more than others. He modified the procedure of Mannila and Toivonen for mining frequency episodes to learn fuzzy frequency episodes. His approach utilizes fuzzy association rules and fuzzy frequency episodes to extract patterns for temporal statistical measurements at a higher level than the data level. Finally he presented the first real-time intrusion detection method that uses fuzzy episode rules.

D. Neural Networks

The application of neural networks for IDSs has been investigated by a number of researchers. Neural networks provide a solution to the problem of modeling the users’ behavior in anomaly detection because they do not require any explicit user model. Neural networks for intrusion detection were first introduced as an alternative to statistical techniques in the IDES intrusion detection expert system to model. In particular, the typical sequence of commands executed by each user is learned. Numerous projects have used neural nets for intrusion detection using data from individual hosts, such as BSM data [18].McHugh et al have pointed out that advanced research issues on IDSs should involve the use of pattern recognition and learning by example approaches for the following two main reasons:

The capability of learning by example allows the system to detect new types of intrusion. • With earning by example approaches, attack”signatures” can be extracted automatically from labeled traffic data. This basically eliminates the subjectivity and other problems introduced by the presence of the human factor. A different approach to anomaly detection based on neural networks is proposed by Lee et al. While previous works have addressed the anomaly detection problem by analyzing the audit records produced by the operating system, in this approach, anomalies are detected by looking at the usage of network protocols.

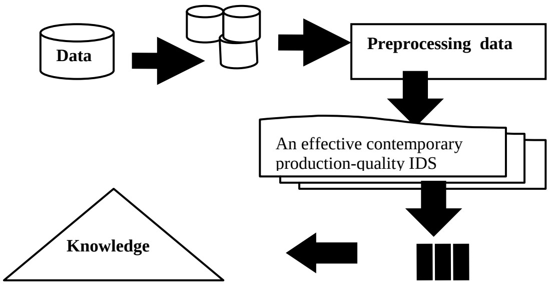

Figure 1: Data to Knowledge

E. Immunological based techniques

Hofmeyr and Forrest present an interesting technique based on immunological concepts. They define the set of connections from normal traffic as the”self”, then generate a large number of ”non-self” examples: connections that are not part of the normal traffic on a machine. These examples are generated using a byte oriented hash and permutation. They can then compare incoming connections using the r-contiguous bits match rule. If a connection matches one of the examples, it is assumed to be in non-self and marked as anomalous. Dasgupta and Gonzalez [45] used a similar approach. The authors generated a set of fuzzy rules using a genetic algorithm. They found that while this approach was not as accurate as a nearest neighbor match with the self-set, it was significantly more efficient. Fan also used a similar approach in [46]. He found that injecting artificial anomalies into the dataset significantly in increased detection of malicious anomalies, including those that had never been seen before.

F. Support Vector Machine

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. They belong to a family of generalized linear classifiers. SVMs attempt to separate data into multiple classes (two in the basic case) though the use of a hyper-plane. Eskin et al. , and Honig et al. [13] used an SVM in addition to their clustering methods for unsupervised learning. The achieved performance was comparable to or better than both of their clustering methods. Mukkamala, Sung, et al. used a more conventional SVM approach. They used five SVMs, one to identify normal traffic, and one to identify each of the four types of malicious activity in the KDD Cup dataset. Every SVM performed with better than 99% accuracy, even using seven different variations of the feature set. As the best accuracy they could achieve with a neural network (with a much longer training time) was 87.07%, they concluded that SVMs are superior to neural nets in both accuracy and speed.

G. Clustering Techniques

Data clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is the classification of similar objects into different groups, or more precisely, the partitioning of a data set into subsets (clusters), so that the data in each subset (ideally) share some common trait - often proximity according to some defined distance measure. Machine learning typically regards data clustering as a form of unsupervised learning. Clustering is useful in intrusion detection as malicious activity should cluster together, separating itself from non-malicious activity. Clustering provides

Some significant advantages over the classification techniques already discussed, in that it does not require the use of a labeled data set for training. Frank breaks clustering techniques into five areas: hierarchical, statistical, exemplar, distance, and conceptual clustering, each of which has different ways of determining cluster membership and representation. Portnoy et al present a method for detecting intrusions based on feature vectors collected from the network, without being given any information about classifications of these vectors. They designed a system that implemented this method**,** and it was able to detect a large number of intrusions while keeping the false positive rate reasonably low. There are two primary advantages of this system over signature based classifiers or learning algorithms that require labeled data in their training sets. The first is that no manual classification of training data needs to be done. The second is that we do not have to be aware of new types of intrusions in order for the system to be able to detect them. All that is required is that the data conform to several assumptions. The system tries to automatically determine which data instances fall into the normal class and which ones are intrusions. Even though the detection rate of the system they implemented is not as high as of those using algorithms relying on labeled data, they claim it is still very useful. Since no prior classification is required on the training data, and no knowledge is needed about new attacks, the process of training and creating new cluster sets can be automated. In practice, this would mean periodically collecting raw data from the network, extracting feature values from it, and training on the resulting set of feature vectors. This will help detect new and yet unknown attacks.

H. Statistical Techniques

Statistical techniques, also known as ”top-down” learning, are employed when we have some idea as to the relationship were looking for and can employ mathematics to aid our search. Three basic classes of statistical techniques are linear, nonlinear (such as a regression-curve), and decision trees. Statistics also includes more complicated techniques, such as Markov models and Bayes estimators. Statistical patterns can be calculated with respect to different time windows, such as day of the week, day of the month, month of the year, etc. or on a per-host, or per-service basis [23]. Denning (1987) described how to use statistical measures to detect anomalies, as well as some of the problems and their solutions in such an approach. The five statistical measures that she described were the operational model, the mean and standard deviation model, the multivariate model, the Markov process model, and the time series model. Javitz and Valdes provide more details on the individual statistical measures used in ID. They also provide formulas for calculating informative statistic metrics. Staniford et al uses a similar approach by employing a Bayes network to calculate the conditional probabilities of various connection features with respect to other connection features.

These probabilities are then used to determine how anomalous each connection is. Mahoney and Chan combined the output of five specific probability measures to determine how anomalous each connection was. In they generate a set of rules for normal traffic where each rule retains the percentage of records in the training stream that support it. When a record is detected that violates a given rule, its anomaly score is the sum of each rules support value times the time since that rule was last violated. Sinclair et al.[12] describe how they used Quinlan’s ID3 algorithm to build a decision tree to classify network connection data. Bloedorn et al and Barbara et al. also use decision tree-based methods.

I. Hidden Markov Models

Much work has been done or proposed involving Markovian models. For instance, the generalized Markov chain may improve the accuracy of detecting statistical anomalies. Unfortunately, it has been noted that these are complex and time consuming to construct [20], however their use may be more feasible in a high-power off-line environment. A hidden Markov model (HMM) is a statistical model where the system being modeled is assumed to be a Markov process with unknown parameters, and the challenge is to determine the hidden parameters from the observable parameters. The extracted model parameters can then be used to perform further analysis, for example for pattern recognition applications. A HMM can be considered as the simplest dynamic Bayesian network.

Hidden Markov Models (HMM) are to detect complex Internet attacks. These attacks consist of several steps that may occur over an extended period of time. Within each step, specific actions may be interchangeable. A perpetrator may deliberately use a choice of actions within a step to mask the intrusion. In other cases, alternate action sequences may be random (due to noise) or because of lack of experience on the part of the perpetrator. For an intrusion detection system to be effective against complex Internet attacks, it must be capable of dealing with the ambiguities described above. The authors describe research results concerning the use of HMMs as a defense against complex Internet attacks. They describe why HMMs are particularly useful when there is an order to the actions constituting the attack (that is, for the case where one action must precede or follow another action in order to be effective). Because of this property, they show that HMMs are well suited to address the multi-step attack problem. In a direct comparison with two other classic techniques, decision trees and neural nets, the authors show that HMMs perform generally better than decision trees and substantially better than neural networks in detecting these complex intrusions.

IX. CONCLUSION

This paper has presented a survey of the various data mining techniques that have been proposed towards the enhancement of IDSs. We have shown the ways in which data mining has been known to aid the process of Intrusion Detection and the ways in which the various techniques have been applied and evaluated by researchers

REFERENCES

[1] Agrawal, R., Imielinski, T., and Swami, A. (1993). Mining Associations between Sets of Items in Massive Databases. In Proceedings of the ACM-SIGMOD 199International Conferencing Management of Data, pages 207{216.

[2] Agrawal, R. and Srikant, R. (1994). Fast Algorithms for Mining Association Rules. In Proceedings of the 20th International Conference on Very Large Databases, pages 487{499.

[3] Berry, M. J. A. and Lino_,G. (1997). Data Mining Techniques. John Wiley and Sons, Inc.

[4] Biswanath Mukherjee, L.Todd Heberlein, Karl .Levitt, ”Network Intrusion Detection”,IEEE, June 1994.

[5] Barbarà, D., Couto, J., Jajodia, S., Popyack, L., And Wu,N., ADAM: Testbed for Exploring the Use of Data Miningin Intrusion Detection, ACM SIGMOD Record, 30(4), 2001,pp. 15-24.

[6] Chittur, A., ”Model generation for an intrusion detection system using genetic algorithms”, High School Honors Thesis, Ossining High School. In cooperation with Columbia Univ, 2001

[7] Cohen, W. W. (1995).Fast E_ective Rule Induction. In Proceedings 12th International Conference on Machine Learning, pages 115{123. Elmasri,R and Navathe, S. B. (1994). Fundamentals of Database Systems. Addison-Wesley

[8] Chittur, A., ”Model generation for an intrusion detection system using genetic algorithms”, High School Honors Thesis, Ossining High School. In cooperation with Columbia Univ, 2001.

[9] Crosbie, M. and E. H. Spafford, ”Active defense of a computer system using autonomous agents”, Technical Report CSD-TR- 95-008, Purdue Univ., West Lafayette, IN, 15 February 1995.

[10] Dasgupta, D. and F. A. Gonzalez, ”An intelligent decision support system for intrusion detection and response”, . In Proc. of International Workshop on Mathematical Methods, Models and Architectures for Computer Networks Security (MMM-ACNS), St.Petersburg. Springer- , 21-23 May,2001.

[11] Dickerson, J. E. and J. A. Dickerson, ”Fuzzy network profiling for intrusion detection”, In Proc. of NAFIPS 19th International Conference of the North American Fuzzy Information Processing Society, Atlanta, pp. 301306. North American Fuzzy Information Processing Society (NAFIPS), July 2000.

[12] Eric Bloedorn et al, ”Data Mining for Network Intrusion Detection: How to Get Started,” Technical paper, 2001.

[13] Eskin, E., Arnold, A., Prerau, M., Portnoy, L., and Stolfo, S. J., A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data, In D.Barbarà and S. Jajodia (eds.), Applications of Data Mining in Computer Security, Kluwer Academic Publishers, Boston, MA, 2002, pp. 78-99.

[14] . Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., editors 1996b). Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press.

[15] Fayyad, U. (1998). Mining Databases: Towards Algorithms for Knowledge Discovery. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 22(1):39-48.

[16] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996a). From Data Mining to Knowledge Discovery in Databases. AI Magazine, 17(3):37 54.

[17] Fayyad, U. M., G. Piatetsky-Shapiro, and P. Smyth, ”The KDD process for extracting useful Knowledge from volumes of data,” Communications of the ACM 39 (11),November 1996, 2734.

[18] Ghosh, A. K., A. Schwartzbard, and M. Schatz,” Learning program behavior profiles for intrusion detection”, In Proc. 1st USENIX, 9-12 April, 1999.

[19] Han, J. and Kamber, M. (2000). Data Mining: Concepts and Techniques, Morgan Kaufmann Publisher.

[20] Kumar, S., ”Classification and Detection of Computer Intrusion”, PhD. thesis, 1995, Purdue Univ., West Lafayette, IN.

[21] G. J. Klir, ”Fuzzy arithmetic with requisite constraints”, Fuzzy Sets and Systems, 91:165175, 1997.

[22] Lee, W. and S. J. Stolfo, ”Data mining approaches for intrusion detection”, In Proc. of the 7th USENIX Security Symp., San Antonio, TX.USENIX, 1998.

[23] W. Lee, S.J.Stolfo et al, ”A data mining and CIDF based approach for detecting novel and distributed intrusions”, Proc. of Third International Workshop on Recent advances in Intrusion Detection (RAID 2000), Toulouse, France.

[24] Lee, W., S. J. Stolfo, and K. W. Mok, ” Mining in a data- flow environment: Experience in network intrusion detection,” In S. Chowdhury and D. Madigan (Eds.), Proc. of the Fifth International Conference on Knowledge Discovery and Data Mining (KDD-99), San Diego, CA, pp. 114124. ACM,12-15 August 1999.

[25] W., S. J. Stolfo, and K. W. Mok, ”Adaptive intrusion detection: A data mining approach,” Artificial Intelligence Review 14 (6), 533567, 2000.

[26] Mannila, H. (1996). Data Mining: Machine Learning, Statistics, and Databases. In Proceedings of the 8th International Conference on Scienti_c and Statistical Database Management, pages 1{8.}

[27] Mannila, H., Smyth, P., and Hand, D. J. (2001). Principles of Data Mining. MIT Press. Mannila, H., Toivonen, H., and Verkamo, A. I. (1997).

[28] Discovery of FrequentEpisodes in Event Sequences Data Mining and Knowledge Discovery, :259-289.

[29] Markou, M. and Singh, S., Novelty Detection: A review, Part 1: Statistical Approaches, Signal Processing, 8(12), 2003, pp. 2481-2497.

[30] Miller, R. and Yang, T. (1997). Association Rules Over Interval Data. In Proceedings of the 1997 ACM- SIGMOD Conference on Management of Data, pages 452{461.

[31] Mithcell Rowton, Introduction to Network Security Intrusion Detection, December 2005.

[32] Mounji, A. (1997). Languages and Tools for Rule-Based Distributed Intrusion Detection. PhD thesis, Faculties Universitaires Notre-Dame dela Paix Namur (Belgium).

[33] Neri, F., ”Comparing local search with respect to genetic evolution to detect intrusion in computer networks”, In Proc. of the 2000 Congress on Evolutionary Computation CEC00, La Jolla, CA, pp. 238243. IEEE Press, 16-19 July, 2000.

[34] Neri, F., ”Mining TCP/IP traffic for network intrusion detection”, In R. L. de M’antaras and E. Plaza (Eds.), Proc. of Machine Learning: ECML 2000,11th European Conference on Machine Learning, Volume 1810 of Lecture Notes in Computer Science, Barcelona, Spain, pp. 313322. Springer, May 31- June 2, 2000.

[35] Noel, S., Wijesekera, D., and Youman, C., Modern Intrusion Detection, Data Mining, and Degrees of Attack Guilt, In D. Barbarà and S. Jajodia (eds.), Applications of Data Mining in Computer Aecurity, Kluwer Academic Publishers, Boston, MA, 2002, pp. 2-25.