Designing a Two-Level Monitoring Method to Detect Network Abnormal Behaviors

Soo-Yeon Ji
Bowie State University
Bowie, Maryland 20715–3319
Email: sji@bowiestate.edu

Seonho Choi
Bowie State University
Bowie, Maryland 20715–3319
Email: schoi@bowiestate.edu

Dong Hyun Jeong
University of the District of Columbia
Washington, DC 20008–1122
Email: djeong@udc.edu

***Abstract*—Monitoring network traffic behavior is very critical for securing computing infrastructures. In this paper, we focus on enhancing the way of detecting anomalous network traffic behaviors by proposing a new two-level detection method that consists of abnormality detection and exact attack type identification. The abnormality detection is performed with the rules generated by Classification and Regression Trees (CART). Then, Support Vector Machine (SVM) is applied to design a predictive model to identify exact attack types (among DoS, U2R, R2L, and Probes). Since feature extraction is an important step for designing an efficient predictive model, we used Higuchi fractal dimension and statistical measures (mean, median, and standard deviation) with an overlapping sliding window operation to extract features. Among the extracted features, only significant features are selected by applying statistical analysis and used to design a predictive model. As results, we found that our approach shows about 80.03% accuracy in detecting network abnormal behaviors. From a comparative study, we concluded that our proposed SVM-based predictive model is superior to a broadly known NN-based predictive model for identifying exact types of attacks.**

I. Introduction

Since the Internet become an important part of our daily lives, it is important to secure our computing resources and infrastructures from any external cyber threats. Due to this importance, numerous researchers studied on understanding and detecting abnormal network behaviors. They proposed techniques to differentiate anomalous network patterns from heavy network traffic load situations. A traditionally known approach to network anomaly detection is based on utilizing attack signatures [1]. Although this signature-based technique produces a lower false alarms rate, it is not as effective as for detecting unknown anomalous network attacks because signature-based detection systems identify network anomalies by referencing built-in attack signatures. If signatures do not exist in the signature-based detection systems, incoming network anomalies cannot be detected. Therefore, signatures need to be updated to the systems continuously to improve the chance of detecting new types of attacks.

Feature-based detection technique is proposed to address the limitation that exists in the signature-based detection systems. It identifies network anomalies by examining network traffic features. Commonly used network traffic features are IP addresses, source/ destination port numbers, and TCP flags. Known feature-based techniques detect anomalous network traffics by observing and comparing them to the features in normal traffic conditions. However, these techniques fall into drawbacks of identifying anomalous network traffic patterns correctly since the patterns include volatile information to hide their uniqueness. For instance, IP address is a unique number assigned to each computer on the network. This IP address information can be changed or hidden by hackers when penetrating network infrastructures. Although several techniques are proposed to identify different characteristics from network traffic patterns, this is still a known research challenge. A detailed explanation about known feature-based network anomaly detection techniques is included in Section 2 Literature Review.

In this paper, we propose a two-level network abnormality monitoring model to identify anomalous network behaviors. First, rules are generated to determine outcomes (normal/ abnormal). Then, a predictive model is designed to identify exact attack types. Specifically, a decision tree is used to generate reliable rules and machine learning (ML) is applied to build a predictive model with utilizing only statistically significant features. With this predictive model, we found that it is possible to identify exact abnormal attack types among denial-of-service (DoS), unauthorized access from a remote machine (R2L), unauthorized access to local administrative privileges (U2L), and surveillance and other probing (Probes).

The rest of this paper consists of four sections. Literature review on network anomaly detection is included in Section II. Section III begins with explaining the network traffic data we used for this study. Then, a detailed explanation about our proposed approach is provided. We conclude this paper after providing experimental results and future works.

II. Literature Review

As mentioned above, there are two globally known techniques in network anomaly detection as signature-based and feature-based detection techniques [2]. The signature-based technique is a common method used in Intrusion Detection System (IDS) that can identify attacks with referencing known signature patterns. Although signature-based approach is good for identifying network anomaly matched to pre-defined signatures, it has a limitation of detecting unknown network anomalies. To avoid this limitation, the feature-based technique is used because it does not require any prior knowledge. Numerous studies have been performed to identify effective approaches for extracting anomalous network traffic features. Among them, statistical methods and machine learning techniques are broadly used to identify anomalous network traffic features.

In the past, researchers studied on increasing the rate of detecting network attacks. Cheng et al. [3] proposed an approach of identifying normal TCP flows by using spectral analysis techniques to protect legitimate TCP flow from Denial of Service (DoS) attacks. Wang et al. [4] proposed statistics-based approach to detect TCP SYN flood attacks, which uses a nonparametric cumulative sum (CUSUM) method. Utilization of statistical approaches is good for maintaining high accuracy with spending reasonably short detection times because it approximately calculates normal traffic patterns to perform a comparison with abnormal traffics. However, it has a difficulty of detecting anomalies caused by network system failures. To resolve this difficulty, Thottan and Ji [5] used a statistical data analysis method with a signal processing technique together to quantify network behaviors to understand network anomalies. They classified network anomalies into two categories as network performance anomalies (e.g. file server failures, paging across the network, broadcast storms) and security-related problems or attacks. They showed that their approach of integrating a signal processing technique is effective for detecting several network anomalies. However, there was no accurate statistical model that was utilized to detect different abnormal traffic patterns. Due to this limitation, researcher started applying various techniques including neural networks, machine learning techniques, data mining, and so forth.

Artificial neural network has been applied broadly because it has a potential of identifying and classifying unknown network activities [6]. Lippmann and Cunningham [7] utilized neural networks to design a detection model by searching for attack-specific keywords in network traffic. Sarasamma et al. [8] used multilevel hierarchical Kohonen Net (K-Map) consisting of three layers to determine different types of attacks. To increase the speed of selecting features, input dataset is divided into three feature sets based on domain knowledge. After applying single-layer K-Map onto each feature set, significant subset features are determined and used to design next hierarchical K-Map. Hand and Cho [9] proposed an approach of employing an evolutionary neural network (ENN) to overcome the limitation of designing a precise topology (i.e. domain specific neural network model) for detecting network attacks. Support Vector Machine (SVM) is also used to classify abnormal network behaviors. Since SVM supports both supervised and unsupervised learning, Shon and Moon [10] applied hybrid approach of integrating the two learning methods with emphasizing the advantages of utilizing both SVM approaches. Similarly, Jain and Abouzakhar [11] designed an approach by utilizing both Hidden Markov Model (HMM) and Support Vector Machine (SVM).

Although numerous network anomaly detection method were proposed, there is no unique solution that maintains higher detection rate and lower false positive and negative rates. Traditional approaches to network anomaly detection utilize information that is directly extracted from network packet header. However, to increase the accuracy of detecting network anomalies, integration with computational feature extraction techniques is emphasized [12]. Shon and Moon [10] used Genetic Algorithm (GA) to extract optimized information from raw internet packets. Jain and Abouzakhar [11] applied J48 decision tree algorithm to determine significant features for anomaly intrusion detection. Hofmann et al. [12] proposed an approach with the combination of evolutionary algorithm (EA) and radial basis-function networks (RBFN). The evolutionary algorithm is used to select an appropriate feature subset, optimize the number of hidden neurons, determine a number of training epochs, and choose a basis function type. Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) are also broadly applied techniques for feature extraction [13]. Most proposed techniques utilize characteristics of network traffics to identify abnormalities precisely. But, performing the real-time network anomaly detection with maintaining higher accuracy is limited due to the complex nature of network traffics. In this paper, we primarily focus on enhancing the way of detecting network anomalies by integrating both a rule extraction technique and a predictive model.

III. Methodology

Identifying abnormal behavior from network traffic is critical to maintaining a network environment secure. Abnormal behaviors detection should be performed with satisfying performance and reliability. Specifically, detailed information about detected abnormal behaviors should be provided when notifying detected abnormal behaviors to protect computing infrastructures efficiently. With our proposed predictive model, identification of exact attack types is performed to address this need. Our approach begins with generating reliable rules for detecting network abnormality (normal/abnormal) with classification and regression trees (CART). Once its abnormality is detected, an over-lapping sliding window operation is applied to extract features. Although the extracted features might contain the characteristics of abnormal behaviors, it is important to utilize only significant features to increase the accuracy of determining exact types of abnormal behaviors. In our study, we used Statistical Analysis System (SAS) to identify significant features, with which our proposed predictive model is designed. Figure 1 illustrates how the predictive model is designed. A detailed explanation about each step is included in the following subsections.

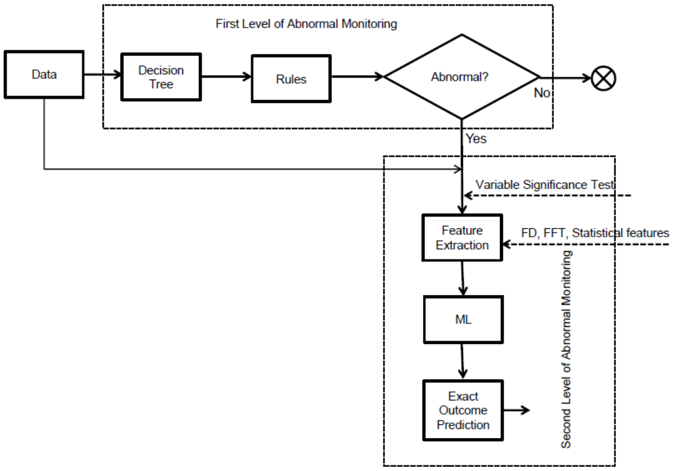

First Level of Abnormal Monitoring
Data
Decision Tree
Rules
Abnormal?
No
Yes
Variable Significance Test
Feature Extraction
FD, FFT, Statistical features
ML
Exact Outcome Prediction
Second Level of Abnormal Monitoring

Fig. 1. The entire process of the proposed predictive model consisting of two levels.

A. Dataset

In this study, we used a new version of KDD dataset, NSL-KDD dataset, which is publicly available for researchers [14],

[15]. Since the KDD Cup'99 intrusion detection dataset includes a lot of redundant records, this redundancy causes an inefficient learning process. For instance, the difference between unauthorized access from a remote machine (R2L) and unauthorized access to local administrative privileges (U2L) is not clear. Due to this reason, it is difficult to distinguish them precisely. To resolve this issue, repeated records are removed from the KDD Cup'99 dataset [14]. We used total 125,973 records for training. For the testing, the rest of the data records (22,544 records) were used. The NSL-KDD dataset includes total 41 attributes (three nominal, six binary, and thirty-two numeric attributes) and four types of attacks (Denial of service (DoS), user to root (U2R), remote to user (R2L), and Probes). DOS attack indicates attempts to disabling network access to remote machines (or computing resources). R2L represents that a remote user gains access to local user accounts by sending packets to a computing machine over the network. U2R explains that an attacker accesses normal users' accounts by exploring the system as a root-user. Probing represents that the network is scanned to gather information to find known vulnerabilities.

TABLE I. DATA DISTRIBUTION OF THE NSL-KDD TRAINING AND TESTING DATASETS.

	Normal	Abnormal			
	-	DoS	U2R	R2L	Probes
Training	53.50%	37.20%	0.04%	0.08%	10.80%
Testing	43.10%	33.10%	0.20%	12.80%	10.70%

Table I shows the sampling distribution of the NSL-KDD training and testing datasets. As shown, the total number of attacks are not equally distributed in the training and testing datasets. For instance, 0.08% and 12.8% of R2L attacks are included in the training and testing datasets, respectively.

B. Methods

As mentioned above, our proposed network traffic detection is designed consisting of two levels. First, reliable rules are generated by using a decision tree method. Then, designing a predictive model is performed with utilizing the reliable rules.

Level 1: abnormal detection In this level, detection of abnormal network traffic behaviors is executed. Specifically, normal and abnormal behaviors are determined. To perform this, rules are generated with classification and regression tree (CART) [16], which uses information theoretic concepts to create a decision tree that captures complex patterns in data. We used four features (duration, protocol_type, service, and flag) to create a tree and to support a rapid decision. There are three attribute values in the protocol_type, seventy attributes in the service, and eleven attribute values in the flag. The service feature denotes specific network services (e.g. HTTP, FTP, etc.) on the destination. The flag indicates network connection status representing how each connection is instantiated and terminated. Total 13 connection states can be identified from network traffic [17].

CART builds a tree for predicting continuous dependent variables (regression) and categorical predictor variables (classification). It is broadly used due to its efficiency in dealing with multiple data types and missing values. CART expression forms explicit and transparent grammatical rules [18], [19]. Also, it uses an exhaustive search of all variables and split values to find optimal splitting values for each node. This splitting stops at the pure node containing fewer examples. Advantages of using CART are a) it does not require any distributional assumptions for dependent and independent variables, b) it deals with multiple types of numerical and categorical variables as inputs and outputs, c) it is not affected by outliers, and d) it efficiently handles high dimensional data. Among the generated trees with CART, only the trees that give high training accuracy are selected. Rules are extracted from the selected trees and tested with the testing dataset. With these rules, abnormality of network traffic is determined.

Level 2: attack type identification The abnormal monitoring depends on identifying network traffic patterns using transparent rules. In this level, we focus on identifying detailed information about abnormal behaviors by generating a predictive model. When generating a model, utilization of feature extraction from the input data is important because features may include informative knowledge representing hidden, but important patterns of the input data.

We used both fractal domain and statistical measurements to extract features. A sliding window (=20 data points) with 67% of overlapping is applied to examine the variation of network traffic flow. Fractal dimension (FD), called a non-linear dynamical method, is based on fractal theory concept. This method is applied broadly in many areas including biology, image segmentation, audio signal analysis, and medicine [20], [21], [22]. FD is a useful method for detecting rapid variations from data [21]. It also presents a self-similarity measurement of data. The self-similarity can be connected with "fractals" and the fractals within the network traffic display shape similarities in time scales. Since fractal features express the degree of self-similarity of the data, any unusual patterns can be detected. There are several algorithms to calculate the FD such as box-counting [23], Katz's [24], and Higuchi [25]. In this study, Higuchi FD is used as a fractal feature extraction method because it requires less computational power, guarantees accurate estimation, supports memory efficiency than other methods. To the best of our knowledge, Higuchi FD has not been used to extract features for determining exact attack types. The Higuchi method first re-generates original input data as a finite time series subset based on pre-defined window size (k=5). For the given the input data, new finite time-series is constructed as follows:

$$x(m), x(m+k), x(m+2k), ..., x\big(m+\left[\frac{N-m}{k}\right]\cdot k),$$
$$m = 1, 2, ..., k$$

where "[ ]" denotes the Gauss' notation, the largest integer in the neighborhood of the number, and both k and m indicate internal and initial time, respectively. The length of the curve x_k^m is defined as

$$L_m(k) =$$
$$\frac{1}{K}\left\{\left(\sum_{i=1}^{\frac{N-m}{k}} |x(m+ik) - x(m+(i-1)\cdot k|\right)\frac{N-1}{\left\lfloor\frac{N-m}{k}\right\rfloor\cdot k}\right\}$$

where $\frac{N-1}{\lfloor\frac{N-m}{k}\rfloor\cdot k}$ presents the normalization factor for the curve length and N is the total length of the signal. $< L(k) >$ defines the length of the curve for the time series k and $< L_m(k) >$ denotes the average value over k.

Statistical features including mean, median, and standard deviation are computed to find meaningful information to determine attack types. After collecting features, a machine learning algorithm, Support Vector Machine (SVM), is used to generate a predictive model. SVM [26], [27] is a supervised machine learning algorithm. It constructs an optimal hyperplane that separates a set of positive examples from a set of negative samples with maximum margin [28]. Due to its effectiveness, SVM is broadly used in pattern recognition, regression-based statistical learning theory, and structural risk minimization. In addition, it has an ability of handling large feature space. Therefore, we utilized SVM to design the predictive model. Neural Network (NN) is also used to determine the effectiveness of our proposed (SVM-based) predictive model. Since the neural network is one of the broadly used techniques for network anomaly detection [29], a NN-based predictive model is designed and its performance is compared with the SVM-based predictive model.

IV. Experimental Results

As described above, the training dataset consists of 125,973 network traffic records. There are 22,544 network traffics in the testing dataset. Figure 2 represents that one of the raw features ('duration') used in this study. It explains how the network traffic data is complex in considering of 'normal' and 'abnormal' activities in the training dataset. The duration indicates length (seconds) of network connection. It represents how long each network session was being established between source and destination.

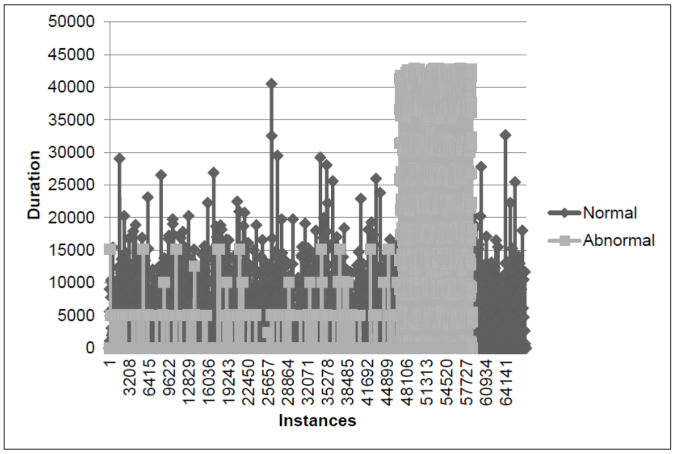
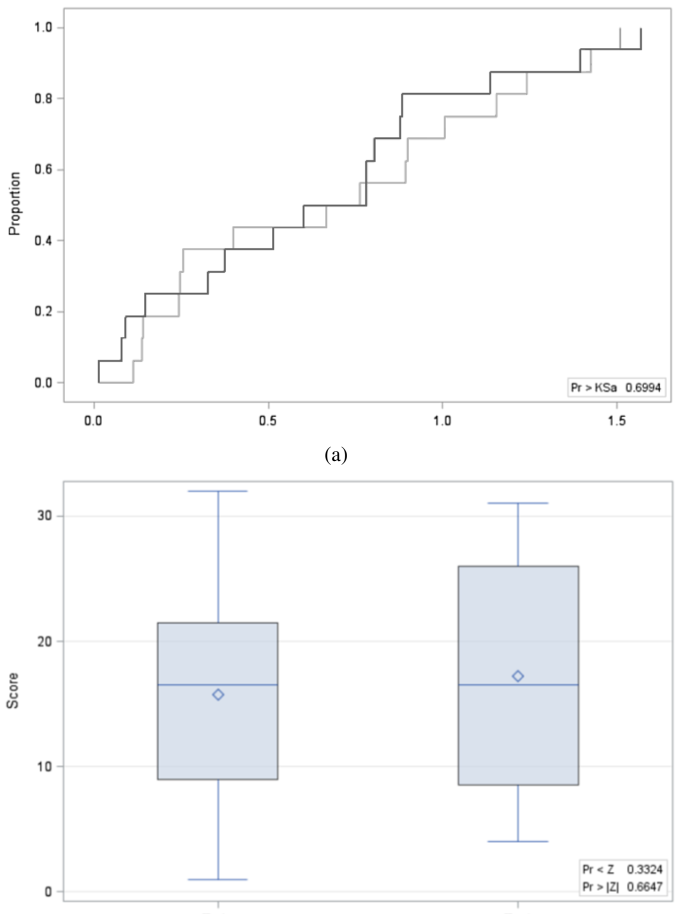


Fig. 2. A raw feature ('duration') distribution comparison of the training dataset.

Non-parametric t-test (Kolmogorov-Smirnov (KS) and Wilcoxon rank-sum test [30]) are performed to examine statistical distribution between the two datasets (training and testing). The hypothesis of the test is that the distribution of the two datasets is the same. Figure 3 presents the Kolmogorov-Smirnov test and the Wilcoxon rank-sum test results of the feature ('duration'). Based on the test results, we can conclude that our hypothesis can not be rejected, indicating that the distribution of the two datasets is similar.



(a)

(b)

Fig. 3. Distribution test results of the feature ('duration') by (a) Kolmogorov-Smirnov test (dark and light gray indicate the training and testing datasets, respectively) and (b) Wilcoxon rank-sum test.

Figure 4 represents how the raw feature ('protocol type') is distributed in the training and testing datasets. TCP and UDP are the most commonly used transport layer protocols in network communication. Since TCP provides a reliable connection, it is used by majority of Internet applications including WWW, FTP, and e-mail. However, UDP is a connectionless communication method that allows the fastest and most simple way of transmitting data to the receiver. Due to this reason, UDP is widely used for online video communication or VoIP applications. Since TCP and UDP is commonly used in the Internet applications, attackers often use TCP- or UDP-based attack techniques. Attackers also use ICMP which is internet layer protocol. ICMP is a message control (or error-reporting) protocol that is used for controlling the Internet between a host server and a gateway. Therefore, ICMP is not directly used or apparent to application users. However, attackers use ICMP to attack networks by sending bad ICMP packets or overloading the targeted network's bandwidth (often called ICMP Smurf or DDos Attack).

As mentioned previously, CART is considered for rule extraction. All generated rules with the training dataset are tested with the testing dataset. As a result, overall testing accuracy was 80.73%. Only the rules with high accuracy and

TABLE II. SOME OF THE GENERATED RULES MAINTAINING HIGH ACCURACY ARE PRESENTED.

Rules	Outcome (Accuracy - records)
$If(PT \neq \gamma \& PT \neq \mu \& (FLG = \epsilon \mid FLG = \varepsilon \mid FLG = \theta \mid FLG = \vartheta) \& SRV \neq \alpha \& SRV \neq \beta)$	Abnormal (99.7% - 6548/6556)
$If(PT = \gamma \& FLG \neq \delta \& FLG \neq \varepsilon)$	Normal (88.0% - 6669/7573)
$If(PT \neq \gamma \& PT \neq \mu \, \& FLG = \rho \& \, duration < 8.5)$	Abnormal (72% - 89/124)

where SRV:service,FLG:flag, PT:protocol type,α:SMTP, β:x11,γ:HTTP,δ:S1,ε:RSTR,ϵ:OTH, θ:SH,ϑ:REJ,μ:IRC,ρ:RSTO.

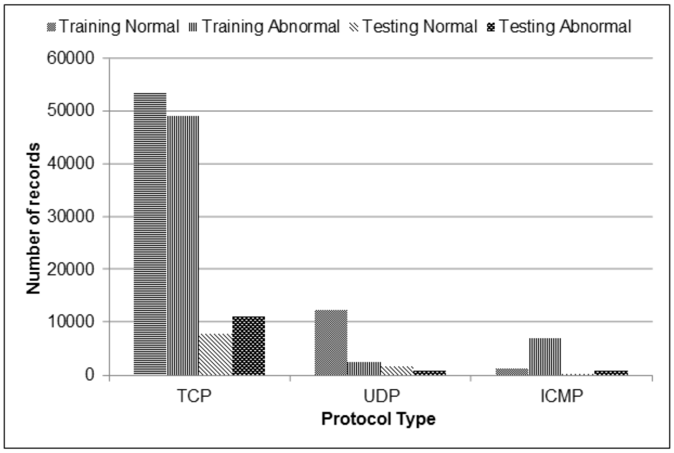
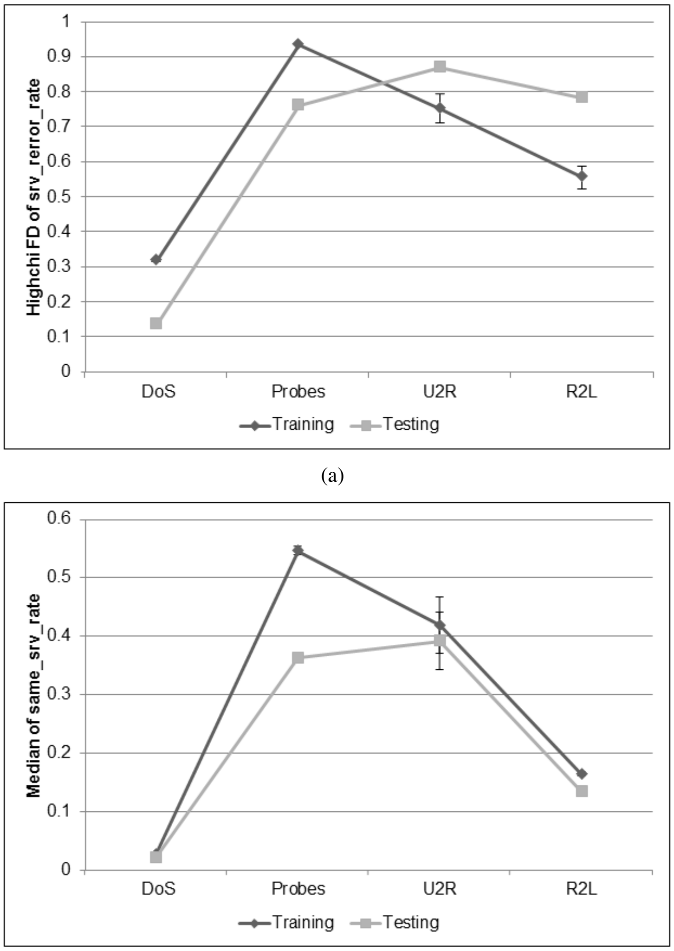


Fig. 4. Distribution of network traffic records depending on the protocol types in the training and testing datasets.

large number of examples are used to form a rule base. Some of the rules are listed in Table II.

After determining the abnormality of network traffic, a predictive model is generated to identify specific attack types. As mentioned in Section III-A, the abnormal attacks (i.e. DoS, U2R, R2L, and Probes) are not equally distributed in the training and testing datasets. Although we know that U2R and R2L attacks are more harmful to networks [14], it is difficult to detect these attacks because a small fraction of such attacks exists in the datasets. Total thirty-seven input features are used to design the predictive model. After applying Higuchi FD and statistical measures to the input features, 148 features are extracted. To enhance the performance of detecting exact attack types by excluding less relevant information, statistical significance of the extracted features is measured with ANOVA. From this statistical significance testing, we can identify how well the features are statistically valuable for determining exact attack types. From the ANOVA test, we found that forty-nine features are statistically significant ($p < 0.05$). These significant features are used to generate a predictive model with SVM.

Since the predictive model is designed with the training dataset, it is important to see how the proposed predictive model is efficient for detecting network anomalies from the testing dataset. To see the effectiveness of the model, we focused on understanding the difference between the training and testing datasets in consideration of the extracted features. Figure 5(a) and 5(b) show Higuchi fractal dimension feature of the network feature (srv_rerror_rate) and the statistical measurement feature (i.e. median) of the network feature (same_srv_rate), correspondingly. The figures indicate how the extracted features from the training and testing datasets are similar. More specifically, Figure 5(a) presents a Higuchi FD pattern of the variable “srv_rerror_rate” between the training and testing datasets. The variable “srv_rerror_rate” indicates connections that have REJ errors ins services. Although there are differences among attacks, the extracted features maintain similar trends in the training and testing datasets. It addresses that detecting abnormal network traffics in the testing dataset can be performed precisely. Figure 5(b) represents mean and the standard error of the mean (SEM) of the feature (same_srv_rate). This feature denotes connections that have the same services on the network. This figure indicates that all attack types except Probes retain similar characteristics in the training and testing datasets. It explains us that our predictive model will maintain high accuracy when identifying abnormal behaviors in network traffic. In addition, we examine the input feature to see the difference between before and after applying Higuchi FD. For this, their mean are compared.



(a)

(b)

Fig. 5. Comparative results (mean ± SEM) of the features from (a) Higuchi fractal dimension and (b) statistical measurement.

Figure 6 shows the result of the feature (srv_rerror_rate) for the DoS attack between the raw feature and its Higuchi feature in the training and testing datasets. It explains that since the Higuchi FD feature well maintains its characteristics compared to the raw feature, the predictive model designed with the extracted Higuchi FD features can successfully distinguish new incoming exact types of attacks.

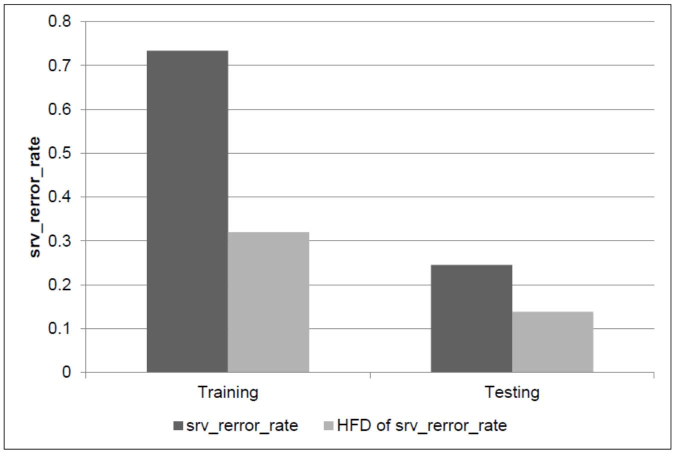


Fig. 6. An example result about the mean of the feature (srv_rerror_rate) for the DoS attack, in consideration of the training and testing datasets.

TABLE III. A PERFORMANCE COMPARISON BETWEEN SVM-BASED AND NN-BASED PREDICTIVE MODELS WITH THE TESTING DATASET.

	Accuracy	ROC area
SVM	77.10%	0.879
NN	51.90%	0.643

Our proposed SVM-based predictive model is compared with a NN-based predictive model. The comparison is performed with the testing dataset focusing on how effectively each model can detect exact attack types among DoS, U2R, R2L, and Probes. Table III shows the comparative result between SVM-based and NN-based predictive models. From the comparison, we found that SVM-based predictive model was outperformed in terms of accuracy. Although the overall accuracy of SVM-based predictive model was 77.1%, we found that true positive of DoS attack and Probes were 99% and 100%, respectively. When generating the SVM-based predictive model with using raw features, we found that the accuracy was dropped to 69% with the area of ROC as 0.744. This explains that our proposed predictive model is able to maintain a higher chance of identifying exact attack types.

V. Discussion and Conclusion

Understanding network traffic is important for securing our computing infrastructures. However, it is difficult to differentiate normal network traffic from abnormal network behaviors. This paper contributes to designing a two-level abnormal network traffic monitoring method. Since the most critical advantage of adapting rule-based method is that transparent rules can provide reasons behind the predictions, we generated rules with CART to differentiate normal and abnormal. We also proposed a predictive model to identify exact attack types of abnormal network traffic. Instead of using the raw feature for the predictive model, Higuchi fractal dimension and statistical measures are applied to extract significant features. Prior to applying them directly to design the predictive model, all extracted features were analyzed to determine their statistical significances. As explained in Section IV, we found that the performance accuracy of detecting network anomaly was high when using the extracted features.

For understanding the effectiveness of our proposed predictive model, it is important to perform a comparative study with other known techniques. In this study, we considered performing a comparison between our proposed SVN-based predictive model and a broadly used NN-based predictive model. From this comparison, we identified that our proposed model shows a better performance than the NN-based model. Although the overall accuracy of the proposed model was 77.1%, true positive of DoS and Probes attacks showed over 90% of accuracy. Since U2R and R2L attacks have fewer numbers of records in the datasets, it is difficult to precisely detect U2R or R2L attacks. Due to this reason, the overall performance accuracy of detecting abnormal behaviors in network traffic was lower than expected. This accuracy can easily be increased if we add more network traffic data related to U2R and R2L attacks. For future works, we plan to find more significant features to detect network attacks with understanding details about the attacks.

Acknowledgment

This work was fully supported by US Army Research Office (ARO) grant W911NF13110143, and partially supported by ARO grant W911NF1210060.

References

[1] A. Kind, M. Stoecklin, and X. Dimitropoulos, "Histogram-based traffic anomaly detection," *Network and Service Management, IEEE Transactions on*, vol. 6, no. 2, pp. 110–121, June 2009.

[2] A. Patcha and J. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Computer Networks*, vol. 51, no. 12, pp. 3448–3470, Aug. 2007.

[3] C.-M. Cheng, H. T. Kung, and K.-S. Tan, "Use of spectral analysis in defense against dos attacks." in *GLOBECOM*. IEEE, 2002, pp. 2143–2148.

[4] H. Wang, D. Zhang, and K. Shin, "Detecting syn flooding attacks," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, June 2002, pp. 1530–1539.

[5] M. Thottan and C. Ji, "Anomaly detection in ip networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 8, pp. 2191–2204, Aug 2003.

[6] J. Cannady, "Artificial neural networks for misuse detection," in *National Information Systems Security Conference*, 1998, pp. 443–456.

[7] R. P. Lippmann and R. K. Cunningham, "Improving intrusion detection performance using keyword selection and neural networks," *Computer Networks*, vol. 34, no. 4, pp. 597 – 603, 2000, recent Advances in Intrusion Detection Systems.

[8] S. T. Sarasamma, Q. A. Zhu, and J. Huff, "Hierarchical kohonenen net for anomaly detection in network security." *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 35, no. 2, pp. 302–312, 2005.

[9] S.-J. Han, K.-J. Kim, and S.-B. Cho, "Evolutionary learning programs behavior in neural networks for anomaly detection," in *Neural Information Processing*, ser. Lecture Notes in Computer Science, N. Pal, N. Kasabov, R. Mudi, S. Pal, and S. Parui, Eds. Springer Berlin Heidelberg, 2004, vol. 3316, pp. 236–241.

[10] T. Shon and J. Moon, “A hybrid machine learning approach to network anomaly detection,” *Inf. Sci.*, vol. 177, no. 18, pp. 3799–3821, Sep. 2007.

[11] R. Jain and N. Abouzakhar, “A comparative study of hidden markov model and support vector machine in anomaly intrusion detection,” *Journal of Internet Technology and Secured Transactions (JITST)*, vol. 2, no. 1/2/3/4, pp. 176–184, 2013.

[12] A. Hofmann and B. Sick, “Evolutionary optimization of radial basis function networks for intrusion detection,” in *Neural Networks, 2003. Proceedings of the International Joint Conference on*, vol. 1, July 2003, pp. 415–420 vol.1.

[13] A. Lakhina, M. Crovella, and C. Diot, “Diagnosing network-wide traffic anomalies,” in *In ACM SIGCOMM*, 2004, pp. 219–230.

[14] “NSL-KDD dataset,” http://nsl.cs.unb.ca/NSL-KDD/, 2014, [Online; accessed 2-April-2014].

[15] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the kdd cup 99 data set,” in *Proceedings of the Second IEEE International Conference on Computational Intelligence for Security and Defense Applications*, ser. CISDA’09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 53–58.

[16] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. New York: Chapman & Hall, 1984.

[17] V. Paxson, “Bro: A system for detecting network intruders in real-time,” *Comput. Netw.*, vol. 31, no. 23-24, pp. 2435–2463, Dec. 1999.

[18] W.-Y. Loh and N. Vanichsetakul, “Tree-structured classification via generalized discriminant analysis,” *Journal of the American Statistical Association*, vol. 83, no. 403, pp. pp. 715–725, 1988.

[19] C. Y. Fu, “Combining loglinear model with classification and regression tree (cart): an application to birth data,” *Computational Statistics & Data Analysis*, vol. 45, no. 4, pp. 865 – 874, 2004.

[20] U. R. Acharya, P. K. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, “Heart rate variability: a review.” *Med. Biol. Engineering and Computing*, vol. 44, no. 12, pp. 1031–1051, 2006.

[21] M. Liu, Y. He, Q. Meng, and Z. Wang, “Research on anomaly detection of network traffic based on fractal technology and vector quantization,” in *Education Technology and Computer Science (ETCS), 2010 Second International Workshop on*, vol. 2, March 2010, pp. 428–431.

[22] C. Gómez, A. Mediavilla, R. Hornero, D. Abásolo, and A. Fernández, “Use of the higuchi’s fractal dimension for the analysis of meg recordings from alzheimer’s disease patients.” *Med Eng Phys*, vol. 31, no. 3, pp. 306–13, 2009.

[23] J. Li, Q. Du, and C. Sun, “An improved box-counting method for image fractal dimension estimation,” *Pattern Recognition*, vol. 42, no. 11, pp. 2460 – 2469, 2009.

[24] M. J. Katz, “Fractals and the analysis of waveforms,” *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145 – 156, 1988.

[25] T. Higuchi, “Approach to an irregular time series on the basis of the fractal theory,” *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277 – 283, 1988.

[26] V. N. Vapnik, *Statistical Learning Theory*. Wiley-Interscience, 1998.

[27] S. Idicula-Thomas, A. J. Kulkarni, B. D. Kulkarni, V. K. Jayaraman, and P. V. Balaji, “A support vector machine-based method for predicting the propensity of a protein to be soluble or to form inclusion body on overexpression in escherichia coli.” *Bioinformatics*, vol. 22, no. 3, pp. 278–284, 2006.

[28] T. Joachims, “Text categorization with suport vector machines: Learning with many relevant features,” in *Proceedings of the 10th European Conference on Machine Learning*, ser. ECML ’98. London, UK, UK: Springer-Verlag, 1998, pp. 137–142.

[29] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009.

[30] E. L. Lehmann and H. D’Abrera, *Nonparametrics : statistical methods based on ranks*, ser. Holden-Day series in probability and statistics. San Francisco: Holden-Day New York Dusseldorf Johannesbourg, 2006, revised edition.