Improving Detection Accuracy for Imbalanced Network Intrusion Classification using Cluster-based Under-sampling with Random Forests

Md. Ochiuddin Miah, Sakib Shahriar Khan, Swakkhar Shatabda, and Dewan Md. Farid*
Department of Computer Science & Engineering, United International University
United City, Madani Avenue, Badda, Dhaka 1212, Bangladesh
Email: mmiah131145@bscse.uiu.ac.bd, skbshariar@gmail.com, swakkhar@cse.uiu.ac.bd, dewanfarid@cse.uiu.ac.bd

***Abstract*—Network intrusion classification in the imbalanced big data environment becomes a significant and important issue in information and communications technology (ICT) in this digital era. Presently, intrusion detection systems (IDSs) are commonly using tool to detect and prevent internal and external network attacks/ intrusions. IDSs are majorly bifurcated into host-based and network-based systems, and use pattern-matching techniques to detect intrusions that known as misuse-based intrusion detection system. Machine learning (ML) and data mining (DM) algorithms are widely using for classifying intrusions in IDS over the last few decades. One of the major challenges for building IDS employing machine learning and data mining algorithms is to improve the intrusion classification accuracy and also reducing the false-positive rate. In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. The proposed method is a multi-layer classification approach, which can process the highly imbalanced big data to correctly identify the minority/ rare class-intrusions. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack then the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. We have used KDD99 intrusion detection benchmark dataset for experimental analysis and tested the performance of proposed method with existing machine learning algorithms like: Artificial Neural Network (ANN), naïve Bayes (NB) classifier, Random Forest, and Bagging techniques.**

Keywords*—*Data Sampling; Ensemble Learning; Intrusion Classification; Imbalanced Big Data; Random Forest;

I. Introduction

In recent times with the encroachment of information and communication technologies, computer networks are facing threats like worms, viruses, adware, spyware, rootkits, trojan horses etc. [1]. These sorts of threats/ intrusions have turn into a major apprehension and need to be perceived before destructing the computer networks. These intrusions are affecting the productivity of computer networks, organizational electronic assets and other resources. Intrusion detection system is a network security technology to identify malicious activities/ intrusions from host or network by inspecting networks and systems [2]. Network-based IDS (NIDS) and host-based IDS (HIDS) are frequently used to identify internal or external malicious activities/ intrusions. NIDS scrutinizes and monitors incoming network traffic to detect network intrusions [3]. HIDS is devoted to a singular host or computer device to detect intrusions by scrutinizing inbound and outbound packets [4]. In detection approach, IDSs are categorized as misuse-based or anomaly-based IDS [5]. Misuse-based IDS is also acknowledged as pattern-based or signature-based IDS [6]. It is incapable of detect anonymous intrusions and detection rate is comparatively low because attackers endlessly try to attack with dissimilar signatures [7]. Otherwise, anomaly-based IDS is extensively used to detect known or unknown network intrusions [3]. Network intrusions are identified by ascertaining hidden intrusion patterns from large volume of network intrusion detection dataset [8]. Many machine learning (ML) and data mining (DM) algorithms such as ANN, SVM, decision tree (DT), genetic algorithm, naïve Bayesian, fuzzy logic etc. [9] have been employed to classify intrusions in IDS over the last few decades. Today's obtainable network intrusion detection datasets are pointedly imbalanced, multi-faceted, dynamic and composed of different inconsistent data that causes problem to handle entire data. Many machine learning algorithms have been miscarried to identify all sorts of attack with high detection accuracy because some of the attacks occurrence in the dataset are too low compared to other attacks. Today's accessible intrusion detection systems have low detection accuracy and high false positive rate for all sorts of known or unknown network attacks/ intrusions [10].

In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack than the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. The foremost purpose of this paper is to enhance the detection accuracy of low-frequency attacks in imbalanced network intrusion classification. We have built several classification models employing ANN [9], NB classifier [11], Random Forest [12], Bagging [13] and compare the performance of these existing learning methods with proposed method. The proposed method achieved 98% detection rate on average on the KDD99 benchmark dataset in compare with existing classifiers.

978-1-7281-3445-1/19/$31.00 ©2019 IEEE

The rest of the paper is organized as follows. Section II presents the related works. Section III presents an effective multi-layer hybrid method. Section IV provides experimental results based on KDD99 benchmark dataset. Finally, Section V presents conclusions and future works.

II. Related Work

Singh et al. [1] presented an Online Sequential Extreme Learning technique to reduce the slow learning drawback of neural network for intrusion detection. They reduce the volume of the training instances and time complexity by applying beta and alpha profiling respectively. They also discarded irrelevant features by using consistent feature selection techniques and ensemble of filtered. KDD 2009 dataset is used to measure the performance of proposed technique. The proposed IDS achieved an accuracy of 98.66% and 97.67% with false positive rate of 1.74% for binary class and multi class respectively. The proposed IDS outperformed other published techniques based on the accuracy and detection time. Farid et al. [14] introduced a new learning method using NB classifier and an ensemble approach for adaptive intrusion detection. The proposed method can classify a known or unknown instance by considering the votes of a series of classifiers. This method can update the weights of training instances based on the outcome of classification error rates, which reduces the false positive (FP) rates and improves the detection rates (DR). They applied existing data mining algorithms and the proposed method by employing on the KDD99 dataset. The proposed method can abate the false positive rates and achieves high detection rates on different sorts of network intrusions. Farid et al. [15] introduced a new learning algorithm for anomaly-based IDS to detect known or unknown intrusions employing decision tree. It splits and adjusts the weights of each instance based on probabilities until all the sub-dataset belongs to the same class. The weights of every instance is same in conventional decision tree algorithms but in this method weights of every instance changes based on posterior probability. The proposed learning procedure achieved 98% detection rate on KDD99 dataset in comparison with other existing algorithms.

Li et al. [9] proposed a two-step hybrid intrusion detection method construct on k-NN technique and binary classification employing NSL-KDD dataset. Firstly, the proposed method uses the C4.5 algorithm to train binary classifier to identify uncertain classes. Secondly, the proposed method uses the k-NN algorithm to classify uncertain classes. By the amalgamation of these steps, the proposed technique outperforms baselines algorithms like Random Forest, naïve Bayes, backward propagation neural network and k-NN. The F1-scores of this proposed technique is greatly greater than these of baselines algorithms for U2R and R2L low-frequency attacks. Zarpelão et al. [16] presented an analysis on intrusion detection system research for IOT. They were experimented that the IoT research for IDS is still impending and their purposes are to recognize open issues, leading trends and the upcoming research possibilities. They were categorized the IDSs based on the IDS placement stratagem, validation strategy, detection scheme and security intimidation. They were discussed signature-based, specification-based, anomaly-based and hybrid detection mechanisms to develop IDSs for IoT. They also discussed about centralized IDS placement, distributed IDS placement and hybrid IDS placement techniques.

III. Learning Algorithms

A. Learning Scheme

We have used Random Forest as learning scheme in our proposed method. Random Forest is an ensemble learning method and used for classification and regression in supervised learning[17]. It has the capability of classifying big data with higher accuracy [18]. It constructs multiple decision trees depending on arbitrarily chosen group of attribute at training time and outputting the class that is picked from the votes of ensemble of trees as furthermost popular class [11]. There is a relationship between the accuracy and the number of trees, higher number of trees supposed to produce higher accuracy rate. One of the big problems of machine learning is overfitting, it occurs when model overfits to the training data that when new test instance comes to the model from outside of training data, it can not classify it correctly, it gives higher performance only if the training data is used as test data. Random Forest algorithm has solved this problem as it takes vote from the ensemble of trees, so it will not overfit the model. As it considers votes of every decision tree it has constructed, it is supposed to achieve much higher accuracy than a single decision tree [11].

B. Proposed Method

In the first layer of our proposed method, we built a model *M* using dataset *D* and Random Forest technique. When a new instance x_{new} comes in, model *M* predicts that weather it is an attack or not. Following that if x_{new} has been identified as normal we will leave it here but if it is predicted as an attack, we have divided the dataset *D* into two different datasets D_{attack} and D_{normal} respectively. We have worked only with the dataset D_{attack} and it includes 4 different category of attacks, $D_{attack} = D_{DOS} \cup D_{U2R} \cup D_{R2L} \cup D_{Probe}$. Among the 22 attacks in the KDD99 dataset, there is a huge imbalance data in the context of quantity. We had to do resampling to make it balance. In order to do that, in the second layer we have applied under sampling and over sampling techniques to the dataset D_{attack}. Cluster centroid technique was applied to do the under sampling, it partitions the instances *X* with a number of clusters $C_1, C_2, ..., Cn$ and takes only cluster centroids $Centroid(C_i)$ so that informative instances are not lost. After the proceeding, we have replaced the labels of the dataset D_{attack} with category of attacks. Then we have built another model M_{attack} using dataset D_{attack} and Random Forest. The category of attacks of x_{new} dataset has been predicted by M_{attack}. Now the domain of attacks has been shorted, so we have separated the data D_{new} from D_{attack} in the third layer. D_{new} contains instances *X* with only attacks of the predicted category of X_{new}. We have built our final model M_{new} with dataset D_{new} and Random Forest. It will predict the actual class or attack of x_{new}. The architecture of the proposed hybrid method is illustrated in Fig. 1 and algorithm is summarised in Algorithm 1. In Fig. 1, (a) the proposed method classify an incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack then the proposed method try to classify main attack types shown in Fig. 1 (b) and later classify the final attack/ intrusion shown in Fig. 1 (c).

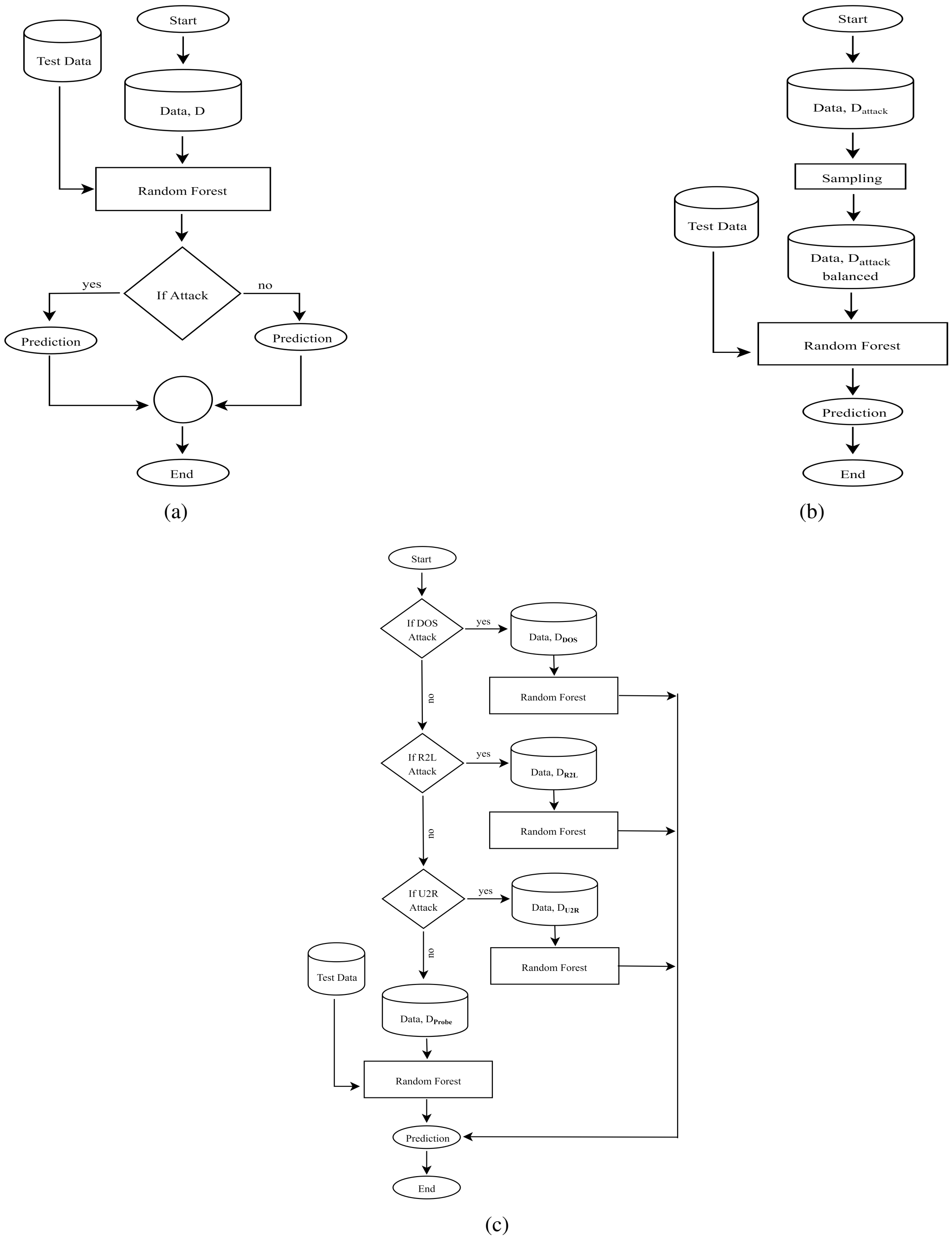

Fig. 1: Proposed method (a) normal or attack detection; (b) main attack types detection; (c) final attack/ intrusion detection.

IV. Experimental Analysis

The performance of classification algorithm is measured with several evaluation metrics in the literature. In this paper, we have used detection rates (DR) [17], false positives (FP) [17] as the standard evaluation metrics. Detection rate and false positive are widely used to estimate the performance of IDS that are shown in Eqs. 1 and 2.

$$DR = \frac{Total\_dected\_attacks}{Total\_attacks} \times 100 \qquad (1)$$

$$FP = \frac{Total\_misclassifed\_process}{Total\_normal\_process} \times 100 \qquad (2)$$

A. Intrusion Detection Dataset

The intrusion detection dataset availability is very limited because it contains information about host, topology and other classified information. The KDD99 standard intrusion detection dataset was used to constitute network intrusion detector to differentiate between normal connections and intrusions in

Algorithm 1 Proposed Hybrid Method

```
Input: Data $D$, and a learning scheme.
Output: Prediction, $C$
Method:
Replace Class labels of D with Normal and Attack;
Derive model, $M$ using $D$ and Random Forest;
Classify, $x_{new}$ using M:
$c = M(x_{new})$; // binary class prediction by $M$
if $c == normal$ then
    return c;
else
    $D_{attack} = \emptyset$;
    for each $x_i \in D$ do
        if $x_i \in attack$ then
            $D_{attack} = D_{attack} \cup x_i$;
        end if
    end for
    Derive model, $M_{attack}$ using $D_{attack}$ and Random Forest;
    Classify, $x_{new}$ using $M_{attack}$:
    $c = M_{attack}(x_{new})$; // prediction by $M_{attack}$
    C = $\{DOS, U2R, R2L, Probe\}$
    for each $k_i \in C$ do
        if $k_i == c$ then
            for each $x_i \in D_{attack}$ do
                if $x_i \in k_i$ then
                    $D_{new} = D_{new} \cup x_i$;
                end if
            end for
        end if
    end for
    Derive model, $M_{new}$ using $D_{new}$ and Random Forest;
    $c = M_{new}(x_{new})$; // prediction by $M_{new}$
    return c
end if
```

the 3rd International Knowledge Discovery and Data Mining Tools Competition [18], [17]. In this dataset, each data point denotes feature values of a class and every class is labeled either normal or attack and all instances are labeled with one of the five types which are mentioned bellow:

Normal: Normal connections are usually behaviors of the daily normal user and generated by visiting web pages, uploading files, downloading files etc [17].

Denial of Service: DoS attack makes the memory resources or power of a device too full or too busy to serve appropriate networking demands. Denial of service is accomplished by flooding the targeted resource with superfluous requests in an attempt to overload systems [11].

Remote to User: In R2L attack, isolated user achieves access of a native account by transferring packets to a device over the internet, it includes xlock, xnsnoop, guest, sendmail dictionary, phf etc [17].

User to Root: In U2R attack hacker takes access of the system with a normal user account and tries to find vulnerabilities for gaining root-user access. Common example of user to root attacks are perl, buffer-overflows, fd-format, load-module, ffb-config, and xterm [18].

Probing: In Probe attack the hacker scans a networking device or a machine to find known vulnerabilities or gather information [13].

KDD99 dataset have 22 dissimilar sorts of attacks that drop into four central categories and total of 41 input attributes and network connections are divided into three groups and they have discrete or continuous values [18]. The basic attributes of each TCP connections are in the first group, it includes the prototype, duration, bytes quantity of IP addresses of source or IP addresses of destination, some flags in TCP connections, and service. The features of network connections suggested by domain knowledge are in the second group and the third group has the statistical attributes that which are calculated by using a two-second time period. We have used training data to train the classifiers and evaluate the classifier models supplying test data from UCI machine learning repository. In testing data, there are some unknown attack examples those are not available in the training data. The number of training and testing examples are shown in Table I.

TABLE I: Number of instances in KDD99 dataset.

Attack Types	**Training**	**Test**
Normal	972781	60593
R2L	1126	7015
U2R	52	39
DoS	3883370	223298
Probing	41102	2377
Total	4898431	293322

B. Result

We have implemented the proposed method in Python 3.7 and used scientific Python development environment (Spyder) 3.3.1 (https://www.spyder-ide.org). We have experienced the performance of some popular machine learning algorithms like Random Forest, Bagging, naïve Bayes and ANN on KDD99 dataset. We have supplied testing data for evaluating our proposed method using detect rate and false positive rate. DR is very low for minority class instances which are mostly in R2L and U2R attack types. Some of the strong classifiers like Random Forest, ANN, Bagging, naïve Bayes failed to classify the minority class instances. For our proposed method, we have momentous improvement and DR for minority class instances in our proposed method is above 81%. The proposed method raises detection rates for R2L and U2R attacks type and achieved 87% and 81% detection rate where some popular algorithms able to achieve less than 30%. The improvement of false positive rate also noticeable for our proposed hybrid method. Comparison of results are tabulated in Table II and detection rates of Proposed Method, Random Forest, Bagging, ANN and naïve Bayes are shown in Figure 2.

V. Conclusions & Future Works

Recently, information security has evolved significantly and turn into a vital concern in information technology. IDSs have been employed to protect information from network or host based attacks by utilizing computational intelligence. However, today's accessible IDSs are pattern based also known as misuse

TABLE II: Performance comparison of different strong classifiers with proposed method.

Method	Normal	Probe	DoS	U2R	R2L
Proposed Method (DR %)	**0.99**	**0.99**	**0.99**	**0.81**	**0.87**
Proposed Method (FR %)	**0.03**	0.02	0.02	**0.01**	**0.01**
Random Forest (DR %)	0.99	0.99	0.99	0.12	0.20
Random Forest (FR %)	0.10	0.02	**0.01**	0.01	0.01
Bagging (DR %)	0.98	0.95	0.99	0.10	0.18
Bagging (FR %)	0.12	**0.01**	0.02	0.02	0.03
NB Classifier (DR %)	0.90	0.92	0.95	0.09	0.15
NB Classifier (FR %)	0.11	0.03	0.01	0.02	0.01
ANN (DR %)	0.99	0.99	0.99	0.16	0.23
ANN (FR %)	0.98	0.95	0.99	0.84	0.77

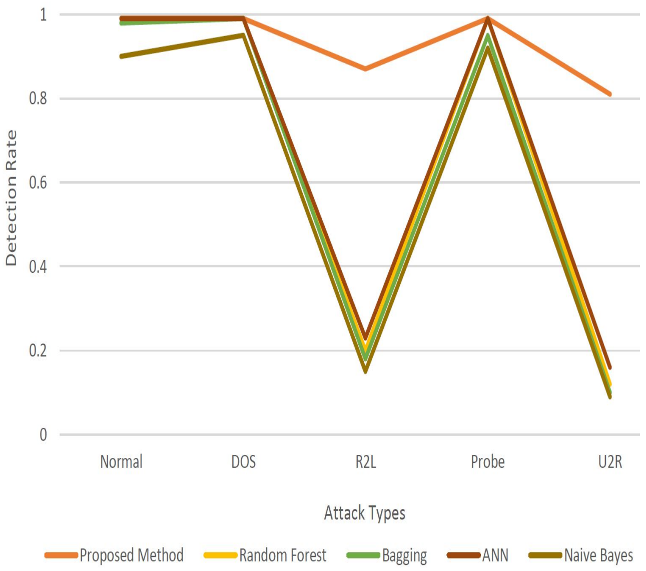

Fig. 2: Detection rates of proposed method, Random Forest, Bagging, ANN, and NB classifier.

based that are not able to detect unknown intrusions. In this paper, we have introduced a new method for improving detection rate to classify minority-class network attacks/ intrusions using cluster-based under-sampling with Random Forest classifier. The proposed method is a multi-layer classification approach, which can process the highly imbalanced big data to correctly identify known or unknown network intrusions. Initially, the proposed method classify a data point/ incoming data is attack/ intrusion or not (like normal behaviour), if it's an attack than the proposed method try to classify attack type and later sub-attack type. We have used cluster-based under-sampling technique to deal with class-imbalanced problem and popular ensemble classifier Random Forest for addressing overfitting problem. The central purpose of this paper is to enhance the detection accuracy of low-frequency attacks in imbalanced network intrusion detection classification. We compared the performance of proposed method with standard data mining algorithms. The experimental results on KDD99 benchmark dataset manifest that the proposed hybrid method raises detection rates and abates false positive rates. The future work will focus on boost the detection rates (DR) of low-frequency attacks in imbalanced dataset and apply this hybrid method into real world NIDS.

References

[1] R. Singh, H. Kumar, and R. K. Singla, "An intrusion detection system using network traffic profiling and online sequential extreme learning machine," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8609–8624, 2015.

[2] W. Wang, J. Liu, G. Pitsilis, and X. Zhang, "Abstracting massive data for lightweight intrusion detection in computer networks," *Information Sciences*, vol. 433, pp. 417–430, 2018.

[3] D. M. Farid, L. Zhang, C. M. Rahman, M. Hossain, and R. Strachan, "Hybrid decision tree and nave bayes classifiers for multi-class classification tasks," *Expert Systems with Applications*, vol. 41, pp. 1937–1946, March 2014.

[4] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *Computers & Security*, vol. 28, no. 1-2, pp. 18–28, 2009.

[5] M. S. Pervez and D. M. Farid, "Feature selection and intrusion classification in nsl kdd cup 99 dataset employing svms," *8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Dhaka, Bangladesh*, pp. 1–6, December 2014.

[6] D. M. Farid, N. H. Hoa, J. Darmont, N. Harbi, and M. Z. Rahman, "Scaling up detection rates and reducing false positives in intrusion detection using nbtree," *International Conference on Data Mining and Knowledge Engineering (ICDMKE), Rome, Italy*, pp. 186–190, April 2010.

[7] D. M. Farid and M. Z. Rahman, "Learning intrusion detection based on adaptive bayesian algorithm," *11th International Conference on Computer and Information Technology (ICCIT), Khulna, Bangladesh*, pp. 652–656, December 2008.

[8] D. M. Farid, M. Z. Rahman, and C. M. Rahman, "Chapter title: Mining complex network data for adaptive intrusion detection," in *Advances in Data Mining Knowledge Discovery and Applications*, September 2012, ch. 15, pp. 327–348.

[9] S. B. Y. H. Longjie Li, Yang Yu and X. Chen, "An effective two-step intrusion detection approach based on binary classification and k-nn," *IEEE Access*, vol. 6, pp. 12 060–12 073, 2018.

[10] D. M. Farid and M. Z. Rahman, "Anomaly detection model for network intrusion detection using conditional probabilities," *6th International Conference on Information Technology in Asia (CITA), July, 2009, Kuching, Sarawak, Malaysia*, pp. 104–110.

[11] D. M. Farid, L. Zhang, A. Hossain, C. M. Rahman, R. Strachan, G. Sexton, and K. Dahal, "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5895–5906, November 2013.

[12] D. M. Farid, M. A. Al-Mamun, B. Manderick, and A. Nowe, "An adaptive rule-based classifier for mining big biological data," *Expert Systems with Applications*, vol. 64, pp. 305–316, December 2016.

[13] D. M. Farid, A. Nowe, and B. Manderick, "Ensemble of trees for classifying high-dimensional imbalanced genomic data," in *Proceedings of SAI Intelligent Systems Conference*. Springer, September 2016, pp. 172–187.

[14] D. M. Farid, C. M. Rahman, and M. Z. Rahman, "Adaptive intrusion detection based on boosting and naïve bayesian classifier," *International Journal of Computer Applications*, vol. 24, no. 3, pp. 12–19, 2011.

[15] D. M. Farid, C. M. Rahman, N. Harbi, E. Bahri, and M. Z. Rahman, "Attacks classification in adaptive intrusion detection using decision tree," *International Conference on Computer Science (ICCS), Rio De Janeiro, Brazil*, pp. 86–90, March 2010.

[16] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in internet of things," *Journal of Network and Computer Applications*, vol. 84, pp. 25–37, 2017.

[17] D. Farid, J. Darmont, N. Harbi, H. H. Nguyen, and M. Z. Rahman, "Adaptive network intrusion detection learning: attribute selection and classification," *International Conference on Computer Systems Engineering (ICCSE), Bangkok, Thailand*, pp. 82–86, December 2009.

[18] D. M. Farid, N. Harbi, and M. Z. Rahman, "Combining naive bayes and decision tree for adaptive intrusion detection," *International Journal of Network Security & Its Applications (IJNSA)*, vol. 2, no. 2, pp. 12–25, April 2010.