Cost-Sensitive Access Control for Detecting Remote to Local (R2L) and User to Root (U2R) Attacks

Doaa Hassan

Computers and Systems Department, National Telecommunication Institute, Cairo, Egypt

Abstract**—** *Remote to local attack (r2l) has been widely known to be launched by an attacker to gain unauthorized access to a victim machine in the entire network. Similarly user to root attack (u2r) is usually launched for illegally obtaining the root's privileges when legally accessing a local machine. One approach for detecting both attacks is to formulate both problems as a binary classification problem by deciding whether to accept or reject access requests from remote sites to local user machine or by accepting or rejecting access as root attempts. However, the cost caused by incorrect decision due to accepting illegitimate access request in a form of the damage that it might lead to is more expensive than the opposite case resulting from rejecting a valid access request. Due to this, in this paper we handle both problems in cost sensitive learning framework. We investigate how various cost-sensitive machine learning methods can be used to produce various cost sensitive detection models for detecting illegitimate remote access and access as a root requests. Those models are optimized for a user-defined cost matrix. Empirical experiment shows that the produced cost sensitive detection models are effective in reducing the overall cost of illegal remote access and access as root detection.*

Keywords*— cost sensitive learning methods, r2L attack, u2r attack.*

I. Introduction

Defining intruders by distinguishing between the normal user behavior and attacker behavior has been one of the main objectives of many network intrusion detection systems (NIDS) [3], [4]. Effective NIDS should define a set of rules that forms its policy for classifying the records of network connections into either normal or anomalies based on the detected attack patterns. Therefore, monitoring and analyzing network traffic for detecting those patterns have been a great challenge to achieve by each deployed NIDS. Recently, there has been various datasets that include a collection of common attack patterns that have been investigated by research on NIDS. KDD-cup 99 dataset [5] is one those datasets that have been extensively used in research on NIDS. This dataset includes various anomaly patterns for four common types of attacks including:

- Denial of Service (DoS): occurs when an attacker tries to deny legitimate users access to a particular service or resource.
- Remote to Local (r2l): occurs when an attacker does not have an account on the victim machine, and tries to gain access by sending packets to a machine over a network in order to generate some vulnerability on that machine that allows him/ her to gain local access as a user of that machine.
- User to Root (u2r): occurs when normal system user illegally gains access to either root's or super user's privileges.
- Probe: occurs when an attacker scans a network in order to gather information or find known vulnerabilities that allows him /her to hack the entire network.

Recently there has been much research work on using data miming and machine learning techniques for detecting the anomaly network connections that have anomaly patterns for any of those types of attack [2], [6]–[9]. The basic concept of this research work was to present the detection of anomaly pattern for each type of attack as a binary classification problem that allows to decide whether the network connection is normal (i.e., free of attack patterns) or anomaly (i.e., has any of the attack patterns for any of the four attacks categories). Some of this research work has been concerned with detecting the anomaly pattern for attacks that allow illegal access to user machine; namely the r2l and u2r attacks [2].

The detection of anomaly pattern for both attacks can be represented as a binary classification problem by deciding whether to accept or reject access to victim's machine either from remote site in case of r2l or unauthorized access to super user's privilege in case of u2r. However, the cost caused by incorrect decision due to allowing illegal access (defined by the damage that it might lead to) is more expensive than the opposite case that results from rejecting a valid access request. Due to this, in this paper we handle such problem in cost sensitive modeling framework using machine learning techniques. We present various cost-sensitive machine learning techniques that can produce detection models for detecting illegitimate access attempts. Those models use different cost sensitive methods and are optimized for a user-defined cost matrix for representing the cost of each type of incorrect decision [15]. The empirical experiments show that the presented cost sensitive access control modeling techniques are different in reducing the overall cost of illegitimate access detection. Among those techniques, the one that uses Metacost method [1] has been found to be the most effective cost sensitive modeling technique in reducing the overall cost of detecting the illegitimate access attempts due to launching r2l attack. Moreover, the one that uses cost evaluation method [16] has been found to be the most effective cost-sensitive modeling technique in reducing the overall cost of detecting the illegitimate access attempts due to launching u2r attack.

The structure of this paper is organized as follows: In Section II, we provide an overview of cost sensitive learning methods that are used in this paper for building cost sensitive models that are able to detect illegal access attempts due to starting either r2l or u2r attacks. In Section III, we introduce the basic experiment by representing the experimental approach and settings, and the evaluation results. In Section IV, we discuss the related work. Finally we conclude the paper in Section V with a direction for future work.

II. Overview of Cost Sensitive Learning Methods

The accuracy of classifying a dataset is commonly used in machine learning field as a metric to evaluate the performance of classifiers. However, some types of misclassifications may affect badly than others. For example, rejecting authorized access to a system may be misleading while authorizing an illegitimate access may be more dangerous and cause very negative consequences. Therefore, using cost sensitive learning in such a scenario for evaluating classifiers performance is much more meaningful, where the cost of every type of misclassification is taken into account in order to avoid the misclassifications that lead to catastrophic situations. To this end, this paper investigates how different cost-sensitive machine learning methods can be used for constructing various cost sensitive detection models for identifying illegitimate remote access and access as root requests caused by launching r2l or u2r attacks receptively. Those models are optimized for a given user-defined cost matrix for representing the cost of each type of misclassifications [15]. Following, we will provide a brief overview of some of the cost sensitive learning methods that will be used in the basic experiment of this paper, we refer to [16] for more details about the first three methods and [1] for more details about the last one:

A. *Cost evaluation*:

In this technique, the cost of particular learning model on a given test set is calculated by just summing the relevant elements of the cost matrix for the models prediction for each test instance. Therefore, the costs are ignored during predictions and only taken into account when evaluating them. For example, given the cost matrix shown in Table I, the total cost of the learning model is equal to (C_{TP} + C_{FP} + C_{FN} + C_{TN})*N, where:

- N refers to the number of instances in the testing set.
- C_{TP} is the cost of true positive due to predicting the authorized access as legitimate.
- C_{FP} is the cost of false positive due to predicting the unauthorized access as legitimate
- C_{FN} is the cost of false negative due to predicting the authorized access as illegitimate.
- C_{TN} is the cost of false negative due to predicting the unauthorized access as illegitimate.

B. *Cost sensitive classification*

A classifier in this technique is built without taking costs into consideration, while it can be used to make predictions that are sensitive to the cost matrix (i.e., costs are ignored at training time but used at prediction time.). This is achieved by adopting the classifier to compute the probability associated with each prediction. Therefore, when the classifier assigns the classes a, and b to a test instance with probabilities p_a and p_b (assuming the cost matrix shown in Table I), then if the classifier predicts a, the expected cost of the prediction is obtained by multiplying the first column of the matrix, [C_{TP}, C_{FP}], by the probability vector, [p_a, 1-p_a], where the sum of p_a and p_b is 1. Therefore, choosing the prediction with the lowest expected cost is equivalent to choosing the one with the greatest probability.

C. Cost sensitive learning

It is an opposite technique to cost sensitive classification, where the cost matrix is taken into account during the training process and costs are ignored at the prediction time. This is achieved by varying the proportion of instances in the training set in order to enforce the learning scheme to minimize the number of costly errors by making a decision that is biased toward avoiding errors on the negative instances.

D. Metacost

This technique consists of two phases: bagging for relabeling each training example with the cost, and retraining the classifier with the cost. For the first phase, a set of samples is generated with replacement from the training set and the class of each instance is estimated by taking the average of votes over all the trained classifiers then each training example is relabeled with the estimated optimal class. For the second phase, the classifier is retrained to the relabeled training set.

TABLE I
AN EXAMPLE OF COST MATRIX.

Cost Matrix	Predicted Class		
Actual Class	$C(a\mid b)$	a	b
	a	C_{TP}	C_{FN}
	b	C_{FP}	C_{TN}

III. BASIC EXPERIMENT

A. *Experimental Approach and Settings*

Our experimental approach starts with preprocessing the KDD-cup 99 dataset by fragmenting it into 4 subsets, each subset contains records of normal and a specific attack category. In this paper, we are concerned with only two of the four subsets: The first one contains 97278 records of normal connections and 1126 records of r2l attack connections, while the second one contains 97278 records of normal connections and 52 records of u2r attack connections. Each dataset subset is split into training and testing subsets with a ratio 2:1. Since both datasets are large and imbalanced, the training subset of each dataset is undersampled [17] by generating a balanced smaller subset of the training subset with records of r2l attack of a ratio 49.77 % in case of r2l dataset subset and with records of u2r attack of a ratio of 49.48% in case of u2r dataset. The testing subsets for both r2l and u2r datasets are kept without undersampling in order to reflect the real distributions of r2l and u2r attacks as in the original r2l and u2r dataset subsets.

Next, four classifiers from different categories are generated using Weka including decision tree (DT), random forest (RF), support vector machine (SVM) and logistic regression (LR) [21]. More precisely J48, RandomForest, SMO and Logistic implementations respectively in Weka. The choice of DT and SVM is due to their robustness to the curse of dimensionality problem when applying them on high dimensional data such as KDD-cup 99 dataset [22], while the choice of LR and RF is due to their previous usage for formulating some problems in cost sensitive framework [23]. Those four classifiers are trained on the experimental training balanced sub-samples of r2l and u2r dataset subsets in order to build four cost sensitive learning models for each one based on the four cost sensitive methods that were presented in section II. Therefore, a total of 16 cost sensitive learning models will be generated in the end for each training subsample. Those models are tested on the two testing subsets of r2l and u2r attacks for identifying the illegal access attempts in terms of the total misclassification cost, where lower cost refers to better performance. A cost matrix similar to the one shown in Table I has been used, where a and b classes are replaced with normal and r2l classes in case of testing on r2l testing subset and with normal and u2r classes in case of testing on u2r testing subset. The cost of right classification in this matrix represented by either C_{TP} or C_{TN} is set with zero value since the identification of any type of both always does not lead to any damage or misleading results.

On the other hand, the cost of C_{FP} and C_{FN} is set with values of a ratio 100:1 since the cost of false positive due to predicting the unauthorized access as legitimate is more dangerous and catastrophic than the inverse case. Moreover, the number of r2l records in r2l testing set is minor in comparison to the number of normal records. Similarly, the number of u2r records in u2r testing set is minor in comparison to the number of normal records.

We have run our approach on a windows laptop machine with 2.6 GHZ processor Intel core (TM)i5 and 4 G Memory Rams. We have used Weka [18], a free open source software data mining tool for generating different classifiers from different categories and train them on the two experimental data set subsets. Weka also allows including the cost matrix in classier evaluation. Moreover, it allows solving the imbalanced data set problem using Weka SpreadSubsample filter [19], [20] for creating a balanced subsample of the training data set.

B. **Evaluation Results**

Tables II and III show the performance of the four experimental classifiers in terms of misclassification cost with each of the cost sensitive methods presented in Section II, for r2l and u2r detection respectively. Best performance for each classifier is represented by the lowest misclassification cost and highlighted in bold. Similarly the best result for each cost sensitive method that leads to the lowest misclassification cost is also highlighted in bold. Figures 1, 2 show the misclassification cost for each classifier with each cost sensitive method in case of r2l and u2r detection respectively. From the results, we can conclude the following:

- DT achieves the lowest misclassification cost of 293 among all classifiers on the testing set of r2l and hence the best performance, using the cost evaluation method.
- RF achieves the lowest misclassification cost of 189 among all classifiers on the testing set of u2r and hence the best performance using the cost evaluation method.
- In general, there is no optimum cost sensitive method that can lead to the lowest misclassification cost with all experimental classifiers. For instance, we found that using the

Metacost method with SVM achieves the lowest misclassification cost of 742 among all other cost sensitive methods used for r2l detection, while it does not achieve the best with DT, since using the cost evaluation method with DT leads to the best performance of 293.

TABLE II
THE PERFORMANCE OF CLASSIFIERS FOR DETECTING R2L ATTACK USING FOUR COST SENSITIVE LEARNING METHODS.

Method	**Classifiers**				
	DT	**RF**	**SVM**	**LR**	**Average**
Cost evaluation	**293**	4605	5770	4991	3914.75
Cost sensitive classification	2559	**1294**	**957**	27621	8107.75
Cost sensitive learning	8598	4577	5770	5129	6018.5
Metacost	**669**	2486	**742**	**909**	**1201.5**
Average	**3029.75**	3240.5	3309.75	9662.5	

TABLE III
THE PERFORMANCE OF CLASSIFIERS FOR FOR DETECTING U2R ATTACK USING FOUR COST SENSITIVE LEARNING METHODS.

Method	**Classifiers**				
	DT	**RF**	**SVM**	**LR**	**Average**
Cost evaluation	**1212**	**189**	**858**	762	**755.25**
Cost sensitive classification	32436	32436	3954	**685**	17377.75
Cost sensitive learning	**1212**	18273	**858**	1273	5404
Metacost	2428	20048	1767	1073	6329
Average	**948.25**	17736.5	1859.25	948.25	

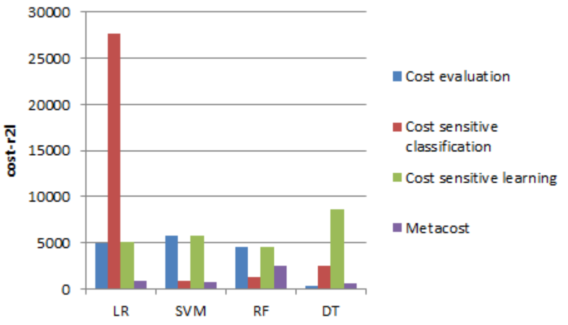


Fig. 1. The misclassification cost of classifiers for r2l testing dataset.

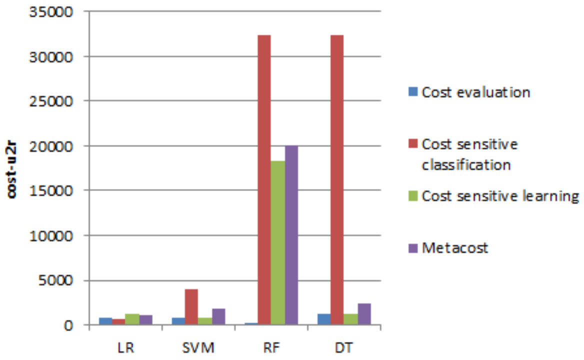


Fig. 2. The misclassification cost of classifiers for u2r testing dataset.

- We have also found that the misclassification cost depends on many factors such as the cost set for FP and FN in the cost matrix and the obtained number of false positives and negatives when detecting either r2l or u2r attacks using any of the four cost sensitive methods as shown in Table IV and Table V. Due to this, we have found that using the Metacost method on the testing set of r2l is preferred over using other cost sensitive methods as it achieves the lowest misclassification cost of 1201.5 on the average and hence the best performance. Clearly, this is because using Metacost results in a reduced number of FP of 0.25 on the average which is 100 times expensive than the cost of FN which is 1152 on the average in this case (assuming we have fixed the cost of FP and FN in the cost matrix during the experiment with 100:1 ratio). Similarly, we have found that using the cost evaluation method on the testing set of u2r is preferred over using other cost sensitive methods as it achieves the lowest misclassification cost of 755.25 on the average and hence the best performance. This is because using cost evaluation results in a reduced number of FP and FN of 0.25 and 730.25 respectively on the average.
- DT outperforms the performance of other classifiers on the average as it achieves the lowest misclassification cost of 3029.75 on the testing set of r2l. Clearly, this is because the average number of obtained FP and FN with DT in this case is 0 and 3029.75 respectively. Similarly, LR outperforms the performance of other classifiers on the average as it achieves the lowest misclassification cost of 948.25 on the testing set of u2r. Clearly, this is also because the average number of obtained FP and FN with LR in this case is 0 and 948.25 respectively.

IV.RELATED WORK

P. Gifty Jeya et el. [2] presented a fisher linear discriminant analysis that was carried out on KDD99 dataset. The analysis relies on applying correlation based feature selection that selects the subset of features that is involved in the classification of the attack categories of KDD-cup 99 dataset. Their proposed approach used the accuracy as a measure for evaluation of classification performance. However, though there approach was approved for improving the classification accuracy for R2L and U2R attacks categories, it did not consider the cost of misclassification as a metric for evaluating the classifier performance.

W. Lee et el. [13] studied the problem of constructing cost sensitive intrusion detection models by investigating the major cost factors associated with an IDS. This includes the development cost, operational cost, the cost of damage effect due to successful intrusions, and the cost of manual and automated response to intrusions. They defined cost models that formulate the total expected cost of IDS using cost-sensitive machine learning techniques that uses low cost features for generating detection rules. Those techniques produce detection models that are evaluated under user-defined cost metrics. Their attack taxonomy categorizes intrusions that occur in the DARPA Intrusion detection evaluation dataset [14].

Mitrokotsa et el. [10] examined how cost-sensitive classification methods can be used in intrusion Detection systems by conducting their experiments on KDD-cup 99 dataset. They performed cost evaluation for four different classifiers, where the performance of each classifier is evaluated in terms of the expected cost under different cost matrices. Next, they examined how the measured cost changes when the relative cost for the misclassification of each attack category in KDD99 versus normal connection increases. This was achieved by investigating the change in false alarm and detection rates due to varying the relative cost of false alarms and false negatives. However, unlike our work, they did not focus on the measuring misclassification cost of illegal access attempts caused by either r2l or u2r attacks.

The most closed work to ours is the one presented by Y-W. Seo and K. Sycara [11]. They examined how two cost-sensitive classification methods, namely costing [12] and Metacost [1] can be used for measuring the cost of illegitimate access attempts of unauthorized insiders to confidential content. They considered the cost of false positive due to accepting an illegitimate access request is more expensive than that of false negative due to rejecting a valid access request. Since the former represents a critical security problem that illegally reveals confidential information.

V. Conclusions

In this paper, we have shown how to build cost sensitive learning models for detecting illegitimate access attempts that take the form of either r2l or u2r attack. We have presented an empirical evaluation of various methods used for building those models. We have evaluated the performance of four experimental classifiers in terms of the total misclassification cost which is the summation of the misclassification costs resulting from false positives due to approving unauthorized access attempt and the false negatives due to preventing an authorized access attempt for each test example in the entire test set. Our results show that generally there is no one optimum method that can be used for building cost sensitive model that achieves the lowest misclassification cost with all the experimental classifiers. However, cost sensitive modeling for detecting r2l attack using Metacost method has been found to outperform other cost sensitive models developed using other methods on the average by achieving the lowest misclassification cost. Moreover, cost sensitive modeling for detecting u2r attack using cost evaluation method has been found to outperform other cost sensitive models developed using other methods on the average by achieving the lowest misclassification cost

As a future work, we are planning to provide similar empirical evaluation for cost sensitive models developed using various cost sensitive machine learning methods, for detecting illegitimate access attempts recorded by the logs of firewall, intrusion detection system or proxy server that are deployed on real network.

References

[1] Pedro Domingos. Metacost: A general method for making classifiers costsensitive. In proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining (KDD99), pages 155164, 1999.

[2] P. G. Jeya, M. Ravichandran and C. S. Ravichandran. Efficient Classifier for R2L and U2R Attacks. International Journal of Computer Applications, Volume 45, No.21, May 2012.

[3] Monowar H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita. Network Anomaly Detection: Methods, Systems and Tools. IEEE COMMUNICATIONS SURVEYS & TUTORIALS, , VOL. 16, NO. 1, FIRST QUARTER 2014.

[4] S.Vijayarani and M. Sylviaa Intrusion Detection System- A study. International Journal of Security, Privacy and Trust Management (IJSPTM), Vol 4, No 1, February 2015.

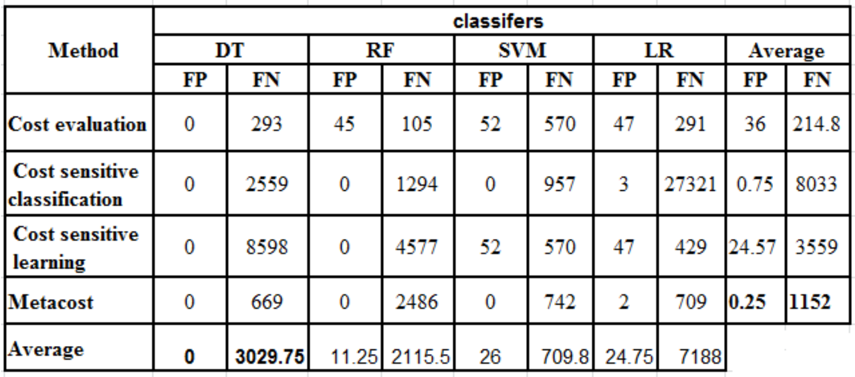
[5] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani. A Detailed Analysis of the KDD CUP 99 Data Set. In Proceedings of the IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009), 2009.

[6] Paul Dokas, Levent Ertoz, Vipin Kumar, Aleksandar Lazarevic, Jaideep Srivastava and Pang-Ning Tan. Data Mining for Network Intrusion Detection. In Proceedings of 2002 NSF Wrokshop on Data Mining, pp. 21.30, 2002.

[7] Ch.Ambedkar and V. Kishore Babu. Detection of Probe Attacks Using Machine Learning Techniques. International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), Volume2, Issue 3 , PP. 25-29, March 2015.

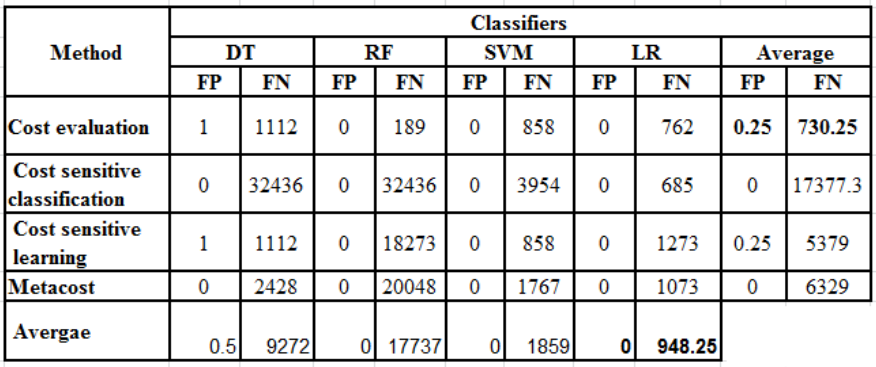
[8] H. Waguih. A Data Mining Approach for the Detection of Denial of Service Attack. IAES International Journal of Artificial Intelligence (IJAI), Vol. 2, No. 2, June 2013, pp. 99 106.

TABLE IV
THE TOTAL NUMBER OF FALSE POSITIVES AND FALSE NEGATIVES OBTAINED BY RUNNING EACH CLASSIFIER ON R2L TESTING SUBSET USING THE FOUR COST SENSITIVE METHODS



Method	classifers									
	DT		RF		SVM		LR		Average	
	FP	FN	FP	FN	FP	FN	FP	FN	FP	FN
Cost evaluation	0	293	45	105	52	570	47	291	36	214.8
Cost sensitive classification	0	2559	0	1294	0	957	3	27321	0.75	8033
Cost sensitive learning	0	8598	0	4577	52	570	47	429	24.57	3559
Metacost	0	669	0	2486	0	742	2	709	**0.25**	**1152**
Average	**0**	**3029.75**	11.25	2115.5	26	709.8	24.75	7188		

TABLE V
THE TOTAL NUMBER OF FALSE POSITIVES AND FALSE NEGATIVES OBTAINED BY RUNNING EACH CLASSIFIER ON U2R TESTING SUBSET USING THE FOUR COST SENSITIVE METHODS.



Method	Classifiers									
	DT		RF		SVM		LR		Average	
	FP	FN	FP	FN	FP	FN	FP	FN	FP	FN
Cost evaluation	1	1112	0	189	0	858	0	762	**0.25**	**730.25**
Cost sensitive classification	0	32436	0	32436	0	3954	0	685	0	17377.3
Cost sensitive learning	1	1112	0	18273	0	858	0	1273	0.25	5379
Metacost	0	2428	0	20048	0	1767	0	1073	0	6329
Avergae	0.5	9272	0	17737	0	1859	**0**	**948.25**		

[9] A. M. Chandrashekhar and K. Raghuveer. Performance evaluation of data clustering techniques using KDD Cup-99 Intrusion detection data set. International Journal of Information & Network Security (IJINS), Vol.1, No.4, pp. 294 305, October 2012.

[10] A. Mitrokotsa, C. Dimitrakakis and C. Douligeris. Intrusion Detection Using Cost-Sensitive Classification. In Proceedings of the 3rd European Conference on Computer Network Defense, Springer Science+Business Media, LLC 2009.

[11] Y-W. Seo and K. Sycara. Cost-Sensitive Access Control for Illegitimate Confidential Access by Insiders. In Proceedings of ISI 2006, pp. 117128, Springer-Verlag Berlin Heidelberg , 2006.

[12] B. Zadrozny, J. Langford, and N. Abe. A simple method for cost sensitive learning, IBM Tech Report, 2002.

[13] W. Lee, W. Fan, M. Miller, S. Stolfo, and E. Zadok. Toward cost sensitive modeling for intrusion detection and response. Journal of Computer Security, Volume 10, Issue 1-2, pp. 5-22, 2002.

[14] R. Lippmann, D. Fried, I. Graf, J. Haines, K. Kendall, D. McClung, D. Weber, S. Webster, D. Wyschogrod, R. Cunninghan, and M. Zissman. Evaluating intrusion detection systems: The 1998 darpa off-line intrusion detection evaluation. In Proceedings of the 2000 DARPA Information Survivability Conference and Exposition, January 2000.

[15] C. Elkan. The foundations of cost-sensitive learning. In Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4-10, 2001, pages 973978, 2001.

[16] I. Witten and E. Frank. Data Mining: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems). Morgan Kaufmann Publishers Inc., 2005.

[17] J. V. Hulse and T. M. Khoshgoftaar. Experimental perspectives on learning from imbalanced data. In Proceedings of International Conference on Machine Learning, 2007, pages 155164, 2007.

[18] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA Data Mining Software: An Update. SIGKDD Explorations, Volume 11, Issue 1, 2009.

[19] http://weka.wikispaces.com/Primer.September,2010.

[20] Weka/SpreadSubsample. Avilable at: https://algorithmia.com/algorithms/weka/SpreadSubsample

[21] P.-N. Tan, M. Steinbach, and V. Kumar. Introduction to Data Mining. Addison-Wesley, 2005.

[22] W. Wang and J. Yang. Mining High-dimensional Data. Data Mining and Knowledge Discovery Handbook, Chapter 27, pp. 793-799, 2005.

[23] A.C. Bahnsen, A. Stojanovic, D. Aouada, and B.E. Ottersten. Cost Sensitive Credit Card Fraud Detection Using Bayes Minimum Risk. In Proceedings of 12th International Conference on Machine Learning and Applications (ICMLA 2013), pp.333-338, Miami, FL, USA, December 4-7, 2013.