Improving Network Intrusion Detection Using Geometric Mean LDA

Elkhadir Zyad[1], Chougdali Khalid[2], and Benattou Mohammed[1]
(Corresponding author: Elkhadir Zyad)
LASTID Research Laboratory, Ibn Tofail University, Kenitra[1]
PO Box 242, Kenitra, Morocco
(Email: zyad.elkhadir@gmail.com)
GEST Research group National School of Applied Sciences (ENSA), Ibn Tofail University, Kenitra[2]
(Received May 15, 2017; revised and accepted Aug. 15, 2017)

Abstract

Anomaly based Intrusion Detection System (IDS) recognizes intrusion by adapting itself to identify normal behavior of the network. It then raises an alarm whenever any suspicious network behaviors are observed. Nonetheless, this kind of IDS is usually prone to small detection rate and high false positive rate due to difficulties involved in building normal network traffic pattern or a model. To avoid as much as possible this issue, many papers exploited a feature extraction method called linear discriminant analysis (LDA) as an intermediate step before constructing the model. Unfortunately, LDA has an important weakness, the class mean vector employed in this method is always estimated by the class sample average. That is not sufficient to provide an accurate estimate of the class mean, particularly with the presence of outliers. In this paper, to overcome that, we propose to use the geometric mean to estimate the class mean vector in LDA modeling. Many experiment on KDDcup99and NSL-KDD indicate that the proposed approach is more effective than numerous LDA algorithms.

Keywords: Geometric Mean; KDDcup99; LDA; Network Anomaly Detection; NSL-KDD

1 Introduction

The quick proliferation of various network tools which communicate and interact with each others have extremely increased the complexity of the network security and leads to the birth of sophisticated attacks. The classical security techniques such as user authentication, firewall and data encryption, are not able to fully cover the entire landscape of network security. As a consequence, they miss many damageable attacks. Hence, another type of protection is highly recommended, such as Intrusion Detection System (IDS). The latter takes part in containing the network breach by respecting appropriate preventive measures before any significant damages caused by the attacker.

IDS can be generally classified into two different categories: Signature-based IDS and Anomaly based IDS. In the first one, the IDS relies on a database of known attack signatures and produces an alarm wherever it exists any malicious network activities that correspond to one or more stored signatures. This kind of IDS has high detection rate against known attacks, but it is not able to detect new attacks. To overcome this limitation, frequent and expensive updates to the signature database are required. On the other hand, anomaly based IDS tries to build a normal behavior or model with the help of system and network characteristics. Here, the attack is considered as any deviation of traffic patterns from normal behavior. The main advantage of anomaly based IDS is it ability to identify new attacks.

Nevertheless, the actual network traffic data which are often enormous in size, are considered as a major challenge to anomaly based IDS. These kind of traffic slow down the entire detection process and lead in most of times to a biased classification accuracy. Such a large scale dataset usually contains noisy and redundant features which present critical challenges to knowledge discovery and data modeling.

To alleviate that, many feature reduction and feature selection methods have been successfully employed. For examples, the paper [5] proposed a cuttlefish based feature selection techniques to ensure data quality features and eliminate redundant and noisy features. The authors of [1] use Ant Colony Optimization algorithm to select important features. As a result, the IDS can accurately detect a broader range of attacks using smaller number of features. Following the same philosophy, the work [13] employed a Discrete Differential Evolution to identify the adequate features. The obtained results show a significant improvement in detection accuracy. In [7], the authors suggest to use Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA) as a pri-

mary step. After that, they classify network connections using k nearest neighbor (K-NN) and decision tree algorithms. In other publication [6], the same authors proposed an improved feature extraction method called PCA Lp using conjugate gradient. Applying this method on the two well-known datasets namely KDDcup99 and NSL-KDD prove the effectiveness of the proposed approach in terms of network attacks detection, false alarms reduction and CPU time minimization.

However, PCA and it variants offer great weights to features with higher variability whether they are effective or not. This fact may bring out the situation where the features have a lake of discriminating characteristics. To deal with that, the scientific community take the advantage of using linear discriminant analysis (LDA) [9] instead of PCA in many pattern recognition problems [3, 15, 21]. The key procedure behind this method is employing the well-known Fisher criterion to extract a linearly independent discriminant vectors and exploit them as basis by which samples are projected into a new space. These vectors contribute in maximizing the ratio of the between-class distance to within-class distance in the obtained space. Recent papers in network security field such [2, 8, 14] exploited an improved variant of this feature extraction method. Hence, this step provides the IDS with an important discrimination power. Meanwhile, it leads to a better attack identification.

In PCA and LDA mathematical formulations, class mean vectors take a significant part. For the first feature extraction technique, the class mean vector contributes in defining the covariance matrix. For the second one, the mean vectors take part in creating the between-class and within-class scatter matrices. However, such vectors are estimated by the class sample averages. Since there are many outliers and some abnormal classes that contain only a few training samples, it becomes difficult to give an accurate estimate of the class mean vectors using the class sample average.

In order to solve the mean calculation issue, a numerous papers had proposed different approaches. For instance, the authors of [12] suggest an algorithm which automatically removes the correct data mean with proved convergence. Experiments on face image datasets show that the approach consistently outperforms many PCA methods. The paper [11] uses a within-class maximum - minimum - median - average vector to construct within-class scatter matrix and between-class scatter matrix instead of within-class mean vector. Recently, the work [23] proposes a harmonic mean based LDA which makes use of weighted harmonic mean of pairwise between-class distance and gives higher priority to maximize small between-class distances. This approach shows a good results when it is applied to many multi-label data sets.

To overcome this weakness in context of intrusion detection, this paper proposes to use the class geometric mean vector [16] to approximate the class mean vector. The class geometric mean vector is less sensitive to outliers. Thus, the geometric mean LDA model should be more robust than the current sample-average based LDA. We will prove this by numerous experiments using two popular data sets namely KDDcup99 and NSL-KDD.

The rest of this paper is organized as follows. In Section 2, we outline LDA. Section 3 presents in details the proposed method. Section 4 introduces the two well known network datasets KDDcup99 and NSL-KDD. Section 5 provides the experimental results and illustrates the effectiveness of the algorithm by comparing it to some LDA approaches. Finally, Section 6 offers our conclusions.

2 Linear Discriminant Analysis

LDA [9] seeks to find a projection matrix G such that the Fisher criterion is maximized after the projection of samples. Suppose X is composed of k classes, $[X_1, .., X_k]$. Every X_i contains n_i samples. The between-class and within-class scatter matrices S_b and S_w, are defined by

$$S_w = (1/n)\sum_{i=1}^{k}\sum_{x\in X_i}(x-m_i)(x-m_i)^T \quad (1)$$

$$S_b = (1/n)\sum_{i=1}^{k}(m_i-m)(m_i-m)^T. \quad (2)$$

m_i is the mean of the ith class, and m is the general mean. They are defined as follow:

$$m_i = \frac{1}{n_i}\sum_{x\in X_i}(x) \quad (3)$$

and

$$m = \frac{1}{n}\sum_{i=1}^{k}\sum_{x\in X_i}(x). \quad (4)$$

The Fisher criterion is defined by

$$G = \arg\max\frac{G^T S_b G}{G^T S_w G}. \quad (5)$$

When S_w is invertible, the solutions to Equation (5) can be obtained by performing the following generalized eigenvalue decomposition:

$$S_w^{-1} S_b g_i = \lambda_i g_i. \quad (6)$$

Where $G = [g_1, \ldots, g_l]$ and l is the number of eigenvectors g_i that correspond to the largest eigenvalues λ_i.

From Equations (1) and (2), it is clear that the class mean vector contributes significantly in formulation of the between-class and within-class scatter matrices. Thus, it precision must have crucial effect on the resulting linear discriminant vectors G. However, the class sample average vector may not approximate precisely the class mean vector (Equations (3) and (4)), when there are only a few samples available for training per class. Furthermore, there are numerous studies which affirms that sample average may not be representative of the true central region

for skewed data or data with outliers. In network intrusion case, since there are some classes such as U2R and R2L attacks which provides a few training samples, the resulting matrix G will be seriously blurred.

3 Geometric Mean LDA Formulation

3.1 Geometric Mean Vector

In probability theory and statistics, the geometric mean m_g of a set of n positive numbers $x_1, x_2, ..., x_n$ is defined as:

$$m_g = (x_1 \times x_2 \times ... \times x_n)^{\frac{1}{n}}. \quad (7)$$

As sample average, the geometric mean [16] can also be used to estimate the central tendency. Furthermore, it is generally considered that this measure is more resistant to outliers (or skewed data). That what we can see in the following example: Suppose we have Data=[3.3, 3.0, 10, 3.1, 1, 3.2, 3.4] with the outliers "1" and "10" then m= 3.857 and m_g=3.186. We observe that 3.186 is more closer to the central tendency ($\frac{3+3.1+3.2+3.3+3.4}{5} = 3.2$) than 3.857. The geometric mean of a non negative matrix:

$$Z = [Z_1, Z_2, .., Z_n] = \begin{bmatrix} Z_{11} & Z_{21} & Z_{31} & \dots & Z_{n1} \\ Z_{12} & Z_{22} & Z_{32} & \dots & Z_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_{1d} & Z_{2d} & Z_{3d} & \dots & Z_{nd} \end{bmatrix}$$

is given by $m_g = (m_{g1}, m_{g2}, .., m_{gn})$ where m_{gi} is the geometric mean of elements on the i-th column of the data matrix Z.

3.2 Geometric Mean LDA

In small sample cases, the class geometric mean vector m_{gi} generally ensures a better representation of true central tendency, in particular when outliers exist in the training samples. Additionally, it is worthwhile to highlight another merit of geometric operator for dealing with outliers. Differing from many outlier-removing methods that just eliminate outliers from the training sample set, the geometric mean operator is able to derive useful information from it.

Based on all the geometric mean merits, m_{gi} and m_g will be used as estimators of the class mean vector m_i and general mean m. To avoid the singularity of the within-class scatter matrix, we would apply PCA [10] on X and get the PCA-projected matrix X_{PCA} with the help of the equation:

$$X_{PCA} = W^T X.$$

Where W is the projection matrix that contains the principal components (PCs). Then, instead of working with X we operate on $|X_{PCA}|$. We apply the absolute value on X_{PCA} in order to make possible the calculation of geometric mean given by Equation (7). After that we compute the new S_w^g and S_b^g with the formulas:

$$S_w^g = (1/n) \sum_{i=1}^{k} \sum_{x \in X_i} (x - m_{gi})(x - m_{gi})^T$$

$$S_b^g = (1/n) \sum_{i=1}^{k} (m_{gi} - m_g)(m_{gi} - m_g)^T.$$

The new proposed Fisher criterion will be defined by

$$G' = \arg\max \frac{G'^T S_b^g G'}{G'^T S_w^g G'}.$$

The solutions to the above problem is reached by:

$$(S_w^g)^{-1}(S_b^g) g_i' = \lambda_i' g_i'.$$

Where $G' = [g_1', \dots, g_l']$. The projection of a new vector x_{new} on the space constructed by our approach is obtained by:

$$t_i = (G')^T x_{new}.$$

Hereafter the algorithm is called geomean LDA.

4 The Simulated Databases

4.1 KDDcup99

The objective of 1999 KDD intrusion detection contest is to create a standard dataset [18] to evaluate research in intrusion detection. The dataset is prepared and managed by DARPA Intrusion Detection Evaluation Program. It is composed of many TCPdump raws, captured during nine weeks.

The first seven weeks were devoted to create training data. The latter represents four gigabytes of compressed binary TCP dump data, equivalent to five million connection records. Similarly, in last two weeks, the program captured around two million connection records and considered it as testing data. The KDD dataset was employed in the UCI KDD1999 competition whose goal is developing intrusion detection system models. the attacks simulated in this competition fall into four main categories: DOS, R2L, U2R, PROBE. In the first category an attacker tries to prevent legitimate users accessing or consume a service via back, land, Neptune, pod Smurf and teardrop. In R2L, the attacker tries to gain access to the victim system by compromising the security via password guessing or breaking. To perform U2R, the intruder tries to access super users (administrators) privileges via Buffer overflow attack. The last type of attack consists in gaining information about the victim machine by checking vulnerability on the victim machine. *e.g.*, Port scanning.

The KDD Cup99 dataset is available in three different files such as KDD Full Dataset which contains 4898431 instances, KDD Cup 10% dataset which contains 494021 instances, KDD Corrected dataset which contains 311029 instances. In this paper, training data are taken from KDD Cup 10% and testing data from KDD Corrected dataset.

Each sample of the dataset is a connection between two network hosts according to network protocols. It is described by 41 attributes. 38 of them are continuous or discrete numerical attributes, the other are categorical attributes. Each sample is labeled as either normal or one specific attack. The dataset contains 23 class labels out of which 1 is normal and remaining 22 are different attacks. The total 22 attacks fall into four categories as forth-mentioned attacks.

KDD Cup 99 features can be classified into three groups:

1) Basic features: This category represents all the attributes that can be extracted from a TCP/IP connection. Most of these features leading to an implicit delay in detection.

2) Traffic features: This category contains features that are computed with respect to a window interval.

3) Content features: The majority of DoS and Probing attacks have many intrusion frequent sequential patterns, this is due to the fact that these attacks establish many connections to the host(s) in a very short period of time. Unlike these attacks, the R2L and U2R attacks do not have any intrusion frequent sequential patterns. The R2L and U2R attacks are embedded in the payload of the packets, and normally include only a single connection. To identify these kinds of attacks, some relevant features are needed to identify suspicious behavior in the packet payload. These features are called content features.

4.2 NSL-KDD

NSL-KDD [19] is a data set proposed to solve some of the shortcomings of the KDD'99 data set discussed in [17]. To summarize, the new dataset proposes a reasonable number of train records (125973 samples) and test sets (22544 samples). This advantage makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. Consequently, evaluation results of different research work will be consistent and comparable. In addition, there is no redundancy sample present in the dataset and testing set contains some attack which are not present in the training set.

5 Experiments and Discussion

In this section, several experiments were designed to demonstrate the effectiveness of our proposed method. In order to show it high accuracy in an all-round way, we compare geomean LDA with other popular methods such as LDA [9], Direct LDA [22], median LDA [20], null space LDA [4]. KDDcup99 and NSL-KDD were selected for evaluation.

To estimate the accuracy of these methods we employ two factors:

$$DR = \frac{TP}{TP+FN} \times 100$$
$$FPR = \frac{FP}{FP+TN} \times 100.$$

(DR) and (FPR) mean Detection Rate and False Positive Rate. True positives (TP) refer to attacks correctly predicted. False negatives (FN) represent intrusions classified as normal instances, false positive (FP) are normal instances wrongly classified, and true negatives (TN) are normal instances classified as normal. Based on the above measures, the most reliable feature extraction method will be the one which improves DR as much as possible and tries to minimize FPR.

Concerning the experiments settings, we decide to vary the number of training samples and keep test dataset unchanged with the following composition (100 normal data, 100 DOS data, 50 U2R data, 100 R2L data, and 100 PROBE). The way we modify training samples consists in increasing the number of DOS and PROBE attacks on the one hand, on the other hand, we set normal training data at 1000 samples. U2R and R2L samples are fixed at 100. In order to get a realistic detection rate (DR) and FPR, the operation of sample selection was done randomly for thirty times. Then DR and FPR took the average. Since our goal is evaluating the efficacy of feature extraction method, we use a simple classifier, the nearest neighbor classifier.

Table 1: Detection rate (%) of geomean LDA, LDA and median LDA in different PCA and LDA space dimensions

The method	PCs	LDs	KDDcup99	NSL-KDD
Geomean LDA	3	3	60.60	60.38
	3	2	58.26	59.09
	3	1	48.52	52.39
LDA	3	3	61.68	54.35
	3	2	59.61	58.43
	3	1	50.45	42.70
Median LDA	3	3	60.63	58.91
	3	2	59.71	56.44
	3	1	40.88	42.12

To avoid the singularity of the within-class scatter matrix in LDA and median LDA, we employ PCA as first dimension reduction, then the algorithms are performed

Table 2: Detection rate (%) Direct LDA and Null space LDA in different LDA space dimensions

Database	The method	5 LDs	4 LDs	3 LDs
KDDcup99	Null space LDA	59.48	60.86	59
	Direct LDA	42.85	46.28	60.28
NSL-KDD	Null space LDA	57.58	58.84	58.37
	Direct LDA	47.14	46.85	50.28

in the PCA-transformed space. The aim of the first experiment is to find the adequate dimension of the subspace transformed by PCA, such that the algorithms can be applied and give optimal results (high DR and less FPR). One proposition to do that is fixing training data at 1000 normal, 100 DOS, 50 U2R, 100 R2L, 100 PROBE. Then we apply LDA in different PCA dimension spaces and pick up the values which ensure good DR. Table 1 shows this manipulation on the two databases, number of PCs means dimension kept after applying PCA, and LDs refer to the number of top discriminant vectors. We observe that choosing three PCs and three LDs contribute significantly to get a higher DR for both median LDA and the proposed approach. LDA excels with three PCs and two LDs on NSL-KDD. In the same frame of mind, we look for the number of LDs that improve the rest of LDA models efficiency. According to TABLE II, we note that three discriminant vectors ensures good DR for Direct LDA. Null space LDA needs four discriminant vectors. Thus, all the stated above LDA models will use these parameters in the next experiments.

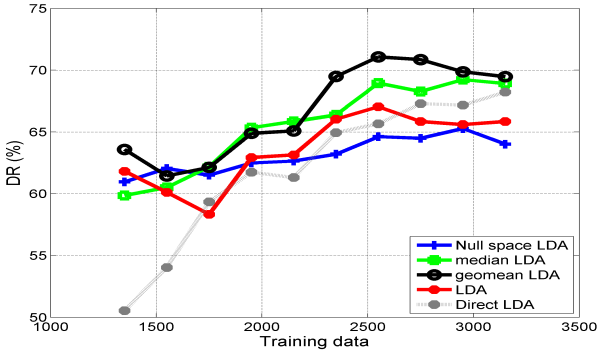

Figure 1: DR of geomean LDA, LDA, Null space LDA and median LDA on KDDcup99

Figures 1 and 2 exhibit the results we found when we compare our approach to the aforementioned LDA models for KDDcup99. According to Figure 1, we observe that geomean LDA overcomes all LDA models once the training data surpasses 2000. The reason behind this phenomenon seems to be that more there are training samples more the effect of outliers is visible. Since the other algorithms except median LDA work with the class sample average vector, they will be more sensitive to outliers, that what decrease their efficiency. Median LDA is also resistant to the skewed data because it estimates the class mean vector by a robust measure which is the median vector. However it still inferior to geomean LDA in this case. A possible explanation of this fact may reside in the distribution nature of data. We expect that the data follows a log-normal distribution which gives an important advantage to geometric mean. Concerning FPR, Figure 2 asserts that the proposed method produces a false positive rate lower than 2.7%. This means that the method is very able to distinguish normal instances from attacks.

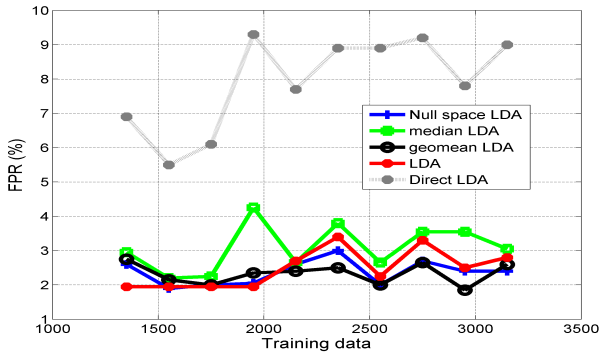

Figure 2: FPR of geomean LDA, LDA, Null space LDA and median LDA on KDDcup99

On NSL-KDD, we see from Figures 3 and 4 that geomean LDA achieves at least 3% improvement over LDA and Null space LDA, 1% over median LDA. The approach takes the lead permanently over Direct LDA. In term of FPR, Figure 4 shows that geomean LDA still gives the fewest values in the company of median LDA and LDA.

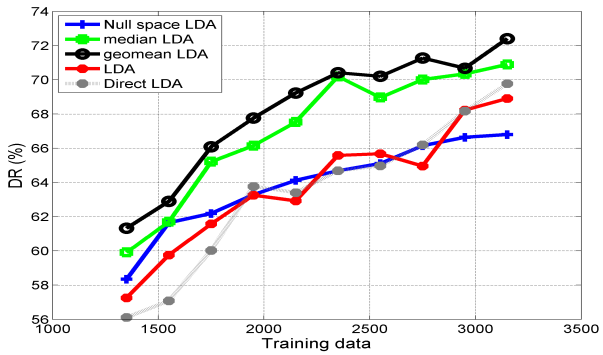

Figure 3: DR of geomean LDA, LDA, Null space LDA and median LDA on NSL-KDD

In the next experiment, we evaluated the proposed method when changing the parameter K of K-NN classifier. In order to make the manipulation possible, we fixed a number of training data having the following settings: 1000 normal, 100 DOS,50 U2R, 100 R2L and 100 PROBE

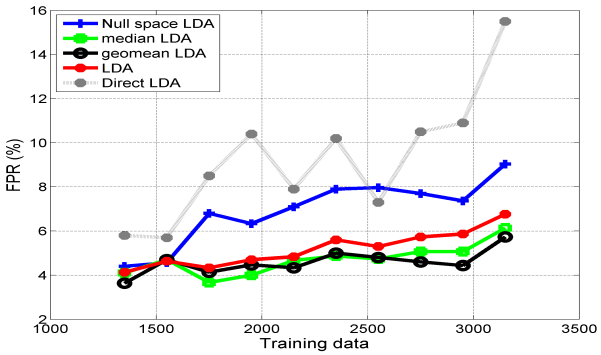

Figure 4: FPR of geomean LDA, LDA, Null space LDA and median LDA on NSL-KDD

instances. Then, we increased K and show it effect on DR and FPR.

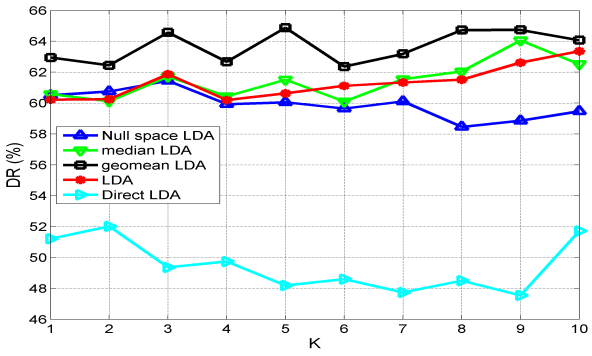

Figure 5: K vs. DR(%) for KDDcup99

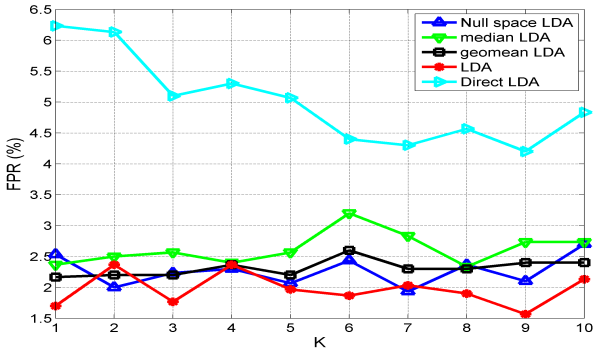

Figure 6: K vs. FPR(%) for KDDcup99

From Figure 5 we can see that geomean LDA preserves it superiority in giving high DR. It produces at least 62% and achieves 65% as a maximal detection rate when $K = 5$. Moreover, the figure asserts that the proposed method overcomes all the other aforementioned LDA variants. In term of false positive rate, we observe from Figure 6 that geomean LDA has the fewest false positive rate in the company of Null space LDA and LDA.

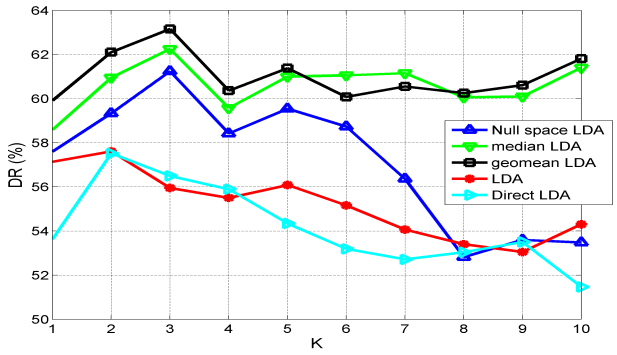

Figure 7: K vs. DR(%) for NSL-KDD

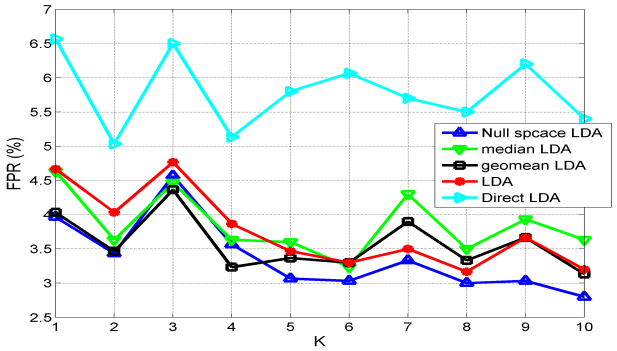

Figure 8: K vs. FPR(%) for NSL-KDD

When we reproduce the same experiment on NSL-KDD, we get the following results: From Figure 7, it is clear that geomean LDA and median LDA are the LDA variants which ensure a better DR. The false positive rate of the proposed method is acceptable. In fact, it produces a less FPR than Direct LDA and median LDA. However, the other LDA methods take the advantage once K surpasses 4.

6 Conclusion

In this paper, we improve the robustness of LDA in detecting network intrusions, by using class geometric mean vector, rather than the class sample average. Therefore, the proposed method called geomean LDA is more robust to outliers and preserve useful discriminant information. Experiments on KDDcup99 and NSL-KDD demonstrate the effectiveness of the proposed model, meanwhile they show it superiority over some Fisher's LDA-based algorithms such as classical LDA, null space LDA, median LDA and Direct LDA.

References

[1] M. H. Aghdam and P. Kabiri, "Feature selection for intrusion detection system using ant colony op-

timization," *International Journal Network Security*, vol. 18, no. 3, pp. 420–432, 2016.

[2] R. A. R. Ashfaq, X. Z. Wang, J. Z. Huang, H. Abbas and Y. L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[3] J. B. Calfa, J. M. Buenaposada and L. Baumela, "Robust gender recognition by exploiting facial attributes dependencies," *Pattern Recognition Letters*, vol. 36, pp. 228–234, 2014.

[4] L. F. Chen, H. Y. M. Liao, M. T. Ko, J. C. Lin and G. J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognition*, vol. 33, no. 10, pp. 1713–1726, 2000.

[5] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2670–2679, 2015.

[6] Z. Elkhadir, K. Chougdali, and M. Benattou, "Network intrusion detection system using pca by lp-norm maximization based on conjugate gradient," *International Review on Computers and Software*, vol. 11, no. 1, pp. 64–71, 2016.

[7] Z. Elkhadir, K. Chougdali, and M. Benattou, "Intrusion detection system using pca and kernel pca methods," in *Proceedings of the Mediterranean Conference on Information & Communication Technologies*, pp. 489–497, 2016.

[8] Z. Elkhadir, K. Chougdali, and M. Benattou, "A median nearest neighbors lda for anomaly network detection," in *International Conference on Codes, Cryptology, and Information Security*, pp. 128–141, 2017.

[9] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic press, 2013.

[10] I. Jolliffe, *Principal Component Analysis*, Wiley Online Library, 2002.

[11] L. Li, H. Ge, and J. Gao, "Maximum–minimum–median average msd-based approach for face recognition," *AEU-International Journal of Electronics and Communications*, vol. 70, no. 7, pp. 920–927, 2016.

[12] F. Nie, J. Yuan, and H. Huang, "Optimal mean robust principal component analysis," in *International Conference on Machine Learning*, pp. 1062–1070, 2014.

[13] E. Popoola and A. O. Adewumi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision.," *International Journal Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[14] A. A. Saad, C. Khalid, and J. Mohamed, "Network intrusion detection system based on direct LDA," in *Third World Conference on Complex Systems (WCCS'15)*, pp. 1–6, 2015.

[15] C. S. Silva, F. D. S. L. Borba, M. F. Pimentel, M. J. C. Pontes, R. S. Honorato, and C. Pasquini, "Classification of blue pen ink using infrared spectroscopy and linear discriminant analysis," *Microchemical Journal*, vol. 109, pp. 122–127, 2013.

[16] S. Stević, "Geometric mean," in *International Encyclopedia of Statistical Science*, pp. 608–609, 2011.

[17] 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 Symposium on Computational Intelligence for Security and Defence Applications*, pp. 324, 2009.

[18] UCI, *KDD Cup 1999 Data*, The UCI KDD Archive Information and Computer Science University of California, Irvine, Oct. 2014.

[19] UNB, *NSL-KDD Data Set for Network-based Intrusion Detection Systems*, Mar. 2014.

[20] J. Yang, D. Zhang, and J. Y. Yang, "Median LDA: A robust feature extraction method for face recognition," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 5, pp. 4208–4213, 2006.

[21] M. Yang and S. Sun, "Multi-view uncorrelated linear discriminant analysis with applications to handwritten digit recognition," in *International Joint Conference on Neural Networks (IJCNN'14)*, pp. 4175–4181, 2014.

[22] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional datawith application to face recognition," *Pattern Recognition*, vol. 34, no. 10, pp. 2067–2070, 2001.

[23] S. Zheng, F. Nie, C. Ding, and H. Huang, "A harmonic mean linear discriminant analysis for robust image classification," in *IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI'16)*, pp. 402–409, 2016.

Biography

Elkhadir Zyad is a PhD student in Faculty of science, Ibn Tofail University, Kenitra, Morocco. He obtained his Master degree in computer science in 2013 from the same Faculty. He is an IEEE member. His main research interest is to develop new feature extraction algorithms for pattern recognition problem such as network intrusion detection.

Chougdali Khalid is an associate Professor of Computer Science at the National School of Applied Sciences, Knitra. In 2010 he obtained his PhD degree from Mohamed V-Agdal university in computer science. His main research interest are network security, pattern recognition and biometrics.

Mohamed Benattou is a Professor of Computer Science at the IBN TOFAIL University KNITRA where he directs the Computer Science and Telecommunication Laboratory. He has also held several positions in his French academic career: University of PAU, University of ORSAY Paris XI, 3IL and Xlim Laboratory. His research interests include distributed testing, secure testing, and software testing.