Intrusion Detection Systems (IDSs) using Multivariate Control Chart Hotelling's T^2 with Dimensional Reduction of Factorial Analysis of Mixed Data (FAMD) and Autoencoder

Kevin Agung Fernanda Rifki [1*], Niam Rosyadi[1], Amanatullah Pandu Zenklinov[1], and Novri Suhermi[1]
[1]Department of Statistics, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia
*Corresponding author: fernandakevin02@gmail.com

Received: 2 September 2023
Revised: 5 Marh 2024
Accepted: 7 Marh 2024

ABSTRACT – Traditional multivariate control charts for network intrusion detection encounter significant challenges including false alarms due to non-conforming network data traffic distributions, limitations in identifying outlier intrusions caused by masking effects, and handling diverse data types. This paper introduces a T^2-based multivariate control chart that leverages dimensional reduction techniques using Factor Analysis of Mixed Data (FAMD) and Autoencoder to address these issues. FAMD reduces data with both quantitative and qualitative variables, while Autoencoder focuses on dimensionality reduction for quantitative variables, enhancing multivariate control chart performance. The proposed chart, a modified T^2, is compared to conventional T^2 with dimensionality reduction through FAMD and Autoencoder. Results from simulating data using UNSW-NB 15 demonstrate T^2's superior performance with dimensionality reduction compared to conventional T^2. Under various conditions, conventional control chart T achieves 64% accuracy, T^2 with FAMD achieves 74%, and T^2 with Autoencoder reaches 76%. Notably, T^2 with FAMD excels in detecting normal activity as intrusion compared to Autoencoder. This approach holds promise for improving network intrusion detection accuracy, especially in mixed-data environments.

Keywords– Autoencoder; FAMD; Hotelling T^2 Control Chart; Intrusion Detection

I. INTRODUCTION

The massive growth in computer networks and application causes many challenges for cyber security, such as intrusion in network system. Intrusions/attacks can be defined as a security event, or a combination of multiple security events, that constitutes a security incident in which an intruder gains, or attempts to gain, access to a system or system resource without having authorization, encompass availability, authority, confidentiality and integrity [1]. Intrusions Detection Systems (IDSs) are systems that try to detect attacks as they occur or after the attacks took place [2].

Statistical Process Control (SPC) has been widely used in many fields, particularly in industry and services [3]. Statistical process control not only can be applied to monitor the manufacturing or industrial processes but also can be utilized for intrusions detection systems. In network monitoring and intrusion detection, statistical process control can be used as a powerful tool to guarantee safety and stability in a network system [4].

Among the statistical process control tools, the most commonly used is the control chart (CC), which is actually a graphical representation of a function of the sample values (say, $g(x)$) of a variable related to the quality of the final product versus the sample number (or time). In addition, the control chart is supplemented with a central line (CL), an upper control limit (UCL), and a lower control limit (LCL). These limits are established using the distributional properties of $g(x)$. The control chart is used to determine whenever the monitored process is statistically In Control (IC) or Out of Control (OC). To achieve this goal, consecutive values of $g(x)$ are plotted against the control limits (UCL and LCL). If the values of $g(x)$ are inside the interval [LCL, UCL] then the process is considered as in control, otherwise the process is considered as out of control.

In case that the process stability/quality is characterized by the values of a single variable then the control chart is called univariate. However, there are many cases in which the simultaneous monitoring of two or more variables is deemed necessary. Statistical process control techniques involving the monitoring of multiple dependent variables are known as multivariate statistical process control techniques. The multivariate control chart Hotelling's T^2 is a commonly used tool for monitoring simultaneously several correlated or uncorrelated quality characteristics of a process [5]. The multivariate control chart Hotelling's T^2 can be exploited in monitoring network attacks as an intrusions detection system.

In the theory, network intrusion detection can be monitored by using Hotelling's T^2 chart technique. Nevertheless, there are two arguments why this method is not suitable to be employed for this case ([6], [7]). Firstly, the intrusion detection system involves large volumes of high-dimensional connection. Secondly, the network monitoring system requires a fast computational process so that an anomaly can be quickly detected. In fact, the effectiveness of conventional multivariate control charts such as Hotelling's T^2 is increased for a small number of quality characteristics. If large number of quality characteristics used then the performance of control chart to detect any shift in a process may be decreased [8]. Large numbers of highly correlated quality characteristics often take place in modern manufacturing processes. As a result, the computation of the T^2 statistic is difficult due to the singularity of the covariance matrix ([9], [10]).

To overcome the problems, arise in monitoring large number of quality characteristic, dimensionality reduction technique has been used to transform high-dimensional data into a lower dimensional space while preserving meaningful characteristics of the original data. Principal component Analysis (PCA) is the most widely used method for dimension reduction [11]. However, it has its limitation on linearity assumption and is unsuitable for data containing both numeric and categorical types of data [12] and non-linear data [13].

Factor analysis of mixed data or factorial analysis of mixed data (FAMD, in the French original: AFDM or Analyse Factorielle de Données Mixtes) is a dimension reduction method that can be used for data with mixed types of variables. The term mixed refers to the use of both quantitative and qualitative variables [14]. An Autoencoder is a type of artificial neural network used to learn efficient coding of unlabeled data (unsupervised learning). The encoding is validated and refined by attempting to regenerate the input from the encoding. The autoencoder learns a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore insignificant data ("noise") [15]. FAMD is used to overcome numeric and categorical types of data while Autoencoder is used to overcome non-linearity data.

In 2005, Qu et al. in [16] used the Hotelling's T² chart to monitor the intrusion of a network. Furthermore, the system so-called real-time Multivariate Analysis for Network Attack (MANA) detection algorithm. The Multivariate Analysis for Network Attack control limits will be updated continuously at certain intervals of time.

In 2006, the Chi-Square Distance Monitoring (CSDM) method is developed by Ye et al [17] and it is applied to monitor the uncorrelated, correlated, autocorrelated, normal, and non-normal distributed data. In general, Chi-Square Distance Monitoring performs better than Hotelling's T² to detect a shift in the mean, especially in uncorrelated, autocorrelated, and non-normally distributed data. Meanwhile, Hotelling's T² has better performance than Chi-Square Distance Monitoring for correlated and normally distributed data.

In 2015, Sivasamy and Sundan in [18] compared the performance of Hotelling's T control charts with Support Vector Machine (SVM) and Triangle Area-based Nearest Neighbors (TANN) methods with result high accuracy Hotelling's for all types of attack classes.

Based on the aforementioned above, the integration between FAMD with T² chart and Autoencoder with T² chart are a good alternative to solve the problems. This paper will focus on creating IDSs using multivariate T² control chart based on FAMD and multivariate T² control chart based on Autoencoder. UNSW-NB15 dataset would be used to evaluate the performance of proposed IDS. Moreover, the performance of the proposed method is compared with the existing T² chart.

II. LITERATURE REVIEW

A. Hotelling's T² Chart

Hotelling's T² Chart [19] is one of the multivariate control charts employed in monitoring the quality of products [8]. Let random vectors x_i are followng the multivariate normal distribution with certain μ and $\Sigma, i = 1, 2, \ldots, n$. Further, the $n \times p$ can be denoted as $X = [x_1', x_1', x_2', \ldots, x_n']'$. The statistics Hotelling's T² is formulated according to Equation (1):

$$T_i^2 = (x_i - \bar{x}_i)' S^{-1} (x_i - \bar{x}_i) \quad (1)$$

Where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $S = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$

By the the asumption that the original data $x_i \sim N_p(\mu, \Sigma)$, the control limit of T² chart can be written as

$$CL = \frac{p(n+1)(n-1)}{n^2 - np} F_{(\alpha, p, (n-p))} \quad (2)$$

with n represents the number of samples, p indicates the number of quality characteristics, and α denotes the false alarm rate.

Table 1 Input Parameters and Levels

	Prediction	
	Intrusion	Normal
Intrusion	True Positive (TP)	False Negative (FN)
Normal	False Positives (FP)	True Negatives (TN)

Moreover, the performance of IDSs would be evaluated by the confusion matrix as shown in Table 1. The accuracy of a classification method could be measured by the degree of accuracy and degree of error. The accuracy in detecting intrusion can be divided into two types:

a. True Positives (TP) is number of successful attacks that is concluded as an attack.

b. True Negatives (TN) is number of normal activities that are successfully detected as normal activity.

The misdetection in intrusion detection can be divided into two types:

a. False Positives (FP) is number of normal activities that are detected as an attack.

b. False Negative (FN) is number of successful attack that are detected as normal activity.

FP cause a false alarm while FN allows an attack on the system. The level of accuracy used is the hit rate that can be calculated as follows:

$$\text{Hit Rate} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{3}$$

Based on the type of inaccuracy, the level of misdetection in intrusion detection can be divided into two types, namely FP rate and FN rate which can be written as follows:

$$FP \text{ Rate} = \frac{FP}{TN + FP} \tag{4}$$

$$FN \text{ Rate} = \frac{FN}{TP + FN} \tag{5}$$

B. FMAD Dimensional Reduction

As we mentioned earlier, when data contains both numeric and categorical variables, FAMD acts PCA in dealing with numeric variables and acts as MCA in dealing with categorical variables. Suppose there are J numeric variables and Q categorical variables, where k_q denotes the number of categorical of the qth qualitative variable. Let P_{k_q} denote the proportion of individuals processing k_q. Let H denote the total categorical for all the qualitative variables.

When processing the data, the numerical variables and the categorical variables are standardized as those described in PCA and MCA, respectively. The weight for each individual is still *1/N*, but instead of assigning the weight of each level to P_{k_q}/H as that in MCA, the weight of each level of a categorical variable. As a result, in space $\mathbb{R}^k$, each numeric variables have inertia of 1 and is represented by a vector; each categorical variable has total inertia of $k_q - 1$ and is represented by k_q vectors. When projecting the total inertia of $k_q - 1$ on each dimension of the subspace of a categorical variable, the projected inertia is 1. Therefore, when searching for the new axes with maximum inertia, the two types of variables are on the equal step.

When search for a new principal component in FAMD, we maximize the sum of the square correlation between numeric variables and the principal component plus the sum of the square correlation ratio between categorical variables and the principal component. The contribution of individual i (or a variable) to a principal component can be calculated in a similar sense as that in PCA and the quality of representation is defined as the cosine of the angle θ_{kj}, which is the correlation coefficients of variable k and variable j.

C. Autoencoder Dimensional Reduction

First of all, an autoencoder is an unsupervised neural network, whose objective is to learn to reproduce input vectors $\{x(1), x(2), \dots, x(m)\}$ as output $\{\hat{x}(1), \hat{x}(2), \dots, \hat{x}(m)\}$. Figure 1 and Figure 2 shows an autoencoder and layer L_2 is the hidden layer, whereby the inputs are compressed into a small number of neurons. Activation of unit i in layer I is given by Equation 6:

$$a_i^{(l)} = f\left(\sum_{j=1}^{n} W_{ij}^{(l-1)} a_j^{(l-1)} + b_i^{(1)}\right) \tag{6}$$

where W and b denote weight and bias parameters respectively. In the first layer, i.e., the input layer, $a^{(1)} = x$, and in the last layer, i.e., the output layer, $a^{(1)} = \hat{x}$. For the activate function f, we used sigmoid function in hidden layers, but in the output layer, we used linear function since we don't pre-scale every input example to a specific interval like [-1,1].

During the training period, we minimize the objective function shown in Equation 7 with respect to W and b. The objective function includes the regularization term, and the parameter λ determines the strength of regularization.

$$J(W,b) = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{2}\left\|x(i) - \hat{x}(i)\right\|^2\right) + \frac{\lambda}{2}\sum_{l=1}^{n_l - 1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_l+1}\left(W_{ji}^{(l)}\right)^2 \tag{7}$$

Where n_i denotes number of layers in the network and s_i denotes number of units in layer L_i

Recently a denoising autoencoder [20], which is one of the extensions of an autoencoder, has been developed. The idea is to learn an over-complete set of basis vectors to represent input vectors, so that our basis vectors can capture structures and patterns inherent in the input data better. At the same time, in order to avoid highly compressed encoding which is usually highly entangled, we can encode the input with a small subset of neurons. We can achieve this by increasing the number of hidden units and adding some noise to the input. There are some ways in adding the noise to each input, but in this work, we destroy the input by randomly choosing a fixed number of components of the input to be 0, which is sometimes called as the salt-and-pepper noise [21].

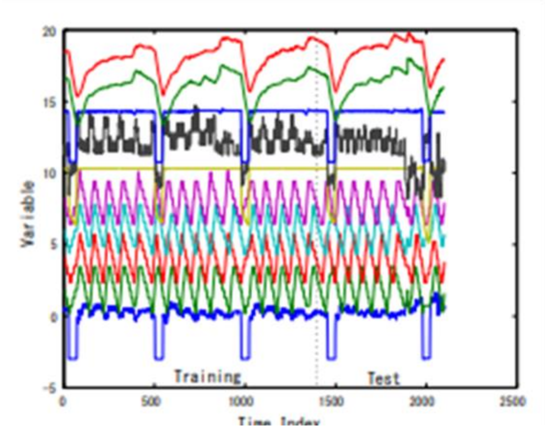

Figure 1 Normal {z(1),z(2),...,z(849)} (blue) amd anomalous (red) data from Lorenz system.

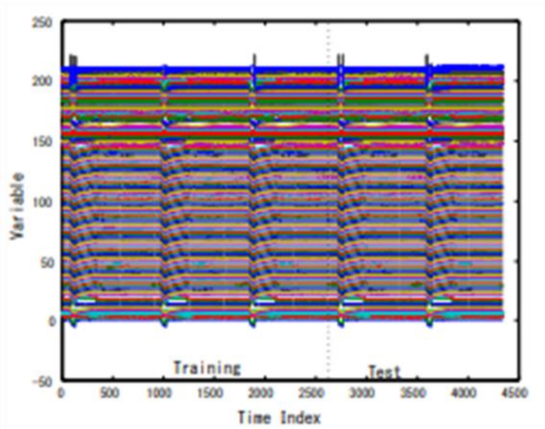

Figure 2 Normalized 25 Dimensional Lorenz system data *x*

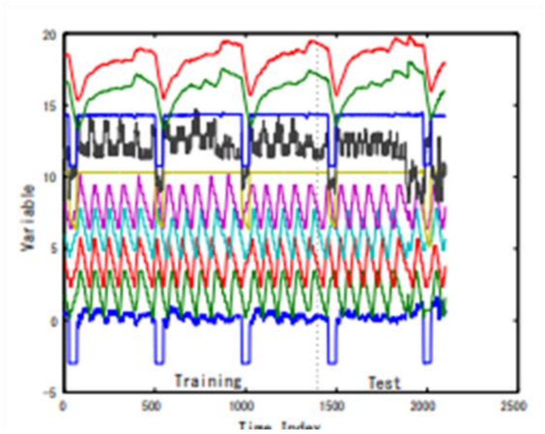

Figure 3 Normalized data of Satellite-A

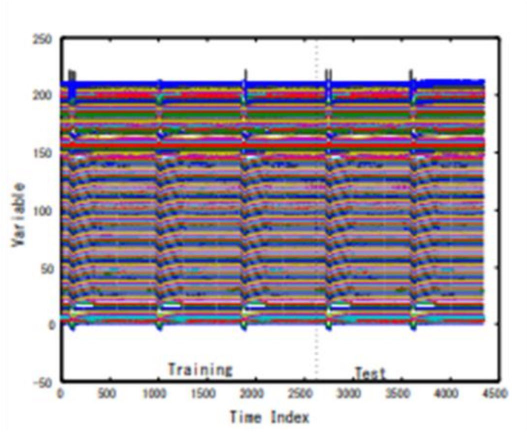

Figure 4 Normalized data of Satellite-B.

III. RESULTS AND DISCUSSIONS

A. Data Representative

UNSW-NB 15 is a large dataset produced by the Australian Network Security Centre (ACCS) in 2015. The data set collected modern normal data and nine types of data traffic attacks. UNSW-NB15 this dataset is enough to propose the actual conditions in the actual network environment. The Data is divided into two, namely Training data and Testing data where each is divided into 82,332 data and 175,341 data with both having 42 attributes and each data is divided into 2, namely normal and no attack. Attacks are divided into 9 types: Fuzzers (1), Analysis (2), Backdoors (3), DOS (4), Exploits (5), Generis (6), Reconnaissance (7), Shellcode (8), and Worms (9).

UNSW-NB15 dataset can also be called a mixture of high-dimensional datasets with 37 continuous numeric attributes and 5 discrete categorical attributes. In the characteristics of the data, the percentage of normal data of 44.94% and 55.06% attack data for Training data while for Testing data has a percentage of normal data of 31.938% and 68.062% attack data. Based on these results for testing data slight unbalance where the proportion of attack data is greater than normal data and for the proportion of each type of attack data is very unbalanced. The characteristics of the dataset can be seen in the following table.

Table 2 UNSW-NB 15 Dataset

Class		Training		Testing	
		Size	%	Size	%
Normal		37000	44.94	56000	31.938
Attack Type	1	18871	22.921	40000	22.813
	2	11132	13.521	33393	19.045
	3	6062	7.363	18184	10.371
	4	4089	4.966	12264	6.994
	5	3496	4.246	10491	5.983
	6	677	0.822	2000	1.141
	7	583	0.708	1746	0.996
	8	378	0.459	1133	0.646
	9	44	0.053	130	0.074
Total		82332	100	175341	100

B. Data Pre-Processing

In this process, the process of checking and overcoming missing values, data transformation, and dimensional reduction will be carried out. The stage of overcoming missing values in Training and Testing data is not done because there is no missing data. Then the data transformation for data that is categorical "object “into categorical”discrete". By using an encoder to transform the data. Here are the transformed features.

The function of data transformation is to facilitate the next stage, namely dimensional reduction. Features are used as many as 42 attributes and dimensional reduction will be done using Factor Analysis of Mixed Data (FAMD) and Autoencoder. For Autoencoder and T^2 using quantitative data as many as 37 because it cannot use categorical data. The selection of dimension reducers will be similar to each other from both FAMD and Autoencoder.

Table 3 Transformed Features

Feature's Name	Description
State	Indicates to the state and its dependent protocol, e.g. ACC, CLO, CON, ECO, ECR, FIN, INT, MAS, PAR, REQ, RST, TST, TXD, URH, URN, and (-) (if not used state)
Proto	Transaction protocol
Service	http, ftp, smtp, ssh, dns, ftp-data ,irc and (-) if not much used service

C. Dimensional Reduction of FAMD

By using PCA and MCA approach, obtained scree plot to determine the number of dimensions specified.

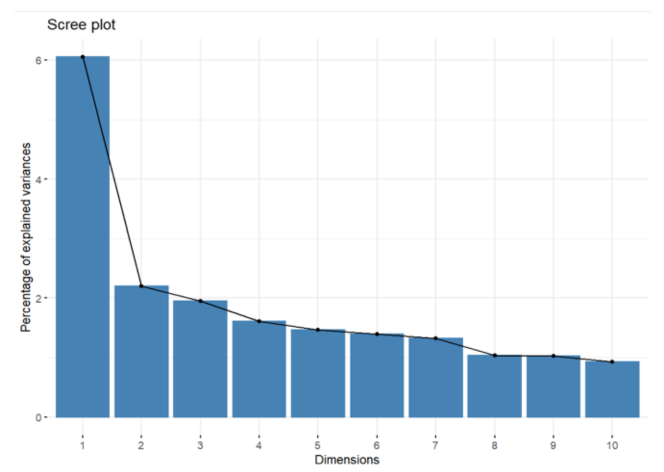

Figure 5 Scree Plot FAMD

Based on the picture above, from Dimension 1 to 10 has decreased and the sharpest decline in dimension 1 to 2 and gradually decreases slowly towards the 10th dimension. Then set with the number of PC 8 because it is enough to explain the existing variance and because it began to slope on PC to 8.

D. Dimensional Reduction of Autoencoder

In the autoencoder, the dimensions will be reduced to the same as in the FAMD, that is, to 8. Here is the configuration of the Autoencoder.

Table 4 Variance contribution of response variables for first PC

Layer (Type)	Output Shape	Parameter
Input Layer	(None, 37)	0
Encode1 (Dense)	(None, 30)	1140
Encode2 (Dense)	(None, 20)	620
Encode3 (Dropout)	(None, 8)	168
Decode1 (Dense)	(None, 20)	180
Decode2 (Dense)	(None, 30)	630
Decode3(Dense)	(None, 37)	1147

Before entering the autoencoder stage, the data will be transformed again using minmax scaler to reduce the range of data on each attribute. Hyperparameter used is by using Adam optimization and tanh activation function with the configuration in the table above. The running autoencoder process is carried out as many as 20 repetitions in order to get representative results with the original data.

E. Hotelling's T^2 Chart

In this process, the calculation graph T^2 for each method is T^2 without dimension reduction, FAMD with T^2, and Autoencoder with T^2.

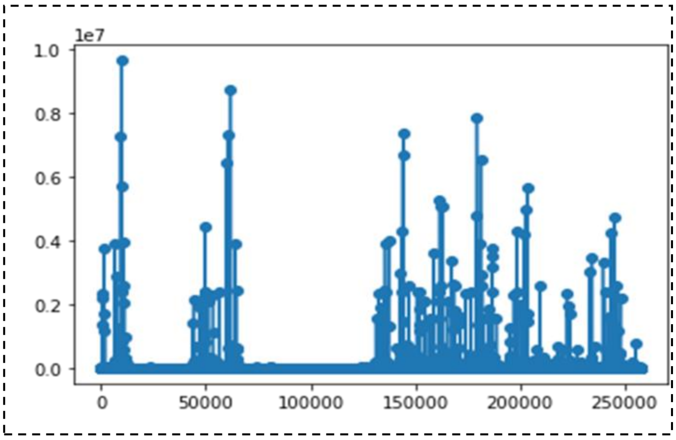

Figure 6 Conventional T^2

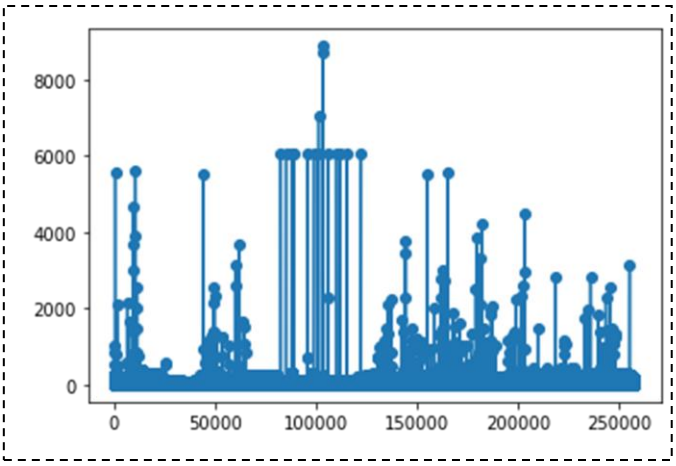

Figure 7 T^2 with FAMD

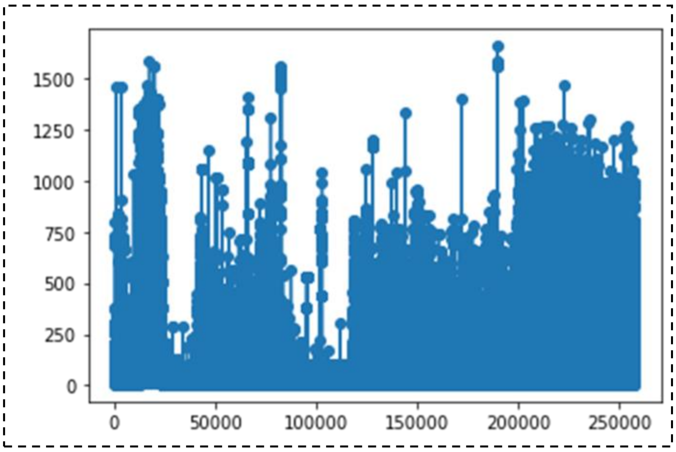

Figure 8 T^2 with Autoencoder

Based on the picture above, there are differences in each treatment. Where: UCL boundary determination using chi square Phase II control limit, if m is large enough (m>100):

$$UCL = \frac{p(m-1)}{m-p} F_{\alpha;p,m-p} \text{ or } UCL = \chi^2_{\alpha,p}$$

because the data is more than 100 and will be compared with alpha values of 0.2, 0.05, 0.00273, and 0.001 and obtained performance as follows.

Table 5 Performance of Conventional T²

	T^2		
	Hit Rate	FP Rate	FN Rate
α = 0.2	0.64	1	0
α = 0.05	0.64	1	0
α = 0.00273	0.64	1	0
α = 0.001	0.64	1	0

Table 6 Performance of Conventional T² with FAMD

	T^2		
	Hit Rate	FP Rate	FN Rate
α = 0.2	0.64	1	0
α = 0.05	0.64	1	0
α = 0.00273	0.64	1	0
α = 0.001	0.64	1	0

Table 7 Performance of T² with Autoencoder

	T^2		
	Hit Rate	FP Rate	FN Rate
α = 0.2	0.64	1	0
α = 0.05	0.64	1	0
α = 0.00273	0.64	1	0
α = 0.001	0.64	1	0

Based on the table above for the T^2 control chart there is no difference for the performance of both classifications in the various options, for the performance of Hit Rate of 0.64, FP Rate of 1 and FN Rate of 0. Which means that the accuracy to recognize normal activity and detected attacks by 64%, then the FP rate is the percentage of normal activity detected as an attack by 100% and the FN rate is the percentage of successful attacks that detected normal activity by 0%. The results show that the method is not suitable for this data. Then with the addition of dimensional reduction FAMD and Autoencoder produces performance in the table above.

For FAMD with T^2 and Autoencoder with T^2, the performance has increased both from Hit Rate, FP rate and FN rate when compared to regular T^2 performance. There are performance differences for FAMD and Autoencoder for a variety of α. The Hit rate for FAMD is highest at α = 0.2 i.e., 0.74 and for Autoencoder is highest at α = 0.001 i.e., 0.76. However, in terms of FN rate and overall FP value, FAMD is minimized when normal activity is detected as an attack and Autoencoder is minimized when an attack is detected as normal activity.

IV. CONCLUSION

In this paper, the analysis and the evaluation of the Hotelling's T^2 and Hotelling's T^2 with dimensional reduction for UNSW-NB15 data sets are discussed. The result shows that Hotelling's T^2 with dimensional reduction perform better result than conventional Hotelling's T^2. In Hotelling's T^2 with dimensional reduction way, T^2 control chart with FAMD performs better to detect normal activity as an intrusion than Autoencoder and the other way T^2 with Autoencoder performs better when intrusion is detected as normal activity. The highest Hit rate for the performance of the control chart with FAMD with T^2 obtained by 0.74 at a value of α = 0.2 and the highest for the performance of the control chart with Autoencoder T2 obtained by 0.76 at a value of α = 0.001. Both have different characteristics where for T^2 control chart with FAMD more minimize when normal activity is detected as an attack and for T^2 control chart with Autoencoder more minimize When attack is detected as normal activity. For future research, it is better to consider the MEWMA type of chart [22]-[24].

REFERENCES

[1] R. Heady, G. Luger, A. Maccabe, and M. Servilla, "The Architectur of a Network Level Intrusion Detection System," 1990, doi: https://doi.org/10.2172/425295.

[2] M. A. Aydın, A. H. Zaim, and K. G. Ceylan, "A hybrid intrusion detection system design for computer network security," Comput. Electr. Eng., vol. 35, no. 3, pp. 517–526, 2009, doi: 10.1016/j.compeleceng.2008.12.005.

[3] N. Kumar, "Advancements in Statistical Quality Control Standards," Int. J. Emerg. Technol. Innov. Res., vol. 6, no. 4, pp. 196–205, 2019.

[4] S. Bersimis, A. Sgora, and S. Psarakis, "The application of multivariate statistical process monitoring in non-industrial processes," Qual. Technol. Quant. Manag., vol. 3703, pp. 1–24, 2018, doi: 10.1080/16843703.2016.1226711.

[5] C. A. Lowry and D. C. Montgomery, "A review of multivariate control charts," IIE Trans. (Institute Ind. Eng., vol. 27, no. 6, pp. 800–810, 1995, doi: 10.1080/07408179508936797.

[6] N. Ye, X. Li, Q. Chen, S. M. Emran, and M. Xu, "Probabilistic techniques for intrusion detection based on computer audit data," IEEE Trans. Syst. Man, Cybern. Part ASystems Humans., vol. 31, no. 4, pp. 266–274, 2001, doi: 10.1109/3468.935043.

[7] R. Rastogi, Z. Khan, and M. H. Khan, "Network Anomalies Detection Using Statistical Technique : A Chi- Square approach," Int. J. Comput. Sci. Issues, vol. 9, pp. 515–522, 2012.

[8] D. Montgomery, Introduction to Statistical Quality Control, 6th Editio. New Jersey: John Wiley & Sons, Inc, 2009.

[9] C T. Kourti, "Application of latent variable methods to process control and multivariate statistical process control in industry," Int. J. Adapt. Control Signal Process., vol. 19, no. 4, pp. 213–246, 2005, doi: 10.1002/acs.859.

[10] R. L. Mason and J. C. Young, Multivariate Statistical Process Control with Industrial Applications. 2002.

[11] I. T. Jollife and J. Cadima, "Principal component analysis: A review and recent developments," Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., vol. 374, no. 20150202, 2016, doi: 10.1098/rsta.2015.0202.

[12] X. Ran, "Using the Dimension Reduction Method FAMD in the Data Pre-processing Step for Risk Prediction and for Unsupervised Clustering," University of Pittsburgh, 2019.

[13] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," ACM Int. Conf. Proceeding Ser., vol. 02-Decembe, pp. 4–11, 2014, doi: 10.1145/2689746.2689747.

[14] J. Pagès, "ANALYSE FACTORIELLE DE DONNÉES MIXTES : PRINCIPE ET EXEMPLE D'APPLICATION," Rev. Stat. appliquée, vol. 5, no. 4, pp. 93–111, 2002.

[15] M. A. Kramer, "Nonlinear Principal Component Analysis using Autoassociative Neural Networks," AIChE J., vol. 37, no. 2, pp. 233–243, 1991, doi: 10.1002/aic.690370209.

[16] G. Qu, S. Hariri, and M. Yousif, "Multivariate Statistical Analysis for Network Attacks Detection," 3rd ACS/IEEE Interenational Conf. Comput. Syst. Appl., pp. 9–14, 2005.

[17] N. Ye, D. Parmar, and C. M. Borror, "A Hybrid SPC Method with the Chi-Square Distance Monitoring Procedure for Large-scale, Complex Process Data," Qual. Reliab. Eng. Int., no. 22, pp. 393–402, 2006, doi: 10.1002/qre.717.

[18] A. A. Sivasamy and B. Sundan, "A Dynamic Intrusion Detection System Based on Multivariate Hotelling ' s T2 Statistics Approach for Network Environments," Sci. World J., vol. 2015, 2015, doi: 10.1155/2015/850153.

[19] H. Hotelling, "Multivariate Quality Control, Illustrated by The Air Testing of Sample Bombsights," Tech. Stat. Anal., pp. 111–184, 1947.

[20] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and Composing Robust Features with Denoising Autoencoders," Int. Conf. Mach. Learn., 2008.

[21] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. A. Manzagol, "Stacked denoising autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," J. Mach. Learn. Res., vol. 11, pp. 3371–3408, 2010.

[22] H. Khusna, M. Mashuri, Suhartono, D. D. Prastyo, and M. Ahsan, "Multioutput Least Square SVR Based Multivariate EWMA Control Chart," J. Phys. Conf. Ser., vol. 1028, no. 1, p. 12221, 2018.

[23] H. Khusna, M. Mashuri, S. Suhartono, D. D. Prastyo, and M. Ahsan, "Multioutput least square SVR based multivariate EWMA control chart: The performance evaluation and application," Cogent Eng., Oct. 2018, doi: 10.1080/23311916.2018.1531456.

[24] N. Sulistiawanti, M. Ahsan, and H. Khusna, "Multivariate Exponentially Weighted Moving Average (MEWMA) and Multivariate Exponentially Weighted Moving Variance (MEWMV) Chart Based on Residual XGBoost Regression for Monitoring Water Quality.," Eng. Lett., vol. 31, no. 3, 2023.