ASEAN Engineering Journal

Full Paper

NOVEL DIFFERENTIAL EVOLUTION FOR FEATURE SELECTION IN ANOMALY-BASED INTRUSION DETECTION

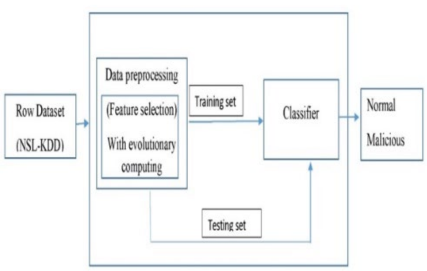
Aqeel Taha Saadoon[a], Azizul Azizan[a]*, Muhammad Yusof Mohd Noor[b], Yusnaidi Md Yusof[a], Mohd Fitri Mohd Yakub[c], Noraimi Shafie[a]

[a]Fakulti of Artificial Intelligence, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 UTM Kuala Lumpur, Malaysia
[b]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[c]Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 UTM Kuala Lumpur, Malaysia

Article history
Received
17 May 2024
Received in revised form
27 August 2024
Accepted
15 September 2024
Published online
31 May 2025

*Corresponding author
azizulazizan@utm.my

Graphical abstract



Abstract

In recent years many organizations and end users suffer from cyber-attacks or intrusions also known as zero-day attacks that aim for damaging resources or theft of data. A well-known tool for detecting such intrusions is anomaly-based Intrusion Detection System (IDS). IDS have integrated Evolutionary Computation (EC) algorithms as dimensionality reduction method to enhance the detection performance. A major limitation in anomaly-based IDS is the high rate of false alarms due to several reasons most importantly is the high volume of training and testing datasets. These high dimensionality datasets could contain irrelevant, duplicate, and redundant features that cause misclassifications and increase the false alarm rate. In this research a new variant of Differential Evolution (DE) algorithm called Differential Evolution – Convergence Extension (DE-CE) is proposed as part of the anomaly-based IDS for dimensionality reduction and feature selection. The new variant of DE adopts a new mutation strategy that ensures the continuously generating new solutions for the current population, thus ensures selecting the most relevant features from the dataset. The well-known NSL-KDD dataset is adopted for training and testing the proposed anomaly-based IDS. Evaluation is performed against previously proposed DE algorithms with different mutation strategies and PSO in terms of number of selected features, Accuracy, False Positive rate (FPR), recall, and precision for five different classifiers. The proposed DE-CE outperformed the classical DE and PSO algorithms in all performance evaluation metrics, where it achieved the highest accuracy of 99.4744% and lowest FPR of 0.3198%.

Keywords: Anomaly detection, feature selection, differential evolution, mutation strategy, NSL-KDD

1.0 INTRODUCTION

Unauthorized access to a system resources or data is usually called an intrusion and the designer of the intrusion is called an intruder. Intruders could be either internal intruders or external intruders. Internal intruders attempt to elevate their limited privileges by abusing it, whereas external intruders attempt to gain unauthorized access to system resources from outside the target network [1]. The usage of network control access using username and password with some rules and policies imposed by the administrator as security, to allow only authorized users is no longer sufficient with the emergence of multiple types of attacks that can impersonate authorized users.

In addition, the attacks resulting from an authorized personal is because of misuse or tampering within the system [2,3]. While minimizing or eliminating security threats is not always possible, an Intrusion Detection System (IDS) is required to identify known and unknown assaults and notify system administrators [4,5].

IDSs are generically divided into two types: signature-based IDS and anomaly-based IDS [6-8]. Signature-based IDS or misuse-based, uncovers common attacks behaviors that are stored in the IDS database [1]. In this scenario, the IDS produces an attack alert to inform the network monitoring administration if a new behavior correlates with the database's existing identities by examining digital signatures for repetitions perceived from earlier attack signals. Unexpected assaults with no signatures in the IDS database cannot be detected by pattern matching techniques, which is their fundamental shortcoming. Successful open-source IDS products such as Snort, Zeek, Suricata and CalmAV/CalmWin are examples of signature-base detection Seng et al., [9].

Anomaly-based IDS, on the other hand, keep records of usual activity patterns in the IDS database. The anomaly-based IDS triggers an alarm to notify network monitoring administration of newly detected threats when it detects any divergence from regular patterns. In these solutions, all network actions are carefully monitored and reviewed. The fundamental advantage of anomaly-based identification systems is their ability to uncover previously undiscovered dangers or as it is known as zero-day attacks. In these systems, any unusual patterns will be recognized as possible assaults, even if they are not. As a consequence, false positives are widespread [10,11].

One major issue in the field of IDS and specifically in anomaly-based IDS is the high dimension of the training-set and testing-set. The high volume of the datasets has a critical impact on the performance of the anomaly-based IDS as a major cause for high false alarms and low accuracy. There is a lack of the anomaly-based IDS in the industry due to the lack of datasets, weak evaluation methods, reproducibility, and comparability [9]. The issue of the high dimensional data could be treated by feature selection method by selecting the most relevant features to the IDS and in turn enhance its accuracy and false alarm rate.

Even though many methods of feature selection are introduced in the anomaly-based IDS using evolutionary algorithms, however it does not necessarily improve the performance of the anomaly detection. Poor selection of features affects the accuracy of the system's classification as it may increase the false positive rates. Moreover, some of these methods increase the computational cost. Hence efficient feature selection methods that can identify relevant and meaningful features for classifier are always analyzed [12-14]. The main contribution in this article is to fully analyze different mutation strategies for differential evolution algorithms as a feature selection method for the anomaly-based IDS. Moreover, this paper also introduces a novel mutation strategy to select the most relevant features using the NSL-KDD dataset to train and test the IDS.

2.0 RELATED WORK

Several algorithms have been proposed for feature selection within anomaly-based IDS. The focus in this article is feature selection methods that utilize the differential evolution algorithm. DE algorithm has been used for feature selection due to its simplicity and promising capabilities as it has less control parameters than many algorithms and therefore computationally affordable. Table 1 shows a summary of the related work.

Al-Yaseen *et al*., [15] used wrapper-based feature selection method by employing the state-of-the-art DE algorithm and used the Extreme Learning Machine-ELM as a classifier to evaluate the selected features. The aim of the proposed feature selection method is to enhance the IDS's accuracy and to reduce the false alarm rate and processing time. The proposed method selects 9 features out of 41 when using the NSL-KDD dataset. Liu *et al*., [16] proposed the use of the state of the art the Self Adaptive Differential Evolution for feature selection in anomaly-based IDS. The self-adaptive DE algorithm automatically controls the values of the control parameters for the DE algorithm which helps avoiding manually tuning these parameters.

In Kathirvel *et al*., [17] the authors addressed the issues of wireless sensor networks (WSN) in particular the high processing energy and storage capabilities, so they proposed an enhanced intrusion detection and response system as they employed the Multi-objective DE algorithm to compute the trust values of the sensor nodes in wireless sensor nodes depending on the energy level of the sensor nodes. A research conducted by Fatani et al., [18] proposed a feature selection and extraction method for anomaly detection by using the Convolutional neural network (CNN) algorithm and the Transient Search Optimization (TSO) algorithm, the proposed algorithm extracts the features by using the operators of the DE algorithm.

Lv et al., [19] used the Kernel Principal Component Analysis (KPCA) algorithm for feature selection in misuse detection. The authors employed a combination of the Gravitational Search Algorithm (GSA) and the DE algorithm to optimize the parameters of the proposed feature selection model. DE algorithm was also used by Shojafar et al., [20] as the authors proposed a clustering algorithm as part of the IDS, the clustering algorithm is based on the concept of coherence and separation. The proposed algorithm is optimized by the Ant Colony Optimization (ABC), Particle Swarm Pptimization (PSO) and DE algorithms.

The authors Hodashinsky *et al*., [21] presented an intrusion detection system based on fuzzy rule-based classifier. The authors addressed the issue of optimizing the parameters of the features selection and for optimizing the parameters of the fuzzy classifier the DE algorithm is used, while the Binary Harmonic Search (BHS) algorithm is used for feature selection. Xue et al., [22] used the state of art Self-adaptive DE (SaDE) algorithm for feature selection with four Candidate Solution Generation Strategy (CSGS). The K-Nearest Neighbour (KNN) is used for evaluation of the selected features. The work conducted in Aburomman et al., [23] compares several methods for creating a multiclass, support vector machine (SVM) classifier from a set of binary SVM classifiers. The authors used the DE optimization algorithm by finding better solutions from the search space, these solutions are reflected through the average fitness value of the population.

Table 1 Summary of related work

Ref.	DE impact on IDS with accuracy and FPR	DE description
[15]	Classical DE used for feature selection and ELM for evaluation. Accuracy achieved 87.70% and FPR 0.7%.	Classical DE
[16]	Self-adaptive DE (SADE) algorithm for feature selection in IDS. SADE automatically adapts the control parameters of DE algorithm maintaining same mutation strategy and cross over.	SaDE
[17]	Multi-objective DE algorithm to compute the trust values of the sensor nodes in WSN. The FPR and accuracy achieved varies according to the number of wireless nodes (50-200) and number of attacks.	Classical DE
[18]	TSO algorithm by employing the classical DE's operators for feature selection. The FPR achieved was 0.586% and accuracy of 77.381%	Classical DE
[19]	PCA is used for dimensionality reduction and the DE is used to optimize the parameters of the PCA. The accuracy achieved is 96.69% and FPR 1.10%.	Classical DE
[20]	Classical DE algorithm is compared with other classical optimization algorithms to optimize the proposed clustering algorithm for comparison reasons.	Classical DE
[21]	DE used to optimize the parameters of the fuzzy rule-based classifier. The FPR scored is 1.58% with accuracy of 97.84%.	Classical DE
[22]	SaDE is adopted for feature selection in IDS. SaDE automatically adapts the control parameters with the same mutation and cross over strategy. The accuracy achieved is 98.71%.	SaDE
[23]	Classical DE algorithm was adopted to tune the hyper parameters of the classifier. The accuracy achieved is 82.62%.	Classical DE
[30]	DE-ME a variant of a new mutation is proposed for IDS with accuracy of 99.56% and the FPR is 0.146%.	DE-ME

3.0 METHODOLOGY

3.1 Differential Evolution Algorithm

DE is a population-based evolutionary strategy which has been successfully used in various fields to improve feature selection to achieve global optimization. DE's effectiveness in tackling complicated challenges is largely driven by the choice of a suitable mutation strategy and the parameter selections; where determining suitable control parameters for the DE is the main task to achieve optimality. Figure 1 shows a simple flow chart of classical DE algorithm.

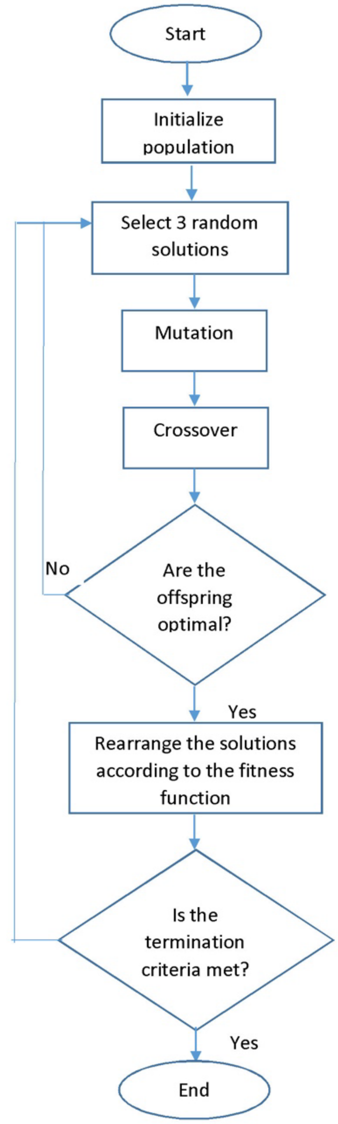


Figure 1 Flow diagram of the DE algorithm

The three primary operations of DE are mutation, crossover, and selecting processes to select meaningful features from a high dimensional dataset that enhances the performance of the classifier. Through loop repetitions of mutation, crossover, and selecting processes, the population is always evolving in the direction of optimal value. While this approach is straightforward and generally productive, it also has some considerable drawbacks. DE is a stochastic optimization strategy that is simpler and more robust to failure.

However, the search behavior of the typical DE algorithm is usually imbalanced [24,25]. In other words, while the DE method's exploration capability is good, its exploitation capability is inadequate, resulting in a slow convergence rate Noman *et al.,* [26]. Additionally, the DE approach does not guarantee convergence to the global solution and is prone to premature convergence Knobloch *et al.*, [27], as well as halting in poorer solutions Neri *et al.*, [28].

3.1.1 Population Initialization

To produce an optimal data population, DE splits the population into several individuals *Wang et al.*, [29], which can be formulated as,

$$X = [x_1, x_2, x_3, \cdots x_N] \tag{1}$$

As a result, each entity in (1) represents a potential solution,

$$x_i(g) = x_{id}(g) \tag{2}$$

Where d = 1, 2, D, this is referring to the size of the solution. Furthermore, $x_{id}(g)$ is the d-th element of the population's i-th developing candidate of the g-th generation. The last expression, (2), must be initialized before the evolution procedure can begin.

$$x_{id}(0) = x_{L,d} + \psi \cdot (x_{H,d} - x_{L,d}) \tag{3}$$

However, x_L and x_H are the lowest and highest quantities, respectively, whereas ψ is a random number from the range [0,1]. After the split, the following step is the mutation-operation. Diverse individuals are produced from each $x_{id}(g)$ in the population using the differential evolution algorithm hence the word evolution in DE.

3.1.2 Mutation

There are several options for the mutation operations Wang et al., [29]. As it can be implemented using random, best selection or others as shown in the equations (4) to (7). There are several options for the mutation operations Wang et al., [29]. As it can be implemented using random, best selection or others as shown in the equations (4) to (7). The mutation outcome for random selection using single differential factors is:

$$y_{id}(g) = x_{rnd1,d}(g) + F \cdot \left(x_{rnd2,d}(g) - x_{rnd3,d}(g)\right) \tag{4}$$

The equation above representing classical DE1, can be re-written in a simple form as DE/rnd/1, where DE stands for "Differential Evolving," and the number "1" relates to the quantity of differential factors. rnd means "random selection," and it refers to the process by which an individual from a bank of populated individuals is selected at random to serve as the basis for a mutation operation.

For a combination of best selection and random selection the mutation outcome for other three classical DE (DE2, DE3 and DE4) are listed as follows:

$$y_{id}(g) = x_{best}(g) + F \cdot \left(x_{rnd1,d}(g) - x_{rnd2,d}(g)\right) \tag{5}$$

$$\begin{aligned} y_{id}(g) = x_{rnd1,d}(g) &+ F \cdot \left(x_{rnd2,d}(g) - x_{rnd3,d}(g)\right) \\ &+ F \cdot \left(x_{rnd4,d}(g) - x_{rnd5,d}(g)\right) \end{aligned} \tag{6}$$

$$\begin{aligned} y_{id}(g) = x_{best}(g) &+ F \cdot \left(x_{rnd1,d}(g) - x_{rnd2,d}(g)\right) \\ &+ F \cdot \left(x_{rnd3,d}(g) - x_{rnd4,d}(g)\right) \end{aligned} \tag{7}$$

DE/best/1, DE/rnd/2, and DE/best/2 are concise descriptions of expressions (4) to (7) that may be represented as DE/option/d, accordingly, where option represents the selection operation from the population (rand or best or other criteria), and d stands for the difference factor.

A recent proposed DE-ME algorithm as in [30] uses the current individual concept with the signum function in their mutation strategy as in (8)

$$y_{i,g} = sgn(x_{best,g}) * \left(\frac{x_{best,g} - x_{i,g}}{x_{best,g} + x_{i,g}}\right) + F \times \left(x_{rnd1,g} - x_{rnd2,g}\right) \tag{8}$$

3.1.3 Crossover

After mutation, cross-processing, or crossover operations, is implemented to increase the diversity of the population. Consequently, a new individual $y_{id}(g)$ is assigned to compete with the created individual $x_{id}(g)$ in accordance with the crossover probability Wang *et al.*, [29].

$$z_{id}(g) = \begin{cases} y_{id}(g) & rnd < \beta \text{ or } d = r \\ x_{id}(g) & otherwise \end{cases} \tag{9}$$

$$z_{id}(g) = \begin{cases} z_{id}(g) & f\left(z_{id}(g) < f\left(x_{id}(g)\right)\right) \\ x_{id}(g) & otherwise \end{cases} \tag{10}$$

In this case, r is a random number selected from the range 1-D. The fitness function, $f\left(x_{id}(g)\right)$can be also used to determine which of the two $x_{id}(g)$ or $y_{id}(g)$ options is the best, which is also known as the objective function.

3.1.4 Proposed Mutation Strategy

Based on the prior state of the art study of anomaly-based IDS and DE algorithms, there have been a variety of approaches for implementing feature selection IDSs for anomaly detection. The differential evolution method that was reported in the listed Table 1 for feature selection is deemed insufficient. Herein, this research considers the other common mutation strategies for DE algorithm to be evaluated for feature selection for IDSs considering its excellent qualities.

Furthermore, it was discovered in most literatures the aims are to enhance the implementations of the DE algorithm for broad optimization issues rather than specific optimization issues such as anomaly-based IDS. In other words, the DE algorithm was not fully evaluated in the field of IDS. To put it in another way, an enhanced version of the DE algorithm has not been deeply considered when dealing with high dimensional data in anomaly-based IDS as a feature selection method.

From the available mutation methods described in (4) to (7), it is possible to construct a novel mutation strategy that may be employed in the feature selection procedure for IDS to improve further the weakness of the current DE namely premature convergence. In this research, a mutation technique termed DE-CE (Differential Evolution – Convergence Extension) algorithm were investigated to get better feature selection outcomes. The proposed mutation strategy is as follow:

$$y_{i,g} = \frac{x_{rnd1,g} - x_{i,g}}{x_{rnd2,g} + x_{i,g}} \times x_{best,g} + F \times \left(x_{rnd2,g} - x_{rnd3,g}\right) \quad (11)$$

where:

- *F* is the scaling control parameter,
- $x_{best,g}$ is the best *g-th* generation individual,
- $x_{i,g}$ is the current *g-th* individual,
- $x_{rnd1,g}$, $x_{rnd2,g}$ and $x_{rnd3,g}$ are three randomly selected individuals of the *g-th* generation.

The first term in (11), $\frac{x_{rnd1,g} - x_{i,g}}{x_{rnd2,g} + x_{i,g}} \times x_{best,g}$, ensures the continuously changing of the iteration, thus, avoiding premature problem which is the main problem of the DE algorithm, the proposed uses a combination of best, current and random individuals to ensure the selection of the best solutions in the population and therefore guarantees the selection of the most relevant features of the dataset . As a pre-processing step before performing feature selection, the NSL-KDD dataset is transformed through several preprocessing steps as labeling the data, categorical encoding, data normalization with a Min-Max scaling function, and data imputation for missing data.

The pseudo code for the proposed DE algorithm is as follow:

Input: population size = 20, number of iterations = 20, set F = 0.5, set *CR* = 0.5, number of samples for FS = 80000.
Output: best features.
Population initialization
While (no. of iteration < *max*) do
  Select 3 random vectors ($x_{rnd1,g}$, $x_{rnd2,g}$ and $x_{rnd3,g}$);
  Select best vector $x_{best,g}$;
  Select current vector $x_{i,g}$;
  Perform Mutation as in (4)-(8) and (11);
  Perform crossover and selection as in (9);
  Update the population with the new vector;
End while

4.0 RESULTS AND DISCUSSION

The proposed feature selection method is built on the DE algorithm with a new mutation strategy in (11). This feature selection method implemented is a wrapper-based feature selection, where the KNN algorithm is chosen to evaluate the lected features. The proposed feature selection method is compared with four classical mutation strategies as in equations (4) to (7), plus a recent feature selection method-based DE [30] namely DE with maturity extension DE-ME in equation (8) and PSO.

The performance comparisons are in terms of number of features selected, accuracy, false positive rate, precision, recall, and f-measure which can be expressed as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$FPR = \frac{FP}{FP + TN} \quad (13)$$

$$Precision = \frac{TP}{TP + FP} \quad (14)$$

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

$$F = \frac{(Precision * Recall)}{(Precision + Recall)} \quad (16)$$

The proposed DE-CE algorithm selects 10 features out of 42 depending on the best fitness values of the new generated offspring. Classical DE1, classical DE2, classical DE3 and classical DE4 select 7, 5, 8 and 7 features respectively. While the PSO algorithm selects 15 features, and the DE-ME algorithm selects 10 features.

The effect of the selected features is reflected on the performance evaluation metrics of the feature selection methods. Five classifiers are used for the anomaly detection performance evaluation which are: SVM with polynomial kernel, SVM with Gaussian kernel, Decision Tree, Ensemble of Decision Trees, and Naïve Bayes.

Figures 2-6 shows that the proposed DE-CE has better performance in all evaluation metrics with all classifiers than the classical DE algorithms and PSO algorithm, however the DE-ME algorithm has slightly better performance than the proposed around (0.01-0.1%) in some metrics with the most stable classifiers as in decision tree and ensemble of trees classifiers. when it comes to the less stable classifiers as in SVM with polynomial classifier the proposed DE-CE highly outperforms the DE-ME algorithm around 9-11% in all evaluation metrics.
In details, the proposed DE-CE achieved 99.4744% as the highest accuracy and the lowest 96.1355%.and PSO achieved accuracy in the range of 79.9512% - 91.7802%. Another important observation lower accuracy was achieved comparatively for the four classical DEs and the PSO algorithm compared to the DE-CE and DE-ME.

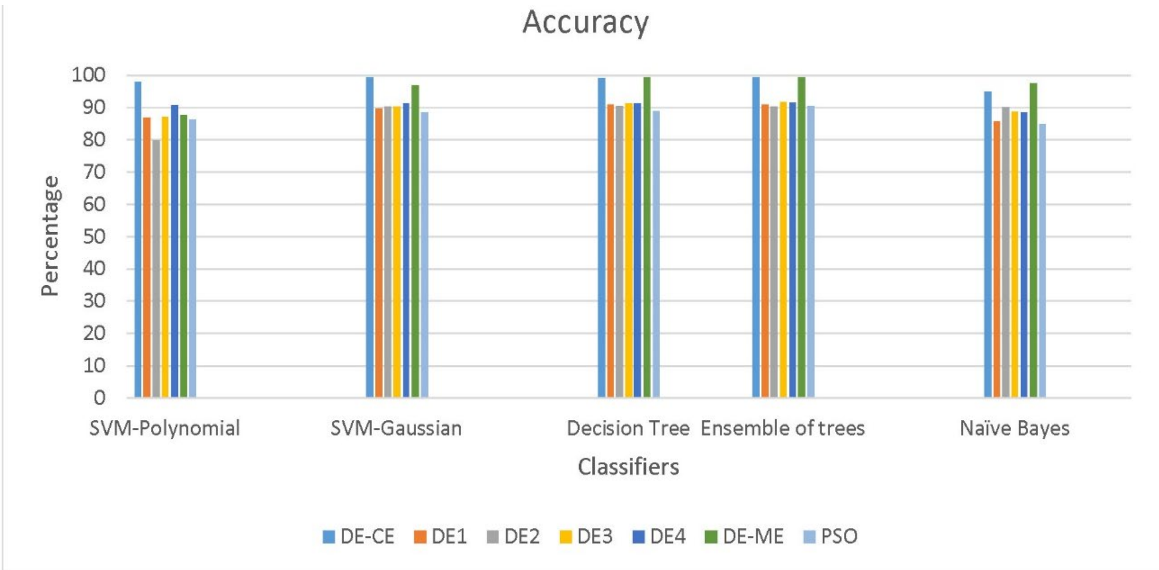


Figure 2 Accuracy for DE-CE and others

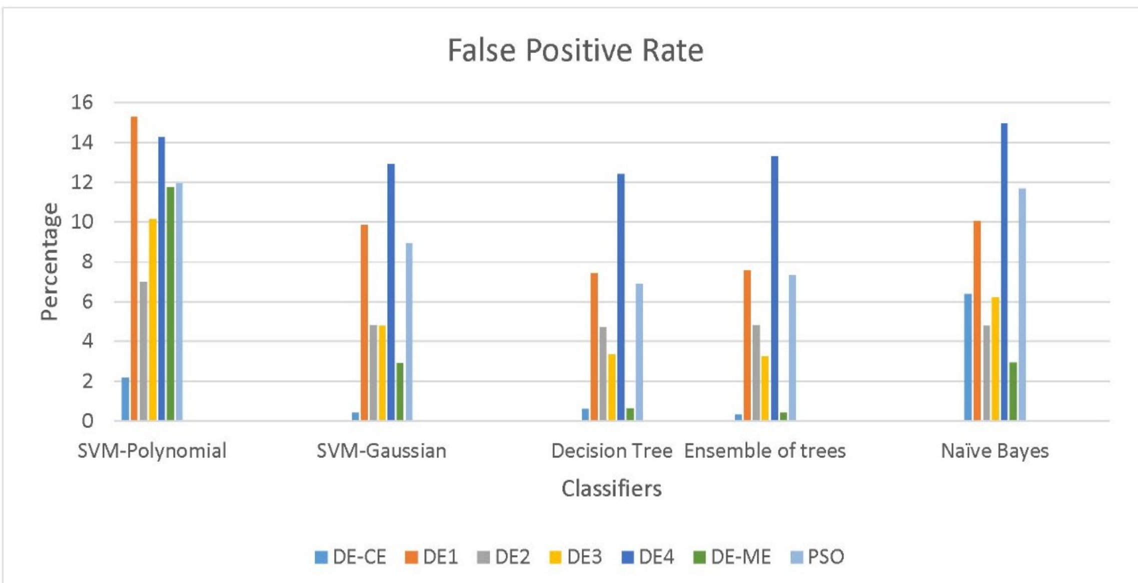


Figure 3 FPR for DE-CE and others

The proposed DE-CE achieved better False Positive Rate (FPR) than all feature selection methods as it scored 0.3198% and the highest was 5.6207% as for the DE-ME algorithm the lowest FPR scored was 0.41342% and the highest was 11.7546%, the rest achieved an FPR in the range 3.24693% - 15.2963%.

The proposed DE-CE and DE-ME also achieved better Recall, Precision and F-measure than all feature selection methods in general. The results also show that our proposed DE-CE also performs better than DE-ME for Recall, Precision and F-Measure especially for the two SVM classifiers

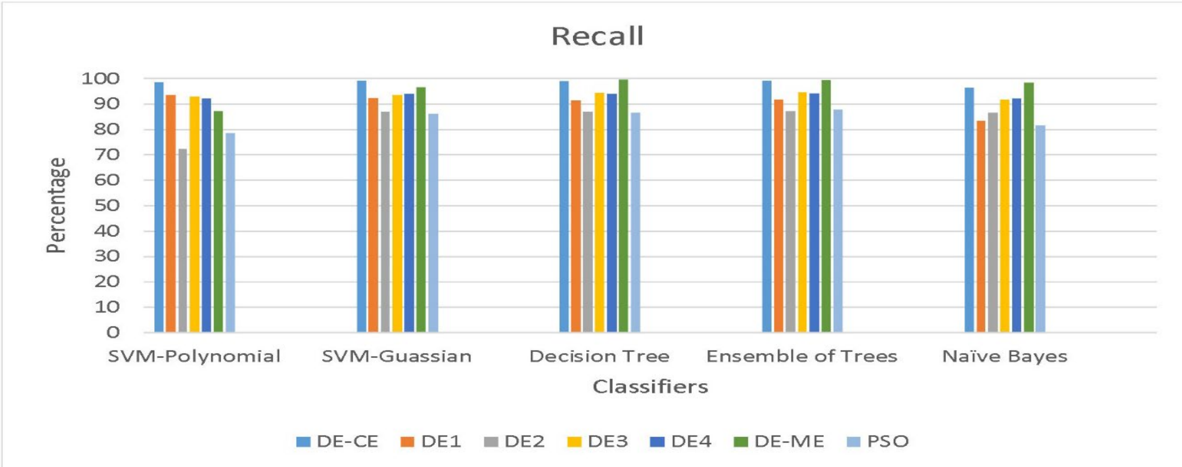


Figure 4 Recall for DE-CE and others

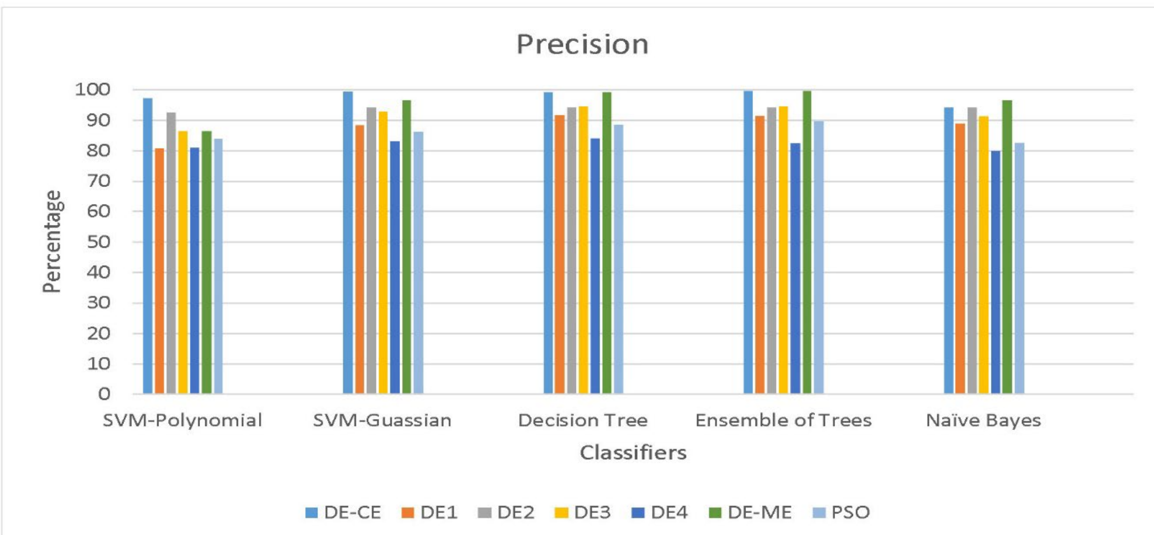


Figure 5 Precision for DE-CE and others

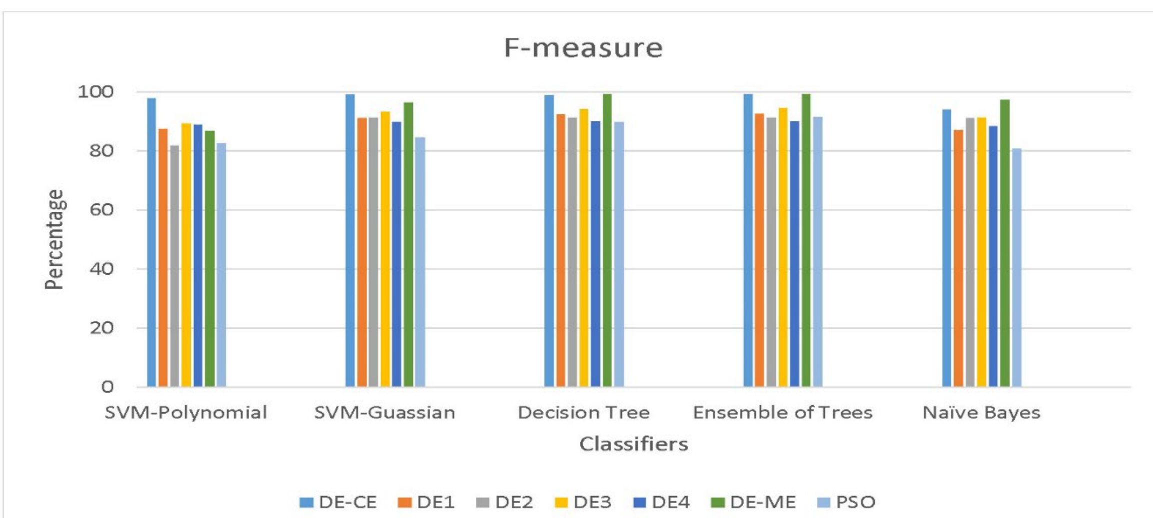


Figure 6 F-measure for DE-CE and others

Table 2 summarizes results for both accuracy and FPR for the proposed DE-CE and the methods mentioned in the literature.

Table 2 summary of accuracy and FPR

Method	Accuracy	FPR
DE-ELM [15]	87.70%	0.7%
TSODE [18]	77.381%	0.586%
KPCA-DEGSA-HKELM [19]	96.69%	1.10%
Classical DE[21]	97.84%	1.58%
SaDE [22]	98.71%	------
WOAR-SVM [23]	82.62%	------
DE-ME [30]	99.56%	0.41342%
DE-CE proposed	99.4744%	0.319%

5.0 CONCLUSION

This paper addresses the issues of improving the accuracy and reducing the high rate of false positive rate in anomaly detection employing DE for feature selection. A new mutation strategy for DE algorithm for feature selection have been proposed focusing on different variants of mutation strategies for DE algorithm to improve the performance of anomaly detection by reducing the features.

DE has been evaluated using only the classical and SaDE variant for feature selection in the field of anomaly detection, hence warranting this research. The proposed feature selection DE-CE with new mutation strategy has enhanced the performance of the anomaly detection classifiers in terms of accuracy and false positive rate as it shows that the DE-CE algorithm's mutation strategy is superior to other classical DEs and the classical PSO algorithm.

Acknowledgement

This research is fully supported by matching grant MG2 – 9.1 with cost center no. Q.K130000.3056.04M27. The authors fully acknowledged Ministry of Higher Education (MOHE) and Universiti Teknologi Malaysia for the approved fund which makes this important research viable and effective.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper

References

[1] Hubballi, Neminath, and Vinoth Suryanarayanan. 2014. "False alarm minimization techniques in signature-based intrusion detection systems: A survey." *Computer Communications* 49: 1-17. DOI: https://doi.org/10.1016/j.comcom.2014.04.012

[2] Pawar, Mohan V., and J. Anuradha. 2015. "Network security and types of attacks in network." *Procedia Computer Science* 48: 503-506. DOI: https://doi.org/10.1016/j.procs.2015.04.126

[3] Safari, Mohammad, Elham Parvinnia, and Alireza Keshavarz Haddad. 2021."Industrial intrusion detection based on the behavior of rotating machine." International *Journal of Critical Infrastructure Protection.* 34: 100424. DOI: https://doi.org/10.1016/j.ijcip.2021.100424

[4] Rassam, Murad A., M. A. Maarof, and Anazida Zainal. 2012. "A survey of intrusion detection schemes in wireless sensor networks." *American Journal of Applied Sciences.* 9(10): 1636. DOI: https://doi.org/10.3844/ajassp.2012.1636.1652

[5] Butun, Ismail, Salvatore D. Morgera, and Ravi Sankar. 2013 "A survey of intrusion detection systems in wireless sensor networks." *IEEE communications surveys & tutorials* 16(1): 266-282. http://dx.doi.org/10.1109/SURV.2013.050113.00191

[6] Sengupta, Nandita, and Jaya Sil. 2020. Intrusion Detection: A Data Mining Approach. *Springer Nature,* DOI: http://dx.doi.org/10.1007/978-981-15-2716-6_1

[7] Meira, Jorge, Rui Andrade, Isabel Praça, João Carneiro, Verónica Bolón-Canedo, Amparo Alonso-Betanzos, and Goreti Marreiros. 2020."Performance evaluation of unsupervised techniques in cyber-attack anomaly detection." *Journal of Ambient Intelligence and Humanized Computing.* 4477-4489. DOI: https://doi.org/10.1007/s12652-019-01417-9

[8] Zoppi, Tommaso, Andrea Ceccarelli, Tommaso Capecchi, and Andrea Bondavalli. 2021. "Unsupervised anomaly detectors to detect intrusions in the current threat landscape." *ACM/IMS Transactions on Data Science* 2(2): 1-26. DOI: https://doi.org/10.1145/3441140

[9] Seng, S., J. Garcia-Alfaro, and Y. Laarouchi. 2021. "Why Anomaly-Based Intrusion Detection Systems Have Not Yet Conquered the Industrial Market?." In *International Symposium on Foundations and Practice of Security,* 341-354. Cham: Springer International Publishing, DOI: https://doi.org/10.1007/978-3-031-08147-7_23

[10] Villalba, Luis Javier García, AL Sandoval Orozco, and J. Maestre Vidal. 2015."Malware detection system by payload analysis of network traffic." *IEEE Latin America Transactions* 13(3): 850-855. DOI: https://doi.org/10.1109/TLA.2015.7069114

[11] Ni, Xiejun, Daojing He, and Farooq Ahmad. 2016."Practical network anomaly detection using data mining techniques." *VFAST Transactions on Software Engineering*. 9(2): 1-6. DOI: https://doi.org/10.21015/vtse.v9i2.403

[12] Ghazy, Rania A., El-Sayed M. El-Rabaie, Moawad I. Dessouky, Nawal A. El-Fishawy, and Fathi E. Abd El-Samie. 2018. "Efficient techniques for attack detection using different features selection algorithms and classifiers." *Wireless Personal Communications*. 100: 1689-1706. DOI: https://doi.org/10.1007/s11277-018-5662-0

[13] Ravi Kiran Varma, P., V. Valli Kumari, and S. Srinivas Kumar. 2018. "A survey of feature selection techniques in intrusion detection system: A soft computing perspective." In *Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2017.* 785-793. Springer Singapore, DOI: https://doi.org/10.1007/978-981-10-7871-2_75

[14] Thaseen, Ikram Sumaiya, and Cherukuri Aswani Kumar. 2017 "Intrusion detection model using fusion of chi-square feature selection and multi class SVM." *Journal of King Saud University-Computer and Information Sciences*. 29(4): 462-472. DOI: https://doi.org/10.1016/j.jksuci.2015.12.004

[15] Al-Yaseen, Wathiq Laftah, Ali Kadhum Idrees, and Faezah Hamad Almasoudy. 2022. "Wrapper feature selection method based differential evolution and extreme learning machine for intrusion detection system." *Pattern Recognition*. 132: 108912. DOI: https://doi.org/10.1016/j.patcog.2022.108912

[16] Liu, Qianqian, Xiaoyan Zhang, Qiaozhi Hua, Zheng Wen, and Haipeng Li. 2022. "Adaptive Differential Evolution Algorithm with Simulated Annealing for Security of IoT Ecosystems." Wireless Communications and Mobile Computing. 2022. DOI: https://doi.org/10.1155/2022/6951849

[17] Kathirvel, Ayyaswamy, Muthusamy Subramaniam, S. Navaneethan, and C. Sabarinath. 2021. "Improved IDR Response System for Sensor Network." *Journal of Web Engineering.* 53-88. DOI: https://doi.org/10.13052/jwe1540-9589.2013

[18] Fatani, Abdulaziz, Mohamed Abd Elaziz, Abdelghani Dahou, Mohammed AA Al-Qaness, and Songfeng Lu. 2021. "IoT intrusion detection system using deep learning and enhanced transient search optimization." *IEEE Access.* 9: 123448-123464. DOI: http://doi.org/10.1109/ACCESS.2021.3109081

[19] Lv, Lu, Wenhai Wang, Zeyin Zhang, and Xinggao Liu. 2020. "A novel intrusion detection system based on an optimal hybrid kernel extreme learning machine." *Knowledge-based systems* 195: 105648. DOI: https://doi.org/10.1016/j.knosys.2020.105648

[20] Shojafar, Mohammad, Rahim Taheri, Zahra Pooranian, Reza Javidan, Ali Miri, and Yaser Jararweh. 2019. "Automatic clustering of attacks in intrusion detection systems." In *2019 IEEE/ACS 16th International*

Conference on Computer Systems and Applications (AICCSA), 1-8. IEEE. DOI: http://doi.org/10.1109/AICCSA47632.2019.9035238

[21] Hodashinsky, I. A., and M. A. Mech. 2018 "Constructing a fuzzy network intrusion classifier based on differential evolution and harmonic search." *International Journal of Computer Networks & Communications (IJCNC)*. 10: 85-91. DOI: https://doi.org/10.5121/ijcnc.2018.10208

[22] Xue, Yu, Weiwei Jia, Xuejian Zhao, and Wei Pang. 2018. "An evolutionary computation based feature selection method for intrusion detection." *Security and Communication Networks*. 2018. DOI: https://doi.org/10.1155/2018/2492956

[23] Aburomman, Abdulla Amin, and Mamun Bin Ibne Reaz. 2017. "A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems." *Information Sciences* 414: 225-246. DOI: https://doi.org/10.1016/j.ins.2017.06.007

[24] Storn, Rainer, and Kenneth Price. 1997. "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces." *Journal of global optimization.* 11: 341-359. DOI: https://doi.org/10.1023/A:1008202821328

[25] Vesterstrom, Jakob, and Rene Thomsen. 2004. "A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems." In Proceedings of the 2004 congress on evolutionary computation (IEEE Cat. No. 04TH8753), 2: 1980-1987. IEEE, DOI: http://doi.org/10.1109/CEC.2004.1331139

[26] Noman, Nasimul, and Hitoshi Iba. 2008. "Accelerating differential evolution using an adaptive local search." *IEEE Transactions on evolutionary Computation.* 12(1): 107-125. DOI: http://doi.org/10.1109/TEVC.2007.895272

[27] Knobloch, Roman, Jaroslav Mlýnek, and Radek Srb. 2017 "The classic differential evolution algorithm and its convergence properties." *Applications of Mathemati*cs. 62: 197-208. DOI: https://doi.org/10.21136/AM.2017.0274-16

[28] Neri, Ferrante, and Ville Tirronen. 2010. "Recent advances in differential evolution: a survey and experimental analysis." *Artificial intelligence review* 33: 61-106. DOI: https://doi.org/10.1007/s10462-009-9137-2

[29] Wang, Tiejun, Kaijun Wu, Tiaotiao Du, and Xiaochun Cheng. 2020. "Adaptive dynamic disturbance strategy for differential evolution algorithm." *Applied Science*s. 10(6): 1972. DOI: http://dx.doi.org/10.3390/app10061972

[30] Faris, M., Mahmud, M. N., Salleh, M. F. M., & Alsharaa, B. 2023. "A differential evolution-based algorithm with maturity extension for feature selection in intrusion detection system." *Alexandria Engineering Journal*, 81: 178-192. DOI: https://doi.org/10.1016/j.aej.2023.09.032