ICIC Express Letters
Volume **17**, Number **12**, December **2023**

ICIC International ©2023 ISSN 1881-803X
pp. 1395–1403

DETECTING IOT ANOMALY USING ROUGH SET AND DENSITY BASED SUBSPACE CLUSTERING

Fokrul Alom Mazarbhuiya

School of Fundamental and Applied Sciences
Assam Don Bosco University
Tapesia Garden, Kamarkuchi, Sonapur, Guwahati, Assam 782402, India
fokrul.mazarbhuiya@dbuniversity.ac.in

Received December 2022; accepted March 2023

Abstract. *IoT technology has significantly contributed in the improvement of quality of life by facilitating various real-life smart applications. IoT consists of large number of interconnecting digital devices which generates the large amount of data and makes computations. However, IoT domain often encounters the issue of anomalies, non-integrity, illegitimacy, etc. Finding anomaly in IoT is a challenging task as the data generated in these domains are often high-dimensional. Clustering and rough set theory have been tried exclusively in many cases with limited success. In this article, we propose a hybrid approach consisting of rough set theory and clustering techniques for IoT anomaly detection. First of all, the rough set theory is employed for the dimension reduction and then a density-based clustering approach is used in the subspace for the anomaly detection. The algorithm automatically supplies anomalies as noise. The efficacy of the method is established by complexity analysis and an experimental study using a real life dataset.*
Keywords: Anomaly detection, Information system, High-dimensional data, Inclusion dominance relation, CORE of attribute set

1. **Introduction.** Internet of Things (IoT) refers to the integration of interconnecting devices that has the ability to capture and transmit data [1]. It is usually viewed as a global set-up which facilitates the connectivity between the physical and digital world based on previous and current set-up [1]. As the data are generating and transmitting continuously, IoT devices are susceptible to external or internal attack in the form of anomalies or intrusions. Also IoT devices depend on the connectivity to other IoT devices and the Internet, making them possible targets of malevolent users. So, the devices must take precautions to prevent and identify intrusions. Due to the rise of anomalous activities, detecting anomalies in IoT devices has emerged as an increasingly crucial component of cybersecurity [2] and many researchers are getting involved in this area.

An anomaly detection is considered as the process of finding an outlier or trend in a dataset. It has been used in a variety of real-time applications, including industry damage detection, credit card fraud detection, sensor node failure detection, abnormal health monitoring, and network intrusion detection. Anomaly detection is usually hindered by high dimensionality. As the number of attributes or features grows, so does the amount of data required to generalize effectively which results in data sparsity. The data sparsity is caused by extra variables or a high amount of noise from several irrelevant attributes, which conceal the actual anomalies. The "curse of dimensionality" [3,4] is a well-known term used for this problem. Many anomaly detection techniques addressing high dimensionality, such as distance-based, density-based, and clustering-based techniques, fail to retain the effectiveness of conventional approaches [5]. In other words, due to the well-known

DOI: 10.24507/icicel.17.12.1395

"dimensionality curses", algorithms proposed for anomaly detection in low-dimensional data are unsuitable for high-dimensional data [6].

Clustering, an unsupervised learning technique, has extensively been used in anomaly detection. There are several approaches of clustering, namely partitioning, hierarchical, and density based approach. In [7], the authors proposed a k-means algorithm-based technique for traffic anomaly detection that uses the weighted Euclidean distance. Several works have been conducted in this direction. In [8], the authors have proposed a fuzzy c-means clustering based method for anomaly detection in mixed data. In [9], the authors have proposed a hierarchical clustering approach for anomaly detection in mixed data. In [10], the authors have proposed a hybrid clustering approach consisting of both partitioning and hierarchical for anomaly detection in mixed data.

Pawlak [11] introduced, the rough set theory, to address uncertainty and vagueness that exist in any datasets. In [12], the author has proposed an efficient algorithm based on neighborhood rough set classification for the detection of anomalies from the datasets with mixed attributes. Thivagar and Richard [13] have proposed the definition of nano topological space with respect to a subset X of universe U in terms of lower and upper approximation of X. In [14], the authors have proposed to generate CORE (a subset of attribute set) of conditional attribute set for medical diagnosis.

Density Based Spatial Clustering of Applications with Noise (DBSCAN) [15] is a cluster discovery technique that uses a density-based approach to find clusters of any shape. DBSCAN's main principle is that the neighborhood of a given radius must contain at least a minimal number of data objects for each cluster object. It has been used for anomaly or outlier identification multiple times because of its robustness to outliers. While DBSCAN works nicely in separating high-density clusters from low-density, it struggles with the high-dimensionality. The authors of [16] presented Novel Anomaly Detection-Density Based Spatial Clustering of Applications with Noise (NAD-DBSCAN), an unsupervised clustering technique that clusters the trajectories of moving objects of various sizes and shapes. The authors of [17] have presented thorough methodologies for anomaly detection as well as other machine learning disciplines such as pattern recognition, outlier detection, spam detection, suspicious detection, fraud detection, deep learning, and novelty recognition. The authors of [18] have proposed HDoutliers, a powerful unsupervised technique for detecting anomalies in high-dimensional data that has a solid theoretical base.

The authors of [19] proposed a thorough literature review of contemporary outlier detection strategies employed in the context of the Internet of Things. The authors of [20] developed a Modified DBSCAN technique that can be used to discover both global and local anomalies in seasonal data. The authors attempted to build unsupervised machine learning models to find anomalies on unlabeled pump measurements using high frequency sampled current and voltage time series data in their thesis [21]. The authors of [22] proposed a combination strategy for dimension reduction that included Principal Component Analysis (PCA) and Density Based Spatial Clustering of Applications with Noise (DBSCAN). The authors of [23] explored the lacunae of existing dimensionality reduction approaches and presented an unsupervised anomaly detection scheme based on Deep Auto Encoder (DAE) and clustering algorithms to model the data.

Although some of the aforesaid methods try to address the high-dimensionality in couple applications efficiently, they fail to address the issues related to IoT anomaly. Also, it is well-known fact that the efficacy of anomaly detection techniques decreases proportionately with the increment in dimensions, finding anomalies in any high-dimensional IoT data can be a challenging work. In this article, we try to address both the problems of high-dimensionality and efficient anomaly detection and propose a hybrid approach consisting of rough set theory and DBSCAN clustering algorithm for this. Our approach is a two-stage method. In the first stage, we apply data pre-processing techniques to converting the information system as set-valued ordered information system, generate a dominance relation, and a nano topology and its basis with respect to it, and then applying criterion reduction process the insignificant attributes are removed to find CORE (subset of the attribute set). In stage two, we apply the algorithm DBSCAN on the CORE to generating clusters. Thus, the insignificant attributes are dropped before the actual clustering process. In other words, the algorithm first uses rough set theoretical approach to find a lower dimensional space by removing irrelevant attributes. Then the algorithm DBSCAN is applied on it to finding the clusters along with the noises. The extracted noises are considered as anomalies. Then, the time-complexity of the method is computed. Finally, the proposed method is implemented using MATLAB and the dataset Kitsune Network Attack Dataset [24] and comparative analysis is made with a well-known algorithm, namely Incremental Possibistic Clustering (IPC) algorithm [25].

The article is prescribed as follows. The problem definition is given in Section 2. The proposed method is given in Section 3. The complexity analysis is given in Section 4. The experimental results and discussions are given in Section 5, and finally, we conclude the paper with conclusions in Section 6.

2. **Problem Definitions.** In this section, we present some important terms and definitions from [14,26,27] used in this paper.

Definition 2.1. *A set-valued information system is given by quadruple $S = (U, A, V, f)$, where U is a non-empty finite set of objects, A is a finite set of attributes, $V = \cup V_a$, where V_a is a domain of the attribute $a \in A$. We define $f : U \times A \to P(V)$, such that $\forall x \in U$ and $a \in A$, $f(x, a) \in V_a$ and $f(x, a) \geq 1$. Also $A = \{C \cup \{d\}; C \cap \{d\} = \phi\}$, where C is the set of conditional attributes and d the decision attributes.*

Definition 2.2. *If the domain of a conditional attribute can be arranged in ascending or descending order of preferences, then such attribute is called as criterion. If every conditional attribute is a criterion, then the information system is known as set-valued ordered information system.*

Definition 2.3. *If the values of some objects in U under a conditional attribute can be ordered according to an inclusion increasing or decreasing preferences, then the attribute is an inclusion criterion.*

Definition 2.4. *Let us consider a set-valued ordered information system with inclusion increasing preference. Also let $R_A^{\geq}$ be a relation defined as [see e.g., [26]]*

$$R_A^{\geq} = \{(y, x) \in U \times U : f(y, a) \geq f(x, a) \, \forall a \in A\} \tag{1}$$

$R_A^{\geq}$ is said to be the dominance relation on U. When $(y, x) \in R_A^{\geq}$, then $y \geq_A x$, which means y is at least as good as x with respect to A.

Property 2.1. *The inclusion dominance relation $R_A^{\geq}$ is i) reflexive, ii) unsymmetric, and iii) transitive.*

Definition 2.5. *For $x \in U$, the dominance class of x is given by*

$$[x]_A^{\geq} = \left\{y \in U : (y, x) \in R_A^{\geq}\right\} = \{y \in U : f(y, a) \geq f(x, a), \forall a \in U\} \tag{2}$$

where $U_A^{\geq} = \left\{[x]_A^{\geq} : x \in U\right\}$ is the family of dominance classes.

Remark 2.1. *$U_A^{\geq}$ is not a partition of U, but induces a covering of U, that is $U = \cup [x]_A^{\geq}$.*

Definition 2.6. *Given a set-valued ordered information system $S = \{U, A, V, f\}$ and a subset X of U, the upper approximation and lower approximation of X are respectively given by [26]*

$$U_A^{\geq}(X) = \left\{x \in U : [x]_A^{\geq} \cap X \neq \phi\right\} \tag{3}$$

and

$$L_A^{\geq}(X) = \left\{x \in U : [x]_A^{\geq} \subseteq X\right\} \tag{4}$$

Also the boundary region of X,

$$B_A^{\geq}(X) = U_A^{\geq}(X) - L_A^{\geq}(X) \tag{5}$$

Definition 2.7. *Given a set-valued ordered information system S, a subset B of A is said to be a criterion reduction of S if $R_A^{\geq} = R_B^{\geq}$ and $R_M^{\geq} \neq R_A^{\geq}$ for any $M \subseteq A$. Otherward, a criterion reduction of S is a minimal attribute set B such that $R_A^{\geq} = R_B^{\geq}$. Also* $\mathrm{CORE}(A) = \left\{a \in A : R_A^{\geq} \neq R_{A-\{a\}}^{\geq}\right\}$.

Definition 2.8. *Let $R_C^{\geq}$ be a dominance relation on U, then $\tau_C^{\geq}(X) = \left\{U, \phi, U_C^{\geq}(X), L_C^{\geq}(X), B_C^{\geq}(X)\right\}$ forms a nano topology [14] on U with respect to X. And $\beta_C^{\geq}(X) = \{U, U_C^{\geq}(X), L_C^{\geq}(X)\}$ is the basis for $\tau_C^{\geq}(X)$. Furthermore,* $\mathrm{CORE}(C) = \left\{a \in C : \beta_C^{\geq} \neq \beta_{C-\{a\}}^{\geq}\right\} = \cap\, \mathrm{red}(C)$ *where* $\mathrm{red}(C)$ *denotes the criterion reduction.*

Definition 2.9. *Let $S = (U, A, V, f)$ be an information system consisting of m entities or objects $x_1, x_2, \ldots, x_m$. Let the attribute set A has n members. Then, S can be viewed as an $m \times n$ matrix in which rows represent objects and columns represent attributes. Attributes can be termed as features or dimension.*

Definition 2.10. $\boldsymbol{\varepsilon}$-Neighbourhood of an Object. *For a given non-negative value ε, the ε-neighbourhood of an object $x \in S$ denoted by $\varepsilon\text{-nbd}(x)$, is defined as $\varepsilon\text{-nbd}(x) = \{y \in S : d(x, y) \leq \varepsilon\}$, where d is any metric on S [see e.g., [27]].*

Definition 2.11. Core Object. *An object $x \in S$ is said to be core object if $|\varepsilon\text{-nbd}(x)| >$ MinPts in S. In other words, a core object has a neighbourhood of user-specified minimum density.*

Definition 2.12. Directly Density-Reachable. *An object $x \in S$ is directly-density-reachable from an object $y \in S$ with respect to ε and* MinPts*, if x is a core object and y is in its ε-neighbourhood [27].*

Definition 2.13. Density-Reachable. *A point $x_i \in S$ is said to be density-reachable from $x_j \in S$ with respect to ε and* MinPts *in S if there is a chain of points $x_1, x_2, \ldots, x_n$ in S such that $x_1 = x_j$, $x_n = x_i$, such that $x_e \in S$ and x_{e+1} is directly-reachable from x_e with respect to ε and* MinPts *in D_i.*

Definition 2.14. Density-Connected. *An object $x_j \in S$ is said to be density-connected to another point $x_j \in S$ with respect to ε and* MinPts *in S if there exists another point $x_k \in S$ such that both x_i and x_j are density-reachable from x_k with respect to ε and* MinPts *in D_i.*

Definition 2.15. Cluster. *A cluster C with respect to ε and* MinPts *is a non-empty subset of S satisfying the following conditions.*

1) For all $x_i, x_j \in S$ if $x_i \in C$ and x_j is density-reachable from x_i with respect to ε and MinPts*, $x_j \in C$.*

2) For all $x_i, x_j \in C$, x_i is density-connected to x_j with respect to ε and MinPts *in S.*

Definition 2.16. *Let $C_1, C_2, \ldots, C_k$ be clusters of S with respect to ε and* MinPts*, and the noise is an object which does not belong to any of the clusters.*

3. **Proposed Method.** For finding anomalies, we use a density-based subspace clustering approach. The method first uses rough set theoretic approach for attribute or dimension reduction and then uses DBSCAN [15] for finding clusters along with noises. The proposed method is as follows. Our dataset $S = (U, A)$ is an information system consisting of both conditional attributes and decision attributes. First of all, we apply data pre-processing techniques to converting the information system as set-valued ordered information system. We then, generate a dominance relation on it. With respect to the dominance relation, we generate a nano topology and its basis. Then we apply the criterion reduction process to generate CORE(A) as a subset of attribute set A. We define a new information system $E = (U, \mathrm{CORE}(A))$ on U which is a lower dimensional space. The pseudocode of the algorithm for the criterion reduction is given in Algorithm 1.

Algorithm 1: Subspace Generation	
Input.	(U, A): the information system, where the attribute set A is divided into C-conditional attributes and D-decision attributes, consisting of n objects
Output.	Subspace of (U, A)
Step1.	Generate a dominance relation $R_C^{\geqq}$ on U corresponding to C and $X \subseteq U$
Step2.	Generate the nano topology $\tau_C^{\geqq}(X)$ and its basis $\beta_C^{\geqq}(X)$
Step3.	for each $x \in C$, find $\tau_{C-\{x\}}^{\geqq}(X)$ and $\beta_{C-\{x\}}^{\geqq}(X)$
Step4.	if $\left(\beta_C^{\geqq}(X) = \beta_{C-\{x\}}^{\geqq}(X)\right)$
Step5.	then drop x from C
Step6.	else form criterion reduction
Step7.	end for
Step8.	generate CORE$(C) = \cap$ {criterion reductions}
Step9.	Generate subspace of the given information system

The above algorithm supplies the CORE of the attribute set by removing insignificant attributes, giving a subspace $E = (U, \mathrm{CORE}(A))$ of $S = (U, A)$. Then DBSCAN [15] is applied on E as follows. Initially, all objects in E are marked as "unvisited". The algorithm randomly chooses an unvisited object p, marks it as "visited", and checks whether the ε-neighborhood of p has at least MinPts objects or not. If not, p is marked as a noise point else new cluster C is created for p, and all the objects in the ε-neighborhood of p are added to a candidate set, N. Algorithm iteratively adds those objects in N to C, which do not belong to any cluster. In this process, for an object p_0 in N that carries the label "unvisited", DBSCAN marks it as "visited" and checks its ε-neighborhood. If the ε-neighborhood of p_0 has at least MinPts objects, those objects in the ε-neighborhood of p_0 are added to N. DBSCAN continues adding objects to C until C can no longer be expanded, that is, N is empty. At this time, cluster C is completed, and thus is output.

To find the next cluster, DBSCAN randomly selects an unvisited object from the remaining ones. This process continues till all objects are visited. The algorithm supplies a set of clusters along with noises. The noises are object or observation which deviate from others which arouse suspicions and they are generally termed as anomalies. The pseudocode of the algorithm DBSCAN [15] is given in Algorithm 2.

4. **Complexity Analysis.** For generating dominance classes and corresponding classes, the algorithm needs to compare values of all the possible pairs of objects from U in all dimensions, there can be at most $|U| \times |U| \times |C|$ number of comparison. So, the computational complexity for step1 is $\mathrm{O}(m^2 \cdot n)$, where $|U| = m$, and $|C| = n$. For generating the nano topology, the lower approximation and approximation of the set has to be generated, which takes computational time $\mathrm{O}\left(|X| \cdot |U|\right)$. So the total computational cost of step1 and step2 is $\mathrm{O}\left(m^2 \cdot n + |X| \cdot |U|\right) = \mathrm{O}\left(m^2 \cdot n\right)$ which is the worst-case complexity.

Algorithm 2: DBSCAN

Input.	E: Information system consisting of n objects and attribute set CORE$(A) \subseteq A$, ε: the radius parameter, and MinPts: the neighborhood density threshold
Output.	Set of clusters along with noises
Step1.	mark all objects as unvisited
Step2.	do
Step3.	{ randomly select an unvisited object in p;
Step4.	mark p as visited;
Step5.	if ($\lvert\varepsilon$-nbd$(p)\rvert \leq$ MinPts) //ε-nbd(p)-ε-neighborhood of p
Step6.	{ create a new cluster C; add p to C;
Step7.	for each point $p_0 \in N$ //N-set of objects in ε-nbd(p)
Step8.	{ if (p_0 not visited)
Step9.	mark p_0 as visited;
Step10.	if ($\lvert\varepsilon$-nbd$(p_0)\rvert \geq$ MinPts)
Step11.	add those points of ε-nbd(p_0) to N;
Step12.	if (p_0 is not member of any cluster)
Step13.	add p_0 to C;
Step14.	} end for
Step15.	output C;
Step16.	}
Step17.	else
Step18.	mark p as noise
Step19.	} (until no object is unvisited)

From step3, do loop starts and it runs over at most all the attribute set. The computation of step4 to step7 takes constant time say $\mathrm{O}(k_1)$, where k_1 = constant. Therefore, the computational cost from step3 to step8 is $\mathrm{O}(k_1 n)$. Similarly, that of step9 and step10 is also constant say $\mathrm{O}(k_2)$, where k_2 = constant. The overall complexity of Algorithm 1 is $\mathrm{O}\big(m^2 \cdot n + k_1 n + k_2\big) = \mathrm{O}\big(m^2 \cdot n\big)$. Again, the average-case run-time and worst-case run-time of DBSCAN algorithm is $\mathrm{O}(m \cdot \log m)$ and $\mathrm{O}\big(m^2\big)$, respectively. So, the worst case computational complexity of our method is $\mathrm{O}\big(m^2 \cdot n + m^2\big) = \mathrm{O}\big(m^2 \cdot n\big)$.

5. **Experimental Settings and Discussions.** The experiment is conducted with Kitsune Network Attack dataset [24] collected from UCI machine repository. The dataset characteristics are given in the Table 1 as follows.

Table 1. Dataset characteristic

Dataset	Dataset characteristic	Attribute characteristic	No of instances	No of attributes
Kitsune Network Attack dataset	Multivariate, sequential, Time-series	Real	27170754	115

The method is implemented using MATLAB. The implementation process consists of three stages, input data pre-processing, subspace clustering, and testing. First of all, the method accepts the input data and converts it to set-valued matrix. The matrix representation of the dataset is the information system. Since rough set cannot deal with continuous attribute, they are discretized at the same time. The Algorithm 1 is then applied to finding the subset of the attribute set by removing the insignificant attributes and by using the concept of dominance relation, nano topology and its basis. The Algorithm 1 gives subset as CORE of the attribute set. Then the algorithm DBSCAN is applied on CORE to finding clusters along with noises. For efficient implementation, two parameters namely ε and MinPts are to be specified. Since we are working on the high-dimensional data, the above two parameters are heavily dependent on the number of attributes in the subspace obtained by Algorithm 1. Usually the MinPts is derived from the number of dimensions in the subspaces. In our case, the MinPts values are taken from 6 to 12 depending on the number of attributes (as MinPts $\geq$ dimension-1). Similarly, for efficiently specifying ε value, we have applied a well-known method using k-distance graph [28] by plotting the distance to the k = MinPts-1 nearest neighbor ordered from the largest to the smallest value. Based on MinPts value, we have chosen the different values of ε. The method is tested multiple times for different sizes of datasets (number of attributes) and the results are recorded. The IPC algorithm [25] is also implemented with the similar parameter settings. The performances of both the methods in terms of accuracies in detection rate are recorded. The findings of the above experiments are expressed in the tabular form in Table 2 and graphically in Figure 1 below.

Table 2. Comparative performances analysis of PC and our method

	% of anomaly detection rate (performances)					
No of attributes	5	10	20	50	80	115
IPC algorithm's accuracy of detection rate	99%	98.8%	96.2%	91.05%	85.07%	80.02%
Our method's accuracy of detection rate	99%	98.02%	97.02%	94.03%	91.07%	88.03%

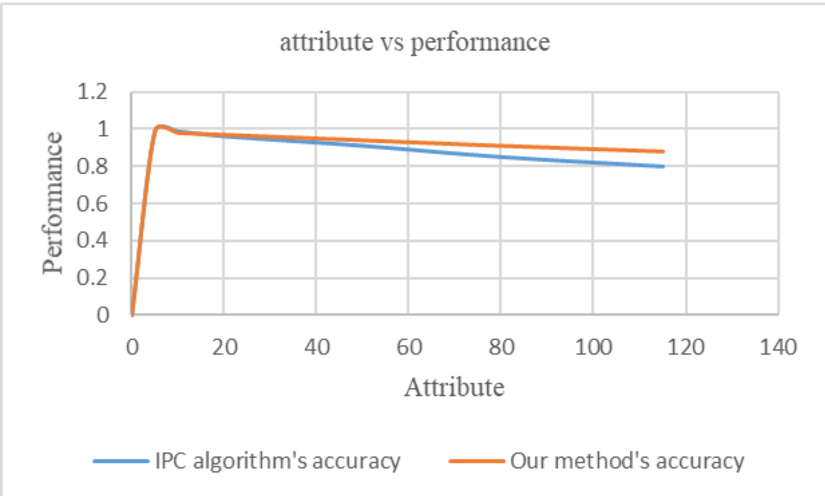

Figure 1. Comparative analysis of IPC and our method in terms of performances

It can be seen from the experimental results that when the number of attributes is less, both the methods have good detection rate. For example, when the numbers of attributes are 5 and 10, the accuracy levels are around 99% and 98%, respectively. However, as the number of attributes increases, both methods' performances decrease, but the rate of decrease of our method is less than that of IPC algorithm [25]. In other words, IPC algorithm's [25] performance decreases rapidly in comparison to our method with the increase in the dimensions. It shows that our method outperforms IPC algorithm [25].

6. **Conclusions.** In this article, we have proposed a method based on hybrid approach consisting of both rough set theory and DBSCAN clustering algorithm for the anomaly detection. The rough set theory is used for attribute reduction and DBSCAN is used for efficient clustering. The method gives clusters along with noises in lower dimensional

space by ignoring insignificant attributes. The noises are assumed to be anomalous object. The method proceeds as follows. First of all, we apply data pre-processing techniques to converting the information system as set-valued ordered information system and generate a dominance relation on it. With respect to the dominance relation, we generate a nano topology and its basis. With the help of these, we reduce attribute set by removing insignificant attributes, which will give us a new information system in lower dimensional space. Then the algorithm DBSCAN is applied on it to finding clusters along with noises. The method's computational complexity is found to be $\mathrm{O}\left(m^2 \cdot n\right)$. The method's efficacy is established with the help of comparative analysis with a well-known method.

Acknowledgment. The author gratefully acknowledges the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] P. Sethi and S. Sarangi, Internet of Things: Architectures, protocols, and applications, *Journal of Electrical and Computer Engineering*, vol.2017, 9324035, DOI: 10.1155/2017/9324035, 2017.

[2] H. Y. Teh, K. I. Wang and A. W. Kempa-Liehr, Expect the unexpected: Un-supervised feature selection for automated sensor anomaly detection, *IEEE Sensors Journal*, vol.21, no.16, pp.18033-18046, DOI: 10.1109/JSEN.2021.3084970, 2021.

[3] S. M. Erfani, S. Rajasegarar, S. Karunasekera and C. Leckie, High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning, *Pattern Recogn.*, vol.58, pp.121-134, 2016.

[4] V. Hodge and J. Austin, A survey of outlier detection methodologies, *Artificial Intelligence Review*, vol.22, no.2, pp.85-126, 2004.

[5] C. C. Aggarwal and S. Y. Philip, An effective and efficient algorithm for high-dimensional outlier detection, *The VLDB Journal*, vol.14, no.2, pp.211-221, 2005.

[6] A. Ramchandran and A. K. Sangaiah, Chapter 11 – Unsupervised anomaly detection for high dimensional data – An exploratory analysis, in *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, Academic Press, 2018.

[7] G. Münz, S. Li and G. Carle, Traffic anomaly detection using k-means clustering, *GI/ITG Workshop MMBnet*, vol.7, no.9, 2007.

[8] W. Ren, J. Cao and X. Wu, Application of network intrusion detection based on fuzzy c-means clustering algorithm, *The 3rd International Symposium on Intelligent Information Technology Application*, pp.19-22, 2009.

[9] F. A. Mazarbhuiya, M. Y. AlZahrani and L. Georgieva, Anomaly detection using agglomerative hierarchical clustering algorithm, in *Lecture Notes in Electrical Engineering*, Singapore, Springer, DOI: 10.1007/978-981-13-1056-0_48, 2018.

[10] F. A. Mazarbhuiya, M. Y. AlZahrani and A. K. Mahanta, Detecting anomaly using partitioning clustering with merging, *ICIC Express Letters*, vol.14, no.10, pp.951-960, 2020.

[11] Z. Pawlak, Rough sets, *International Journal of Computer and Information Sciences*, vol.11, pp.341-356, 1982.

[12] F. A. Mazarbhuiya, Detecting anomaly using neighborhood rough set based classification approach, *ICIC Express Letters*, vol.17, no.1, pp.73-80, 2023.

[13] M. L. Thivagar and C. Richard, On nano forms of weakly open sets, *International Journal of Mathematics and Statistics Invention*, vol.1, no.1, pp.31-37, 2013.

[14] M. L. Thivagar and S. P. R. Priyalatha, Medical diagnosis in an indiscernibility matrix based on nano topology, *Cogent Mathematics*, vol.4, no.1, 1330180, 2017.

[15] M. Ester et al., A density-based algorithm for discovering clusters in large spatial databases with noise, *Proc. of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD'96)*, pp.226-231, 1996.

[16] R. Ranjith, J. J. Athanesious and V. Vaidehi, Anomaly detection using DBSCAN clustering technique for traffic video surveillance, *The 7th International Conference on Advanced Computing*, pp.1-6, 2015.

[17] S. Thudumu, P. Branch, J. Jin and J. Singh, A comprehensive survey of anomaly detection techniques for high dimensional big data, *Journal of Big Data*, vol.7, 42, 2020.

[18] P. D. Talagala, R. J. Hyndman and K. Smith-Miles, Anomaly detection in high-dimensional data, *Journal of Computational and Graphical Statistics*, vol.30, no.2, pp.360-374, 2021.

[19] M. A. Samara, I. Bennis, A. Abouaissa and P. Lorenz, A survey of outlier detection techniques in IoT: Review and classification, *Journal of Sensor and Actuator Networks*, vol.11, no.4, pp.1-31, 2022.
[20] P. Jain, M. S. Bajpai and R. Pamula, A modified DBSCAN algorithm for anomaly detection in time-series data with seasonality, *The International Arab Journal of Information Technology*, vol.19, no.1, pp.23-28, 2022.
[21] P. Hanna and E. Swartling, *Anomaly Detection in Time Series Data Using Unsupervised Machine Learning Methods: A Clustering-Based Approach*, Degree Project in Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden, 2020.
[22] A. Y. Srihari and Sashirekha. K, Dimensional reduction of data for anomaly detection and speed performance using PCA and DBSCAN, *International Journal of Engineering and Advanced Technology*, vol.9, no.1S2, pp.39-41, 2019.
[23] C. Zhang, J. Liu, W. Chen, J. Shi, M. Yao, X. Yan, N. Xu and D. Chen, Unsupervised anomaly detection based on deep autoencoding and clustering, *Hindawi Security and Communication Networks*, vol.2021, pp.1-8, 2021.
[24] Y. Mirsky, T. Doitshman, Y. Elovici and A. Shabtai, Kitsune: An ensemble of autoencoders for online network intrusion detection, *Network and Distributed System Security Symposium 2018 (NDSS'18)*, 2018.
[25] T. Yang, S. Liu and J. Liu, Network traffic anomaly detection based on incremental possibilistic clustering algorithm, *Journal of Physics: Conference Series*, vol.1284, 012067, 2019.
[26] Y. Qian, C. Dang, J. Liang and D. Tang, Set-valued ordered information systems, *Information Sciences*, vol.179, no.16, pp.2809-2832, 2009.
[27] A. K. Pujari, *Data Mining Techniques*, University Press, 2001.
[28] G. L. Miller, S.-H. Teng, W. Thurston and S. A. Vavasis, Separators for sphere-packings and nearest neighbor graphs, *Journal of the Association for Computing Machinery*, vol.44, no.1, pp.1-29, DOI: 10.1145/256292.256294, 1997.