Hindawi
Security and Communication Networks
Volume 2022, Article ID 9149164, 15 pages
https://doi.org/10.1155/2022/9149164

Research Article

iReTADS: An Intelligent Real-Time Anomaly Detection System for Cloud Communications Using Temporal Data Summarization and Neural Network

Gotam Singh Lalotra,[1] **Vinod Kumar**,[2] **Abhishek Bhatt**,[3]
Tianhua Chen,[4] **and Mufti Mahmud**[5,6,7]

[1]*Department of Computer Science, Govt. Degree College Basohli, University of Jammu, Jammu, India*
[2]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India*
[3]*Department of Electronics and Telecommunication Engineering, College of Engineering, Pune, India*
[4]*Department of Computer Science, School of Computing and Engineering, University of Huddersfield, UK*
[5]*Department of Computer Science, Nottingham Trent University, Nottingham, UK*
[6]*Medical Technologies Innovation Facility, Nottingham Trent University, Nottingham, UK*
[7]*Computing and Informatics Research Centre, Nottingham Trent University, Nottingham, UK*

Correspondence should be addressed to Gotam Singh Lalotra; singh.gotam@gmail.com and
Vinod Kumar; drvinodkumar2019@kluniversity.in

Received 20 September 2021; Revised 31 October 2021; Accepted 23 December 2021; Published 22 January 2022

Academic Editor: Thippa Reddy G

A new distributed environment at less financial expenditure on communication over the Internet is presented by cloud computing. In recent times, the increased number of users has made network traffic monitoring a difficult task. Although traffic monitoring and security problems are rising in parallel, there is a need to develop a new system for providing security and reducing network traffic. A new method, iReTADS, is proposed to reduce the network traffic using a data summarization technique and also provide network security through an effective real-time neural network. Although data summarization plays a significant role in data mining, still no real methods are present to assist the summary evaluation. Thus, it is a serious endeavor to present four metrics for data summarization with temporal features such as conciseness, information loss, interestingness, and intelligibility. In addition, a new metric time is also introduced for effective data summarization. Finally, a new neural network known as Modified Synergetic Neural Network (MSNN) on summarized datasets for detecting the real-time anomaly-behaved nodes in network and cloud is introduced. Experimental results reveal that the iReTADS can effectively monitor traffic and detect anomalies. It may further drive studies on controlling the outbreaks and controlling pandemics while studying medical datasets, which results in smart healthy cities.

1. Introduction

In the last one and half decade, computer technology has significantly overpowered the conventional ways of handling the daily routines in almost every walk of life. All the daily routine activities like reading newspapers, shopping, running a business, studying, and a lot of official works have taken a shift over to the computer networks like LAN, WAN, MAN, Internet, and cloud computing. As these computer networks offer daily routine services to people across the globe, at the same time, attackers have also joined the international community on the same platform, but to disturb the streamlined activities over the networks. These kinds of regular attackers or hackers not only affect the daily routine activities but also disturb the business or government networks [1]. Countering hackers and ensuring the smooth working of computer networks need the construction of a new security mechanism to provide security to the network users and their own secret stored data. For safe and secure communication services for exponential growing e-business and electronic transactions, the use of intrusion detection and prevention systems, encryption, firewalls, authentication, and effective security mechanisms has been done [2]. Data have a significant role in every domain. The storage requirements over networks, in addition to the analysis of data, are of utmost importance to obtain knowledge [3]. A new secure arrangement method is presented, which is based on matrix eigenvalue. The aim of these arrangements is to generate a secret position for each user for storing data, which is called a secure arrangement [4]. Sometimes, the input data could be faster and easier to examine for attaining similar knowledge. For instance, a network administrator over the computer network requires surveillance and supervising the activities of the network [5]. Yet, for a small corporation network like HTTP, FTP, e-mail, and P2P applications, the quantity of data generated is enormous and difficult to analyze [6]. Moreover, the network traffic is increasing at a very rapid rate, which in turn becomes infeasible to monitor a network in real time by administrator [7]. Therefore, to analyze the current scenario of the network, a summary of the network traffic is quite useful to immediately review the situation of the network.

For a large volume of different kind of data generated from different resources like wireless sensors and cloud [8, 9], a summary is essential [10]. The intent of summarizing is to present a crisp dataset as input [11, 12]. Summarization is extensively traversed in different domains such as network data streams [11], intrusion detection systems (IDS) [13, 14], point of sales (POS) data [15], and natural text processing [16]. The summarization has been applied to various domains like healthcare, transport, security, logistics, and daily life [17] and has been demonstrated to be efficient in getting useful data out of huge datasets generated through IoT (Internet of Things) and cloud applications, which is easier to understand or interpret. It becomes even more important to summarize when the whole world is facing a pandemic outbreak and each and every sector of human life is affected.

In the same manner, a network sniffer protects and collects packets in an indiscriminate way, and an intrusion detection system (IDS) does the same. An IDS has the capability to detect different kinds of network attacks in the presented environment. The malicious network activities are identified by analyzing the packets collected by IDS, which gives alert signals to the system administartor and attack connections are blocked to avoid additional destruction from attacks. In general, intrusion detection algorithms are categorized as misuse detection and anomaly detection [18]. Misuse detection algorithm identifies attacks on the basis of known attack signatures. These algorithms are efficient in identifying known attacks with low errors. These algorithms are unable to identify newly created attacks, which do not have similar properties to the known attacks. On the other hand, the anomaly detection method relies on the hypothesis that the attackers have different behavior than a normal user. This paper, being part of *i*ReTADS, presents the following contributions:

(i) An existing metric named information loss has been modified that is biased towards recurring attributes and proposed a novel summarization technique that is based on a newly defined time metric for data summarization purpose

(ii) The newly proposed metric has been employed to split the dataset into different time intervals

(iii) A novel method named modified synergetic neural network (MSNN) has been designed for effective anomaly detection

Further, the paper is organized as follows: Section 2 offers the related works. Section 3 discusses the overall system architecture. Section 4 explains the proposed work. Section 5 gives the results and discussion. Section 6 has the conclusion and future enhancements.

2. Related Works

The techniques of data summarization and anomaly detection have already been extensively researched. For association rule mining and clustering various data, summarization techniques are used, and different metrics have been proposed to improve the technique of data summarization. The authors in [19] have demonstrated that there are no universally excepted standards on the subject of what is a good summarization technique or a good summary. The aim of their technique is to represent a transactional database, implementing the notion of hyperrectangles, the Cartesian product of a set of transactions, and a set of items. To define the effectiveness of hyper and hyper + techniques. They calculate the conciseness and the quality as the ratio of coverage per cost of each hyperrectangle so that the final summary cannot be compared to another summary as there is no measure to evaluate it. In [11], the authors explored the technique of compacting the specified number of transactions to a smaller set of summaries so that every summary entitles a subset of the input transactions in such a way that every transaction is represented in the summary. The original dataset is considered as the summary by the Bottom-Up Summarization (BUS) algorithm. In the beginning, frequent item set mining is employed on the input dataset, and then item sets are searched greedily, replacing those with minimum information loss and maximum compaction gain, and represented data points are replaced with them in the summary. The process is repeated till the desired compaction gain is achieved. Here, the metric used is the same for all the techniques of compaction gain. Information loss is also used by the authors to evaluate the results for measuring the amount of information not present in the summary of the original data. Table 1 represents some of the summarized contribution towards anomaly detection. The problem of summarization of a dataset of transactions, where two objective functions, compaction gain and information loss, were used with categorical attributes, is an optimization problem by authors in [11]. In order to describe the output of any summarization algorithm, a new metric was presented by them, and for addressing this problem, they investigated two approaches. In the first approach, clustering was implemented, and for the second approach, the frequent

Table 1: Summarized contribution towards the anomaly detection.

SL	Title	Method	Dataset	Pros	Cons
1	RADS: Real-time anomaly detection system for cloud data centres [20]	OpenStack-based cloud data centre, one-class classification (RF, SVM, and naïve Bayes), and window-based time series analysis	Twitter dataset	Achieved 90–95% accuracy with a low FPR of 0–3%	Two metrics, precision and recall, need to be investigated for proper evaluation of the system
2	Real-time anomaly detection using ensembles [21]	Base learner (i) perceptron; (ii) ML-OzaBagadWin; (iii) binary class SVM	KDD CUP 99	Accuracy attained 89.9% by MLP	Only a few base learners were used
3	A real time anomalies detection system based on streaming technology [22]	Spout architecture	Data are one-hour (22:00–23:00) exported flow data in *L* province, China	Real-time anomalies detection from mass stream data in a scalable manner	Up to 4 GB dataset is tested
4	Adapted K-nearest neighbors for detecting anomalies on spatio-temporal traffic flow [23]	K-nearest neighbors	Urban traffic flow Beijing dataset	Able to detect the real distribution of flow outliers. Outperforms the baseline algorithms for high-urban traffic flow	Does not work well with high dimensions because of the inherent feature of k-NN
5	Real-time anomaly detection based on long short-term memory and Gaussian mixture model [24]	LGMAD, based on long short-term memory (LSTM) and Gaussian mixture model (GMM)	NAB public dataset and synthetic dataset	A novel idea of the health factor *alpha* is proposed additionally to describe the health level of the system	Evaluated on precision, recall, F1-measure, and overlooked accuracy
6	Malware traffic classification using convolutional neural network for representation learning [25]	Convolution Neural Network (CNN)	USTC-TRC2016 traffic dataset	Malware traffic classification	Study the CNN parameters tunings
7	Adaptive real-time anomaly detection in cloud infrastructures [26]	Robust PCA and SVD	Amazon CloudWatch and Yahoo!	Accuracy: 87.24%; F-measure: 86%	Precision, recall, and metrics are ignored in evaluation
8	ADSaS: comprehensive real-time anomaly detection system [27]	Seasonal autoregressive integrated moving average (SARIMA) model and seasonal trend decomposition using loess (STL)	Numenta Anomaly Benchmark (NAB)	ADSaS performed well in terms of precision, recall, and F1-score	Error range variation is large in precision: 2.5%–97%; F1-score: 4.9%–95.1%; recall: 22.2%–100%

item sets from the association analysis domain were used. In their work of summarization, they proposed one of the applications in the field of network data in which they showed how their technique could be efficiently used for summarizing network traffic into a meaningful and compact representation.

In [28], a recent investigation was done to find the possibility of anomaly detection in the context of real-time big data preprocessing and machine learning techniques. This survey includes the essential components of real-time processing of big data for anomaly detection, taxonomy of real-time big data processing, and various research challenges.

The authors in [29] studied the concept of information gain for network summarization and put forward a measure called information entropy for measuring the quality of a resolution. A probabilistic model of the information contained in a network was developed, and a formula is derived based on this model for information entropy. The network summarization method determines the computational complexity of computing network entropy; for simple deterministic node-reduction summarizations, they developed an O(E) algorithm. In order to decide the most appropriate level of summarization, analysts use network entropy. With the help of information entropy, the information is measured over the network; this network information is combined with information provided by other attributes like geospatial labels for providing a complete scenario of the information enclosed in a particular network resolution.

In [30], a hierarchical data summarization data structure is presented; it was labeled hierarchical as the data structure implemented the concept of subcomponents to systematically attain conceptually larger components. The methods proposed herein acquire a bigger component repeatedly induced by the domain understanding of the users. So, for hierarchical data summarization, the rules implemented in the creation of data structure like *B*+ trees were also considered, and various data structures were implemented in hierarchical data summarization. Authors in [31] proposed estimation and a real-time loss performance monitoring scheme. Asymptotic relationship between the buffer size for both Markovian and self-similar traffic and Common Language Runtime was used in the proposed scheme. Results obtained by implementing the proposed scheme showed that it required less monitoring time and obtained improved accuracy in comparison to the existing schemes.

The paper [32] presented a group of techniques and methods for traffic data collection, preprocessing, transformation, and integration till the data is forwarded for processing and transfer further for integration or fusion. Real-time data is very imperative for encouraging model accuracy, comprehensive use of assignment models, and historical traffic data for assisting online services and operations. The reliability of information and output from data fusion and processing as proposed by authors in [33] proposed the concern for analyzing the large network data for packet loss in real time, irrespective of any device installed at the network node in advance at monitoring place. However, it is found that the proposed system needs some sort of training in advance for adopting the features of the traffic to be monitored. The time series models are used for training which efficiently represent the high-speed traffic; with the help of these models, important conclusions like how to sample the data can be drawn by simulating the similar behavior shared by traffic.

This work [34] proposed four metrics, conciseness, information loss, interestingness, and intelligibility. These could be used for characterizing data summarization results. However, they modified the information loss metric because of its biased nature towards the recurring attributes. With the use of these four metrics, they assessed the existing summarization techniques on renowned network traffic datasets. A summarization method based on an already existing method was proposed, but here it is taken as an objective function; further, the classification of summarized datasets is carried out to reveal the usability of the metrics. Authors [35], with the help of the Bayesian Network, explained anomaly detection and getting learned by the real world automated identification system (AIS) data and from the additional data, resulting in the production of static and dynamic Bayesian network model. In their finding, they proposed that learned networks were pretty easy to inspect and verify in spite of the large number of variables being incorporated. In order to improve the anomaly detection performance, they confirmed the combination of both static and dynamic modeling approaches for improving the coverage of the overall model.

This work focused [36] on reducing security risk and presented two techniques of the network traffic anomaly detection in cloud communication, and these techniques, with the assistance of synergetic neural networks and the catastrophe theory, understand the dynamic behavior of the network traffic. In synergetic neural networks, a synergetic dynamic equation along with a set of ordered parameters is implemented for describing the complex nature of the network traffic system over cloud communications. Once this equation is solved, the ordered parameters confirmed by the primary factors can converge to 1, which results in anomaly detection. Catastrophe theory makes use of catastrophe potential function to explain the catastrophe dynamic process of the network traffic in cloud communications.

State of the network traffic derives from the normal one; whenever there is an anomaly in the network, the catastrophe distance index is used to assess the derivation, which helps in detecting the anomaly. They assessed the two approaches by implementing these techniques over standard Defense Advanced Research Projects Agency datasets, and it proved to be effective in detecting the network traffic anomaly and accomplished the high detection probability and low false alarm rate.

This contribution [37] presented a new increasing mapping-based hidden Markov model (IMHMM) in order to monitor the dynamic traffic efficiently. An increasing mapping is set up between the observation sequence and possible state sequence. In spite of FB variables, these mappings are used for obtaining the reestimation formulas for the model parameters. The IMHMM can be used to make fault detection and process monitoring framework to deal with large-scale dynamic processes. The IMHMM needs less storage space, and it is simple in comparison to HMM. Kim et al. [3] presented a new hybrid intrusion detection method that integrates hierarchically an anomaly detection model and a misuse detection model. Based on the C4.5 decision tree algorithm, a misuse detection model is built; later, using this model, normal training data is crumbled into smaller subsets. After that, these subsets are used to make multiple one-class SVM models. This method considerably optimizes the high time complexity of training and testing processes.

In this paper [38], a new technique SVM-L is given for anomaly detection in network traffic. Based on the concept of the dual formulation of kernel SVM and Linear Discriminant Analysis, an optimization model was proposed to adjust the hyperparameter of the classifier. Experimentally, 99% accuracy was claimed over network traffic dataset.

The work [39] proposed the ANN-based techniques for anomaly detection in Apache Spark, which works effectively for complex scenarios with multiple types of anomalies, like CPU contention, cache thrashing, and context switching anomalies, and showed 98–99% *F*-scores. Also, they claimed that a random duration, random start instant, and overlapped anomalies do not cause a significant influence on the performance of the proposed method.

The authors of [40] demonstrated the Convolution Neural Network features with bidirectional long short-term memory for real-time anomaly detection in the surveillance system. They claimed a 3.41% and 8.09% upsurge in accuracy on UCF-Crime and UCF-Crime2Local databases when compared to the newest methods.

The authors of [41] did a multiperspective review over smart anomaly detection in sensor systems and discussed the potential of computing (machine learning models), efficiency in communications medium, and engineering (constraints) in development of a smart anomaly detection system.

The research work [42] proposed a novel multistage anomaly detection ensemble technique named BFA-PDBSCAN for the incessant execution of computations on IoT-based applications. This selection of relevant features from the dataset is carried out by the Boruta method and extended k-medoid with a firefly-inspired strategy for performing partitioning. They claimed the effectiveness of the proposed model over several datasets.

Blockchain and smart contract-based dependable and efficient lightweight certificateless signature (CLS) scheme, which is more secure than CLS protocol alone (Susceptible to security risks), is popular for resource-constrained Industrial Internet of Things (IIoT) protocol design [43].

This paper presents a new model for anomaly detection in the cloud using data summarization and neural network. Network traffic in the cloud is monitored with the help of data summarization; it also collects the necessary data. This summarized data further can be sent to the proposed Modified Synergetic Neural Network for anomaly detection.

3. Proposed Work

In this paper, a new model called *i*ReTADS for detecting a real-time anomaly in the cloud during communications using data summarization and neural network with temporal features is introduced. Temporal Data summarization monitors the network traffic in the cloud in real time and collects the necessary data, and the summarized data can be sent to the proposed Modified Synergetic Neural Network for anomaly detection.

3.1. System Architecture of iReTADS. This proposed system architecture comprises seven key components, namely, data collection agent, cup dataset, network trace data, data summarization module, anomaly detection module, temporal information manager, and knowledge base, as shown in Figure 1.

3.1.1. Data Collection Agent. The network data are collected from the network layer or from the KDD'99 cup dataset by the data collection agents. This data is further sent to the data summarization module for summarizing the data.

3.1.2. Data Summarization Module. Data summarization module is comprised of three chief components as a quality threshold, that is, setting, optimization, and clustering. These components use different algorithms for quality threshold setting, optimization of the data based on the threshold, and clustering the data, and then this summarized data is sent for anomaly detection module.

3.1.3. Anomaly Detection Module. Anomaly detection module for efficient classification of the dataset makes use of proposed classification techniques. This module has a component that sets the rules on the basis of fuzzy concepts by integrating the various combinations of summarized datasets to have efficient classification. For setting the time interval, this module bears the responsibility of remaining in contact with the temporal information manager.

3.1.4. Temporal Information Manager. This module has the accountability of allotting the time interval for summarizing the data; with the assistance of the knowledge base, the time-based fuzzy rules also formed the knowledge base. The knowledge base has a set of rules to answer the queries being fired by the users and execute efficient decision-making. It has contained rules in order to make a decision regarding the summarization process and classification. The decision manager manipulates and maintains the knowledge base.

3.1.5. Decision Manager. The whole process of this proposed system is monitored by the decision manager. In collaboration with the temporal information manager and knowledge base, the decision manager takes all the decisions regarding the classification and data summarization. It has all the control over data summarization, collection agent, and anomaly detection module.

3.2. Data Summarization Technique. This section is presented with a proposed metric along with four existing metrics in order to serve the purpose of data summarization. Although it is well noted that the data summarization obtains the data as input and outputs the data as well, the output cannot be misinterpreted with information or knowledge. Summary here is simply a precise form of the input data, which is made to use as a replacement for competence reasons. As a result, the whole of the measures dealing with data summarization should be objective in nature. This section is comprised of five objective measures for data summarization, namely, time, conciseness, information loss, interestingness, and intelligibility.

3.2.1. Time. This paper presents a new metric, time, which is used for data summarization. In order to monitor traffic, the time interval is a very significant parameter, as a large number of the users access the network or cloud, and many times a situation arises when the network is not traffic-free. In such conditions, when online traffic is being monitored, it should be managed with the help of time intervals amongst the data groups. Already existing techniques are performed remarkably well, even without concentrating on real-time traffic control. These existing techniques have not considered the retrieval time and randomly retrieving the information from the database. Datasets could be prepared between the particular time intervals, whether it is one week, two weeks, one day, or one hour from the server. From datasets, we can find different numbers of time appearances in various time slots. Based on this, the metric subset has to be broken based on the time intervals; using this approach, data can be retrieved based on the exact data occurrence in various time intervals.

3.2.2. Conciseness. The metric conciseness explains the compact data summary as per the dataset; conciseness is discussed and explained with different terminologies like summarization, compression ratio, and compaction gain in various works and different papers like [11, 34, 44]. In another paper, the conciseness is calculated for a set of records at a specified time interval in accordance with the [25]; the calculation is done in a similar manner as of three

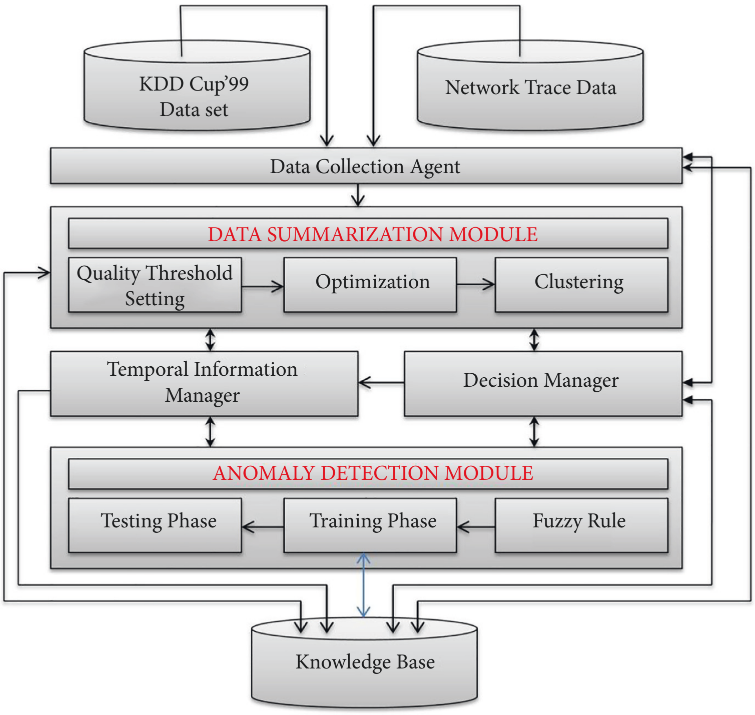

Figure 1: System architecture.

previous approaches, which is expressed as the ratio of the input dataset size to the summarized dataset size. Let M be the number of data points in the input dataset and S be the number of tuples in the summary. Starting time and ending time are denoted by t_1 and t_2.

$$\text{Conciseness}(t_1, t_2) = \frac{M}{S}. \tag{1}$$

3.2.3. Information Loss. Information loss denoted the amount of information loss in the process or in very simple terms the information that is absent in the summary. According to Ha-Thuc et al. [44], the amount of information lost is known as distortion. As per their definition, distortion is the total of the squared Euclidean distance between every data point and centroid of the cluster to which these points belong. It is quite evident from the definition of distortion in this particular work [44]that this method of information loss calculation can only be used in clustering-based summarization, hence not applicable as a general summarization metric. As per Chandola and Kumar [11, 44], information loss is given by an individual summary for a transaction as "the weighted sum of all the features which are absent in the individual summary." Hence, aggregating the information loss of each transaction results in the total information loss of the summary. In this work, the information loss for the set of records is calculated in accordance with the work of Chandola and Kumar [11, 44] for the specific time interval further; the results are normalized. Sum of all the ratios of attributes not present to the total attributes represented per summary gives the information loss. Let S be the number of individual summaries, t_i be the number of different attributes represented by summary i, and l_i be the number of different attributes not present in summary i. Starting time and ending time will be indicated by t_1 and t_2. Then,

$$\text{Information Loss}(t_1, t_2) = \frac{1}{s}\sum_{i=1}^{s}\frac{l_i}{t_i}. \tag{2}$$

3.2.4. Interestingness. The interestingness metric has been discussed extensively in literature; it has been an area of interest for researchers, particularly for finding interesting classification and association rules in data mining [2, 45]. Interestingness is taken up as a broad concept in the literature. It focuses on conciseness, peculiarity, coverage, diversity, reliability, surprisingness, utility, novelty, and actionability. Hoplaros et al. [34] defined interestingness as follows:

$$\text{IRAE}(t_1, t_2) = \sum_{i=1}^{m}\frac{n_i(n_i - 1)}{N(N - 1)}. \tag{3}$$

Let n_i be the derived count attribute of a summary tuple, m be the number of tuples in a summary, and N be the number of total input data points. The starting time and the ending time are represented by t_1 and t_2.

3.2.5. Intelligibility. Intelligibility defines the characteristics of the summary, that is, the level of sense a summary makes out of the data, on account of the number of ANY attributes present in the summary. Every tuple in the summary represents a subset of the input dataset. If a tuple contains the most closely related data, then there will be fewer ANY attributes present. Let m, be the number of tuples in a summary the i tuple having the attributes n_i, and l_i be the number of non-ANY attributes present in tuple i. As per our approach towards information loss and interestingness, intelligibility should be normalized. We define

$$\text{Intelligibility}\left(t_1, t_2\right) = \frac{1}{m} \sum_{i=1}^{m} \frac{l_i}{n_i}. \tag{4}$$

3.3. Temporal Data Summarization Algorithm. This algorithm is comprised of three phases, namely, threshold setting, optimization, and clustering. All these three phases make use of various algorithms proposed by different authors on the basis of different metrics. These algorithms are modified in our proposed method by using temporal constraints.

3.4. Proposed Temporal Data Summarization Method. This proposed real-time data summarization method according to [34] is the combination of four phases. These four phases also contain different algorithms, namely, modified quality threshold summarization algorithm, BUS algorithm, K-means clustering algorithm, and data summarization algorithm. We used the data summarization technique, which was proposed by Hoplaros et al. [34]. A new metric called time is introduced for handling real-time data. This new metric plays a major role in the quality threshold summarization algorithm and summarization algorithm, which are present in phase 1 and phase 2, respectively. This metric uses a modified data summarization method that plays necessary roles for handling real-time data in cloud. The proposed data summarization method has introduced a new metric called time interval in Algorithm 1. This new metric is used in the quality threshold algorithm.

3.5. Modified Synergetic Neural Network. This section presents a new system for anomaly detection in network and cloud communications which is a combination of temporal data summarization and a Modified Synergetic Neural Network. A Modified Synergetic Neural Network is introduced according to [46]. A new layer based on fuzzy rules for all sets of data is introduced in the framework of SNN in the starting and ending time based on Ganapathy et al. [47]. For anomaly detection, fuzzy intervals will be implemented for making efficient and appropriate decisions; when it is compared with time series, these fuzzy rules have been framed on the basis of different time intervals. There is no denying the fact about the dynamic nature of the network traffic; it is not only dynamic in nature but also complex dynamic. The network traffic has shown nonlinear, nonstationary, and complex dynamic behavior. There are many factors involved in describing the behavior or nature of this network at the broader level [1, 48]. Therefore, the network problem is treated as a high-dimensional problem. Network generation in the cloud communication environment is a task that involves many factors, and with the assistance of all these factors or parameters, it could be achieved. There are various factors that dominate the network equally, and the changes over the network are normal, but the network traffic reflects large randomness. Attackers or abnormal users dominate the key factor because when anomalies happen, all the above-discussed factors do not contribute to the network traffic at par. The network at an abnormal state shows strong certainty. Synergetic, says primary, factors contribute to the generation of the order parameter. Randomness and similarity in the cloud communication and network traffic are the characteristics of order parameters in the network. Interdisciplinary science named synergetic demonstrates the organization and formulation of structures and patterns in an open system, which are far from thermodynamics equilibrium. This science focuses on bringing temporal, spatial, and functional structures on macroscopic scales of the various individual factors of a dynamic system. According to synergetic [46], a dynamical system can be expressed as follows:

$$\begin{aligned} q &= -\frac{\partial V}{\partial q+}, \\ q+ &= -\frac{\partial V}{\partial q}, \end{aligned} \tag{5}$$

where q is the system state; $q+$ is the adjoint state of q; V is the potential function of the system; q is the differential of q; and the other is the same on the following equations in this paper.

According to the control principle of synergetic, stable models have a dependency on the unstable models. Whenever there is the process of evolution of the system, the number of certain unstable models keeps on increasing; on the other hand, the number of stable models starts decreasing. When the number of unstable models becomes large, they start behaving as the primary factor in the system, which, as a consequence, transforms the high-dimensional problem into the low dimension problem and the values of the unstable models known as order parameters. These unstable models with the highest original order parameter decide the final state of the system. Synergetic science explains the fundamental building principle for pattern recognition and comes up with an opinion: the process of pattern recognition is pattern formation, which is a top-down approach for analyzing a problem or system. How pattern recognition works are explained here, macroqualitative variation of the system corresponds to the pattern formation, and transformation of the process from testing data to the training data is equivalent to pattern recognition, which reflects the similarity between pattern recognition and pattern formation. MSNN is the technique of pattern recognition when it comes to network traffic anomaly detection in cloud communication based on MSNN. Identified patterns and prototype patterns are presented by testing data and training data, respectively, in the process of MSNN. And the technique to identify the testing data is to map the testing data to some already existing training dataset. The

```
Phase 1: modified quality threshold summarization algorithm.
    Input: dataset $D$, threshold $T$, time interval.//Threshold setting will be different for different time interval summarized data.
    **Output:** cluster centroids $\{C_1, C_2...C_k\}$
    Step 1: initialize the cluster centroid $C = \varnothing$;
    Step 2: initialize the threshold $t = 0$;
    Step 3: initialize the $k_0 = 1$;
    Step 4: choose data item from $D$ and set $I_0 = d$ for the particular time interval.
    Step 5: $(C_t, E_t, <t_1, t_2>) = \text{K-means}\ (D, k_t, I_t)$;
    Step 6: $K_{t+1} = K_t$;
    Step 7: $I_{t+1} = C_t$;
    Step 8: for i $\longleftarrow$ 1 to $k_t$ do
    Step 9: if $Et_i$ less than $T$, then
    Step 10: $C = C\ \{C_{ti}\}$
    Step 11: remove cluster $i$ out of D
    Step 12: $K_{t+1} = K_{\ t+1} - 1$
    Step 13: $l_{t+1} = l_{t+1} - \{C_{ti}\}$
    Step 14: end
    Step 15: if $D$ equals $\varnothing$, then
    Step 16: return centroids $\{C_1; C_2...C_k\}$;
    Step 17: randomly choose a data point $d$ approximately close to the centroid of the largest cluster;
    Step 18: insert $d$ to $I_{t+1}$;
    Step 19: $k_{t+1} = k_{t+1} + 1$;
    Step 20: $t = t + 1$;
    Step 21: go to 5;
Phase 2: apply BUS algorithm
Phase 3: apply K-means clustering
Phase 4: apply data summarization algorithm that incorporates all metrics [29].
```

Algorithm 1: Temporal data summarization method.

identification of testing network traffic patterns q can be explained as a dynamic process in cloud communications. After mapping the initial testing data $q(0)$ from intermediate state $q(t_1,t_2)$ to a training data vector v_k, the training data vector v_k is most near to $q(0)$. The process can be described as $q(0) \longrightarrow q(t_1,t_2) \longrightarrow v_k$, where $q(0)$ is the testing network traffic data, v_k is the stored normal or abnormal traffic network traffic, and the intermediate state $q(t_1, t_2)$ is the order parameter ω_k. To be precise, this process can be represented by a dynamic equation (2). The assumption is made that the number of the training data vectors is M and the dimension of the training data vector is N. For maintaining the linear independence of the M training data vector, $M \leq N$ is required.

$$q = \sum_{k=1}^{M} \gamma_k V_k \left(V_k^{+} q\right) - B \sum_{k=1}^{M} \sum_{k1=1,k=k1}^{M} \left(V_k^{+} q\right) 2 \left(V_k^{+} q\right) V_k - C\left(q^{+} q\right) q + F\left(t_{1,} t_2\right), \tag{6}$$

where q is the testing network traffic data vector in cloud communications with the original input data value q (0). Scalar value γ_k is the attention parameter because when it is positive, only testing data can be identified. F (t_1, t_2) is the fluctuation factor for the particular record login and logout time of the network traffic in cloud communications and can be ignored. Scalar values B and C are specified coefficients and must be greater than 0. v_k is the training data vector, $v_k = (v_{k,1}, v_{k,2}, \ldots, v_{k,N})^{\mathrm{T}}$, where superscript T is vector transposition. $v_k{}^{+}$ is the adjoint vector of v_k.

$$V_k^{+} V_{k,} = \delta_{k,k'} = \begin{cases} 1, & k = k', \\ 0, & kk'. \end{cases} \tag{7}$$

v_k should be prepared with normalization and zero mean:

$$\sum_{l=1}^{N} V_{k,1} = 0, \sqrt{\left(\sum_{l=1}^{N} V^2{}_{k,1} = 1\right)}. \tag{8}$$

In order to reduce the dimensionality of the system, the order parameters γ_k are features extracted from vector q. q can be represented by the order parameters γ_k, a training data vector v_k, and the remaining vector w:

$$= \sum_{k=1}^{M} \gamma_k V_k + W, V_k^{+} W = 0. \tag{9}$$

The adjoint vector of q is defined as follows:

$$q^{+} = \sum_{k=1}^{M} \gamma_k V_k + W^{+}, W_k^{+} V_k = 0. \tag{10}$$

There is a relationship:

$$V_k^{+} q = q^{+} V_k. \tag{11}$$

Typing (5) into (7), according to the orthogonal relationship, the order parameter $_{\mathrm{k}}$ is defined as follows:

$$\gamma_k = V_k^{+}. \tag{12}$$

Style described in (2) is a powerful dynamic equation. If we neglect the fluctuation factor F(t_1, t_2) during the particular time interval of the network traffic in cloud communications, according to equations (1) and (2), the potential function can be described as follows:

$$V = -\frac{1}{2}\sum_{k=1}^{M}\gamma_k\left(V_k^{+}q\right)^2 + \frac{1}{4}B\sum_{k=1}^{M}\sum_{k'=1,k\neq k}^{M}\left(V_k^{+}q\right)^2 + \frac{1}{4}C\left(\sum_{k=1}^{M}\left(V_k^{+}q\right)^2\right). \tag{13}$$

According to equations (1), (2), and (8), correspondingly, the dynamic equations and the potential function described by the order parameter are as follows:

$$\omega_k = \gamma_k\omega_k - B\sum_{k'=1,k\neq k'}^{M}\omega_k^2\omega_k - C\left(\sum_{k'=1}^{M}\omega_k^2\right)\omega_k, \tag{14}$$

$$V = -\frac{1}{2}\sum_{k'=1}^{M}\gamma_k\omega_k^2 + \frac{1}{4}B\sum_{k=1}^{M}\sum_{k'=1,k'\neq k}^{M}\omega_{k'}^2\omega_k^2 + \frac{1}{4}C\left(\sum_{k'}^{M}\omega_k^2\right)^2. \tag{15}$$

At the lowest potential energy of a system, the system controlled by the order parameters reaches the most stable state. Here, the order parameters attain their extreme value. The stable state of the network system is described by the following formula:

$$\omega_k = 0, \quad 0 \leq k \leq M. \tag{16}$$

That is,

$$\omega_k = \gamma_k\omega_k - B\sum_{k'\neq k}\omega_k^2\omega_k - C\left(\sum_{k'=1}^{M}\omega_k^2\right)\omega_k = 0. \tag{17}$$

If we define

$$D = (B + C)\sum_{k'=1}^{M}\omega_k^2, \tag{18}$$

then the following equations can be inferred from (10) and (12):

$$\begin{gathered}\omega_k = \omega_k\left(\gamma_k - D + B\omega_k^2\right),\\ \omega_k\left(\gamma_k - D + B\omega_k^2\right) = 0.\end{gathered} \tag{19}$$

As per the (15), there are four layers in the MSNN architecture in Figure 2. The top layer is the input layer in which unit j receives component $q_j(0)$ of need recognized pattern vectors original value $q(0)$. All the order parameter components form the middle layer, where the order parameter ω_k is obtained by summing all angle indexes j through each input value $q_j(0)$, multiplying its joint unit $v^{+}{}_{kj}$. The active order parameter ω_k recognizes the special training data chosen by the angle index k. According to the dynamical equation, MSNN will be evolved to the end state that only one of the order parameters survives, and q_j is obtained through reciprocity and competition of D. The down layer is the output layer in which output pattern can be expressed as $q_j(t_1, t_2) = \sum_k \omega_k(t_1, t_2)V_{kj}$, where q_j is active of output unit j and ω_k is the end of the state of the middle layer. There is $\omega_k = 1$ if $k = k_0$; otherwise, $\omega_k = 0$. v_{kj} is the component of j of the training data vector.

Time series of the network traffic in cloud communications are represented by $y_1, \ldots y_N$; these are sampled in bytes or bits or packets per time unit. The fuzzy temporal information manager of the proposed system MSNN deals with the fuzzy time interval for detecting anomalies. Fuzzy time intervals are set up as shown in Figure 3. Normal and abnormal network traffic are set for constructing a training dataset that may contain M components from a specific time interval. Network traffic fuzzy time intervals also share the same size N. This proposed anomaly detection method is designed in such a way to make a distinction between normal and abnormal network traffic.

The process of anomaly detection includes two stages as shown in Figure 4: the training stage is to learn the training data of the normal and abnormal network traffic, and the testing stage is to detect the network traffic anomaly. The detailed detection steps are given as follows.

3.5.1. *The Training Stage*

(a) Choose the training data vectors { $y_1, \ldots y_{\mathrm{N}}$ } from the trained dataset for specific time intervals

(b) Deal with the training pattern vectors { $y_1, \ldots y_{\mathrm{N}}$ } with normalization and zero mean and then compute the training data vectors v_k for the particular time interval records

(c) Compute the corresponding adjoint $\mathrm{v}^{+}{}_{\mathrm{k}}$ of the training data vectors v_k at particular time intervals

3.5.2. *The Testing Stage*

(a) Test on the testing data vector $q(0)$ consisted of the testing network traffic data in cloud communications dealt with normalization and zero mean.

(b) Compute the corresponding order parameter ω_k of each training data according to (8), which is in the particular time interval.

(c) Evolve by the following order parameter dynamic equation (17) until the order parameters start converging to a specific training data and then to the specific training data vector $q(0)$ in a particular time interval. Thus, the processes of the network traffic anomaly detection in cloud communications based on MSNN have been completed.

$$\omega_k(n+1) - \omega_k(n) = \beta\left(\gamma_k - D + B\omega_k^2(n)\right)\omega_k(n), \tag{20}$$

where β is the iterative step.

The steps of the training stage are as follows:

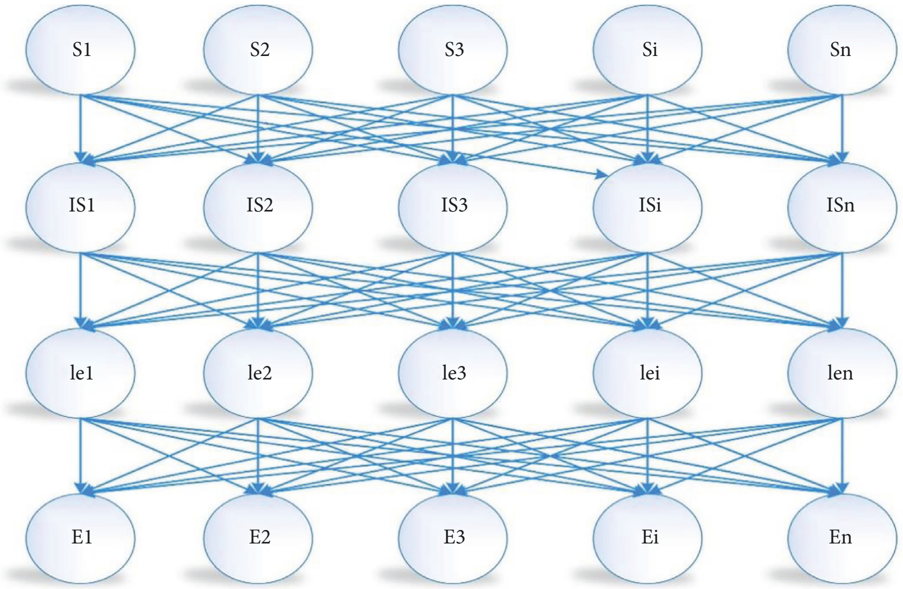

Figure 2: The framework of MSSN.

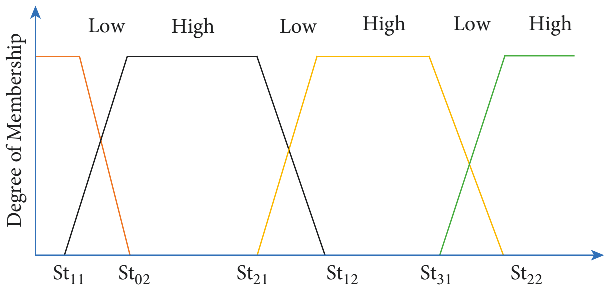

Figure 3: Fuzzy time interval.

(1) Consider the fuzzy time intervals $y_1,\ldots y_N$ of the training data, for each time interval between t_1 and t_2, and construct the vector set { $Y_t^P = (y_{t-p}+_1,\ldots,y_t)$ $|t = 1,\ldots,N-p+1\}$, which is the time interval window

(2) Obtain the set of the state variable $\{x_t\}$ and the control variables $\{u_t\}$ and $\{v_t\}$ based on normalized features extracted from each vector Y_{pt}

(3) Compute the parameters a and b of the cusp catastrophe model using the series $\{x_t\}$, $\{u_t\}$, and $\{v_t\}$

In the testing stage, the main steps are as follows:

(1) Construct the vector Y_t^P (with the same time window Win_p in the training stage) of the testing data at the observed time I, which is labeled as observed point P_i.

(2) Extract the selected normalized features to present the state variable x_i and control variables u_i and v_i.

(3) Compute the catastrophe distance between the observed point $P_i(x_i, u_i, v_i)$ and the bifurcation set $G(x, u, v)$, labeled as D_p. The catastrophe distance D_p is defined as follows: assuming that $P_i(x_i, u_i, v_i)$ is an observed point in the testing data of the traffic in cloud communications and $P_t(x_t, u_t, v_t)$ is a point of the equilibrium surface $G(x, u, v)$, the distance between two points $P_i(x_i, u_i, v_i)$ and $P_t(x_t, u_t, v_t)$, labeled as $D_E(P_i, P_t)$ is computed by the Minkowski distance. The catastrophe distance D_p between the observed point $P_i(x_i, u_i, v_i)$ and the equilibrium surface $G(x, u, v)$ is defined as

$$D_P\left(P_i, G(x,u,v)\right) = \min_{Pt\in G(x,u,v)}\left\{D_E\left(P_i, P_t\right)\right\}. \qquad (21)$$

As catastrophe distance D_p is more than a threshold, then anomaly can be claimed at the observing point P_i (x_i, u_i, v_i). The threshold is obtained by training.

4. Results and Discussion

In this section, the different experimental results carried out for data summarization and anomaly detection have been discussed. For data summarization, different experiments by using the four different metrics in algorithms called quality threshold algorithms (BUS algorithm and K-means clustering algorithm) have been performed. Finally, these algorithms are combined to propose Modified Synergetic Neural Network for providing better classification accuracy.

The proposed method is the combination of the very well-known and state-of-the-art techniques and, while using the strength of the neural network, gives a very good performance.

4.1. Datasets. The dataset for this experiment is taken from the third International Knowledge Discovery and Data Mining Tools Competition (KDD Cup 99) [18, 49]. Every connection record is characterized by 41 attributes. All these attributes are both discrete and continuous in nature; these variables are drastically varying to each other on the basis of statistical distributions, turning it to be a challenging task for intrusion detection [50].

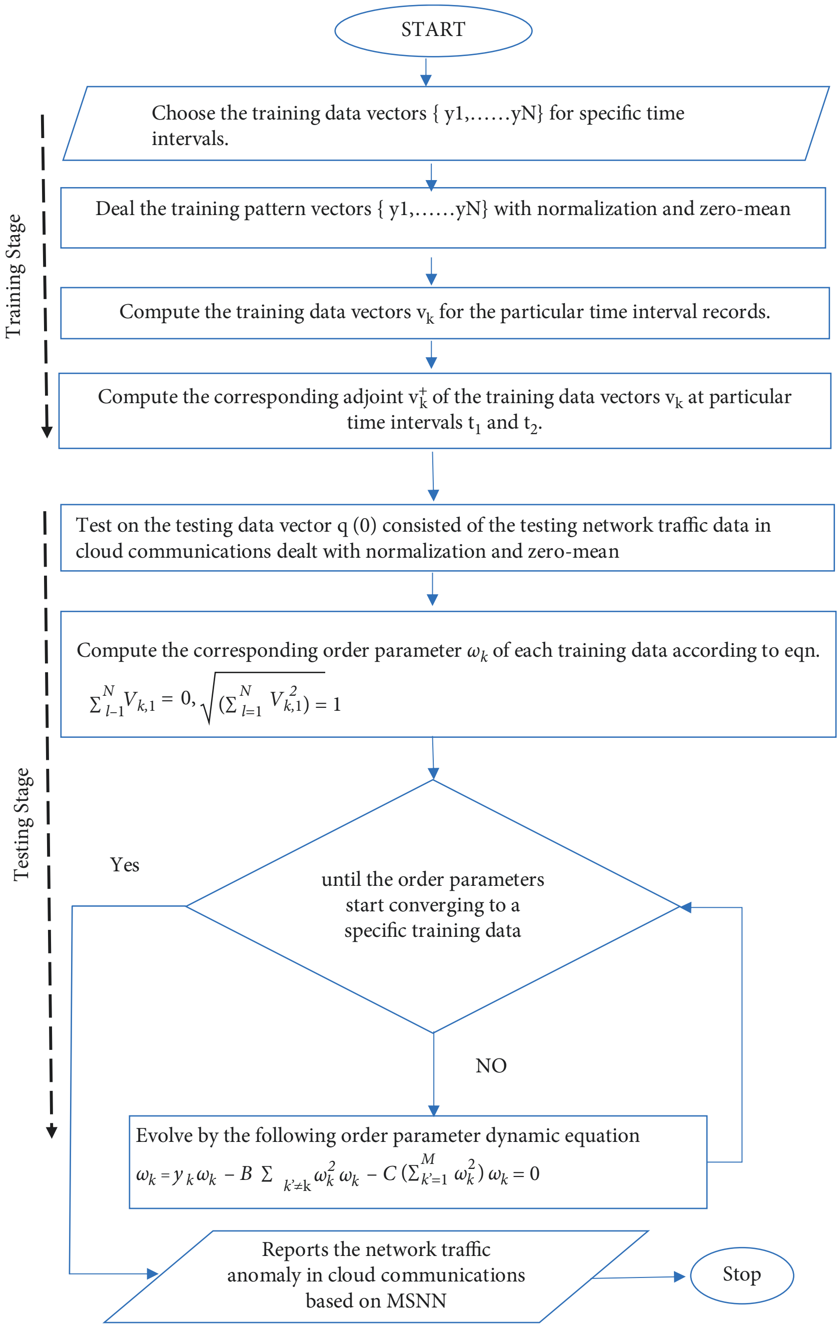
START
Choose the training data vectors { y1,......yN} for specific time intervals.
Deal the training pattern vectors { y1,......yN} with normalization and zero-mean
Compute the training data vectors v_k for the particular time interval records.
Compute the corresponding adjoint v_k^+ of the training data vectors v_k at particular time intervals t_1 and t_2.
Training Stage
Test on the testing data vector q (0) consisted of the testing network traffic data in cloud communications dealt with normalization and zero-mean
Compute the corresponding order parameter ω_k of each training data according to eqn.
$\Sigma_{l-1}^{N} V_{k,1} = 0, \sqrt{(\Sigma_{l=1}^{N} V_{k,1}^{2})} = 1$
until the order parameters start converging to a specific training data
Yes
NO
Evolve by the following order parameter dynamic equation
$\omega_k = y_k \omega_k - B \Sigma_{k' \neq k} \omega_k^2 \omega_k - C (\Sigma_{k'=1}^{M} \omega_k^2) \omega_k = 0$
Testing Stage
Reports the network traffic anomaly in cloud communications based on MSNN
Stop

Figure 4: MSNN process.

This dataset contains 5 million network connection records like password guess, land attack, Neptune attack, and port scan. The twenty-two categories of attack are from the following four classes: DOS, R2L, U2R, and probe. The 41 features explain the fundamental information regarding network packet, network traffic, host traffic, and content information. Each record has 5 class labels, namely, normal, probe, DOS, R2L, and U2R. It has 391458 DOS attack records, 52 U2R attack records, 4107 probe attack records, 1126 R2L attack records, and 97278 normal records only in 10 percent of this dataset.

4.2. Experimental Results. Table 2 shows the performance of quality threshold data summarization. From this table, we can see the different four metrics such as conciseness, information loss, interestingness, and intelligibility values when considering the different threshold values with a number of clusters considered for summarization of data.

Table 3 shows the performance of BUS in data summarization. Here, the different four metrics such as conciseness, information loss, interestingness, and intelligibility values were obtained when considering the different kinds of summary sizes for summarization of data.

Table 2: Results for quality threshold summarization phase on the KDD Cup'99 dataset.

Threshold	Clusters	Conciseness	Information loss	Interestingness	Intelligibility
15000	12	10493.31	0.9891	0.14876	0.2712
10000	16	7869.52	0.9893	0.088	0.29513
5000	24	5243.97	0.98267	0.06621	0.3068
2500	43	2925.04	0.97348	0.04912	0.3579
1000	75	1675.42	0.9651	0.03017	0.3745
500	124	1010.128	0.9567	0.02197	0.40649
250	187	669.47	0.94652	0.01662	0.4392
100	297	420.512	0.9262	0.0119	0.46725
50	437	284.17	0.89734	0.0065	0.48343

Table 3: Results for BUS phase on the KDD Cup'99 dataset.

Summary size	Conciseness	Information loss	Interestingness	Intelligibility
12	10493.31	0.9891	0.2225	0.2624
16	7869.52	0.9893	0.09351	0.29132
24	5243.97	0.98267	0.12832	0.3265
43	2925.04	0.97348	0.01284	0.3768
75	1675.42	0.9651	0.095721	0.2763
124	1010.128	0.9567	0.03672	0.4216
187	669.47	0.94652	0.07419	0.4552
297	420.512	0.9262	0.2032	0.26821
437	284.17	0.89734	0.05071	0.52242

Table 4: Results for K-means clustering phase on the KDD Cup'99 dataset.

Clusters	Conciseness	Information loss	Interestingness	Intelligibility
12	10497.75	0.99242	0.11231	0.30672
16	7873.3125	0.98968	0.08921	0.34128
24	5248.875	0.98151	0.0762	0.39312
43	2929.6046	0.9778	0.0523	0.43345
75	1679.64	0.96495	0.0348	0.49213
124	1015.9112	0.95392	0.0217	0.53279
187	673.6524	0.94646	0.0108	0.5725
297	424.1515	0.93889	0.0087	0.6218
437	288.2677	0.93130	0.0056	0.6617

Table 4 shows the performance of the K-means clustering algorithm in data summarization. From this table, we can see the different four metrics such as conciseness, information loss, interestingness, and intelligibility values during the consideration of different thresholds for the different number of clusters for summarization of data.

Figure 5 shows the comparison of performance analysis between the existing quality threshold algorithm and the combination of the proposed MSNN with the quality threshold algorithm. Figure 2 explains that the MSNN framework with quality algorithm has outperformed the existing quality threshold algorithms in terms of classification accuracy. The proposed anomaly detection method provides better anomaly detection accuracy significantly while considering different thresholds for anomaly detection.

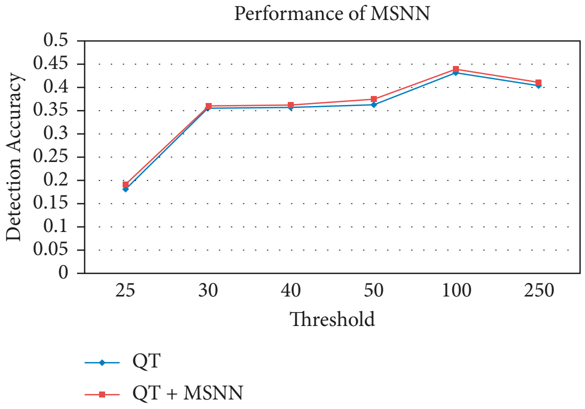

Figure 5: Comparison of performance analysis between QT and MSNN with QT algorithm.

Figure 6 shows the comparison of performance analysis between the existing K-means clustering algorithm and the combination of the proposed MSNN with the K-means clustering algorithm.

From Figure 6, it can be observed that the classification accuracy of the proposed MSNN with the K-means clustering algorithm is better than the existing K-means clustering algorithm. The proposed anomaly detection method provides better anomaly detection accuracy quite significantly during the consideration of different thresholds for anomaly detection.

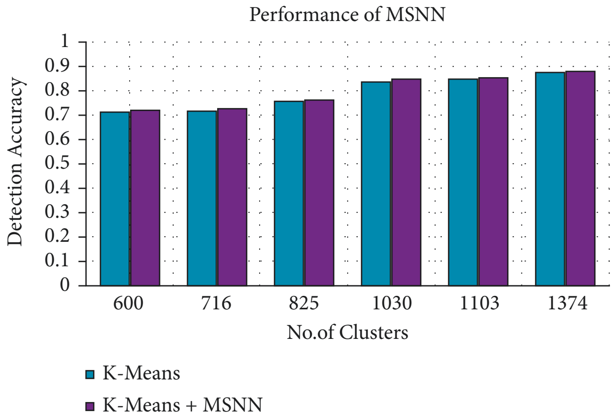

Figure 6: Comparison of performance analysis between K-means and MSNN with K-means.

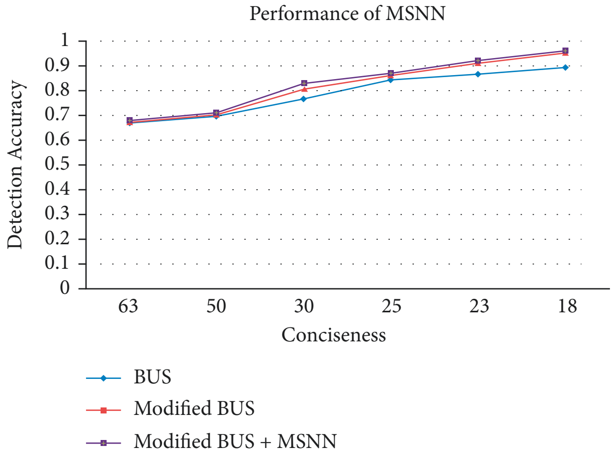

Figure 7: Comparison of performance analysis between BUS, modified BUS, and MSNN with modified BUS.

Figure 7 shows the comparison of performance analysis for the existing BUS algorithm, modified BUS algorithm, and the proposed MSNN with modified BUS algorithm. From this figure, it can be observed that the classification accuracy of the proposed MSNN with modified BUS algorithm is better than the existing BUS algorithm and modified BUS algorithm. The proposed anomaly detection method provides better anomaly detection accuracy significantly during the consideration of different thresholds for anomaly detection.

5. Conclusions and Future Work

This paper presents *i*ReTADS, an intelligent real-time anomaly detection technique. As a part of it, a new metric time interval is introduced for data summarization in addition to four existing metrics for the same purpose. We proposed a novel neural network framework with fuzzy temporal features comprised of four layers, and this handles the fuzzy time interval for classification. Finally, the demonstration for classification of summarized datasets using the proposed neural network was carried out for assessing its effectiveness. Time taken while using summarized data can be a fraction of the total time taken over the original dataset, getting approximately the same results, which can save time in a critical application. These methods should be optimized and parallelized. However, the system is tested with large datasets of network traffic, which revealed another necessity. More focus should be given to real-time anomaly detection, and the research efforts will be directed to stream data summarization and anomaly detection methods. In future works, we will explore a new effective real-time anomaly detection method using soft computing techniques.

Data Availability

The data used in this study are available at https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. You and K. Chandra, “Time series models for internet data traffic,” in *Proceedings of the 24th Conference on Local Computer Networks. LCN’99*, pp. 164–171, Lowell, MA, USA, October 1999.

[2] G. Kim, S. Lee, and S. Kim, “A novel hybrid intrusion detection method integrating anomaly detection with misuse detection,” *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2014.

[3] C. Zins, “Conceptual approaches for defining data, information, and knowledge,” *Journal of the American Society for Information Science and Technology*, vol. 58, 2007.

[4] J. Song, Z. Han, W. Wang, J. Chen, and Y. Liu, “A new secure arrangement for privacy-preserving data collection,” *Computer Standards & Interfaces*, vol. 80, Article ID 103582, 2022.

[5] Q. Tao, G. Xiaohong, L. Wei, and W. Pinghui, “Monitoring abnormal traffic flows based on independent component analysis,” in *Proceeding of the 2009 IEEE International Conference on Communications*, Dresden, Germany, June 2009.

[6] K. Keys, D. Moore, and C. Estan, “A robust system for accurate real-time summaries of internet traffic,” *ACM SIGMETRICS - Performance Evaluation Review*, vol. 33, no. 1, pp. 85–96, 2005.

[7] Z. Lv, L. Wang, Z. Guan et al., “An optimizing and differentially private clustering algorithm for mixed data in sdn-based smart grid,” *IEEE Access*, vol. 7, Article ID 45773, 2019.

[8] H. Patel, D. Singh Rajput, G. Thippa Reddy, C. Iwendi, A. Kashif Bashir, and O. Jo, “A review on classification of imbalanced data for wireless sensor networks,” *International Journal of Distributed Sensor Networks*, vol. 16, no. 4, 2020.

[9] D. S. Rajput, S. M. Basha, Q. Xin et al., “Providing diagnosis on diabetes using cloud computing environment to the people living in rural areas of India,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2021.

[10] P. Karras, “Multiplicative synopses for relative-error metrics,” in *Proceedings of the Twelveth International Conference on Extending Database Technology: Advances in Database Technology*, pp. 756–767, Saint Petersburg, Russia, March 2009.

[11] V. Chandola and V. Kumar, “Summarization - compressing data into an informative representation,” *Knowledge and Information Systems*, vol. 12, no. 3, pp. 355–378, 2007.

[12] R. Saint-Paul, G. Raschia, and N. Mouaddib, “General Purpose Database Summarization,” in *Proceedings of the 31st International Conference on Very Large Data Bases*, pp. 733–744, Citeseer, Trondheim Norway, September 2005.

[13] A. Singhal, *Data warehousing and data mining techniques for cyber security*, Springer Science & Business Media, vol. 31, New York, NY, USA, , 2007.

[14] R. Zhu, “Intelligent rate control for supporting real-time traffic in wlan mesh networks,” *Journal of Network and Computer Applications*, vol. 34, no. 5, pp. 1449–1458, 2011.

[15] R. Agrawal, T. Imielinski, and A. Swami, “Mining association rules between sets of items in large databases,” in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 207–216, Washington D.C. USA, May 1993.

[16] L. Yu and F. Ren, “A Study on Cross-Language Text Summarization Using Supervised Methods,” in *Proceedings of the 2009 International Conference on Natural Language Processing and Knowledge Engineering*, pp. 1–7, IEEE, Dalian, China, September 2009.

[17] D. S. Rajput and R. Gour, “An IoT framework for healthcare monitoring systems,” *International Journal of Computer Science and Information Security*, vol. 14, no. 5, p. 451, 2016.

[18] O. Depren, M. Topallar, E. Anarim, and M. K. Ciliz, “An intelligent intrusion detection system (ids) for anomaly and misuse detection in computer networks,” *Expert Systems with Applications*, vol. 29, no. 4, pp. 713–722, 2005.

[19] Y. Xiang, R. Jin, D. Fuhry, and F. F. Dragan, “Summarizing transactional databases with overlapped hyperrectangles,” *Data Mining and Knowledge Discovery*, vol. 23, no. 2, pp. 215–251, 2011.

[20] S. Barbhuiya, Z. Papazachos, P. Kilpatrick, and D. S. Nikolopoulos, “Rads: Real-time anomaly detection system for cloud data centres,” 2018, https://arxiv.org/abs/1811.04481.

[21] R. R. Reddy, Y. Ramadevi, and K. Sunitha, “Real time anomaly detection using ensembles,” in *Proceedings of the 2014 International Conference on Information Science & Applications (ICISA)*, pp. 1–4, IEEE, Seoul, Republic of Korea, May 2014.

[22] Y. Du, J. Liu, F. Liu, and L. Chen, “A real-time anomalies detection system based on streaming technology,”vol. 2, pp. 275–279, in *Proceedings of the Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*, vol. 2, IEEE, Hangzhou, China, August 2014.

[23] Y. Djenouri, A. Belhadi, J. C.-W. Lin, and A. Cano, “Adapted K-nearest neighbors for detecting anomalies on spatio-temporal traffic flow,” *IEEE Access*, vol. 7, Article ID 10015, 2019.

[24] N. Ding, H. Ma, H. Gao, Y. Ma, and G. Tan, “Real-time anomaly detection based on long short-term memory and Gaussian mixture model,” *Computers & Electrical Engineering*, vol. 79, Article ID 106458, 2019.

[25] W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, “Malware traffic classification using convolutional neural network for representation learning,” in *Proceedings of the International Conference on Information Networking (ICOIN)*, pp. 712–717, IEEE, Da Nang, Vietnam, January 2017.

[26] B. Agrawal, T. Wiktorski, and C. Rong, “Adaptive real-time anomaly detection in cloud infrastructures,” *Concurrency and Computation: Practice and Experience*, vol. 29, no. 24, Article ID e4193, 2017.

[27] S. Lee and H. K. Kim, “Adsas: comprehensive real-time anomaly detection system,” in *Proceedings of the International Workshop on Information Security Applications*, pp. 29–41, Springer, Jeju Island, Republic of Korea, August 2018.

[28] R. A. Ariyaluran Habeeb, F. Nasaruddin, A. Gani, I. A. Targio Hashem, E. Ahmed, and M. Imran, “Real-time big data processing for anomaly detection: a survey,” *International Journal of Information Management*, vol. 45, pp. 289–307, 2019.

[29] J. F. Olson and K. M. Carley, “Summarization and Information Loss in Network Analysis,” in *Proceedings of the Workshop on Link Analysis, Counter-terrorism, and Security Held in Conjunction with the SIAM International Conference on Data Mining (SDM)*, Citeseer, Atlanta, Georgia, USA, 2008.

[30] E. Tanin and M. E. Ali, “Hierarchical Data Summarization,” *Encyclopedia of Database Systems*, Boston, MA, USA, Article ID Springer, 2009.

[31] G. Mao, “A real-time loss performance monitoring scheme,” *Computer Communications*, vol. 28, no. 2, pp. 150–161, 2005.

[32] J. Lopes, J. Bento, E. Huang, C. Antoniou, and M. Ben-Akiva, “Traffic and mobility data collection for real-time applications,” in *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems*, pp. 216–223, IEEE, Funchal, Portugal, September 2010.

[33] T. Vafeiadis, A. Papanikolaou, C. Ilioudis, and S. Charchalakis, “Real-time network data analysis using time series models,” *Simulation Modelling Practice and Theory*, vol. 29, pp. 173–180, 2012.

[34] D. Hoplaros, Z. Tari, and I. Khalil, “Data summarization for network traffic monitoring,” *Journal of Network and Computer Applications*, vol. 37, pp. 194–205, 2014.

[35] S. Mascaro, A. E. Nicholso, and K. B. Korb, “Anomaly detection in vessel tracks using bayesian networks,” *International Journal of Approximate Reasoning*, vol. 55, no. 1, pp. 84–98, 2014.

[36] W. Xiong, H. Hu, N. Xiong et al., “Anomaly secure detection methods by analyzing dynamic characteristics of the network traffic in cloud communications,” *Information Sciences*, vol. 258, pp. 403–415, 2014.

[37] Z. Li, H. Fang, and L. Xia, “Increasing mapping based hidden Markov model for dynamic process monitoring and diagnosis,” *Expert Systems with Applications*, vol. 41, no. 2, pp. 744–751, 2014.

[38] Q. Ma, C. Sun, B. Cui, and X. Jin, “A novel model for anomaly detection in network traffic based on kernel support vector machine,” *Computers & Security*, vol. 104, Article ID 102215, 2021.

[39] A. Alnafessah and G. Casale, “Artificial neural networks based techniques for anomaly detection in Apache Spark,” *Cluster Computing*, vol. 23, no. 2, pp. 1345–1360, 2020.

[40] W. Ullah, A. Ullah, I. U. Haq, K. Muhammad, M. Sajjad, and S. W. Baik, “Cnn features with bi-directional lstm for real-time anomaly detection in surveillance networks,” *Multimedia Tools and Applications*, vol. 80, pp. 1–17, 2020.

[41] L. Erhan, M. Ndubuaku, M. Di Mauro, and W. Song, M. Chen, G. Fortino, O. Bagdasar, and A. Liotta, Smart anomaly detection in sensor systems: a multi-perspective review,” *Information Fusion*, vol. 67, 2020.

[42] S. Garg, K. Kaur, S. Batra, G. Kaddoum, N. Kumar, and A. Boukerche, “A multi-stage anomaly detection scheme for augmenting the security in iot-enabled applications,” *Future Generation Computer Systems*, vol. 104, pp. 105–118, 2020.

[43] W. Wang, H. Xu, M. Alazab, T. Reddy Gadekallu, Z. han, and C. su, “blockchain-based reliable and efficient certificateless

signature for IIoT devices," *Journal of latex class files*, vol. 14, no. 8, 2015.

[44] V. Ha-Thuc, D. C. Nguyen, and P. Srinivasan, "A quality-threshold data summarization algorithm," in *Proceedings of the IEEE International Conference on Research, Innovation and Vision for the Future in Computing and Com- Munication Technologies*, pp. 240–246, IEEE, Ho Chi Minh City, Vietnam, July 2008.

[45] K. McGarry, "A survey of interestingness measures for knowledge discov- ery," *The Knowledge Engineering Review*, vol. 20, 2005.

[46] H. Haken, *Synergetic computers and cognition: A top-down approach to neural nets*, Vol. 50, Springer Science & Business Media, , New York, NY, USA, 2004.

[47] S. Ganapathy, R. Sethukkarasi, P. Yogesh, P. Vijayakumar, and A. Kannan, "An intelligent temporal pattern classification system using fuzzy temporal rules and particle swarm optimization," *Sadhana - Academy Proceedings in Engineering Sciences*, vol. 39, 2014.

[48] S. Belarouci and M. A. Chikh, "Medical imbalanced data classification," *Advances in Science, Technology and Engineering Systems*, vol. 2, 2017.

[49] Q. Yang and X. Wu, "10 challenging problems in data mining research," *International Journal of Information Technology and Decision Making*, vol. 5, no. 4, pp. 597–604, 2006.

[50] A. Rehman, S. U. Rehman, M. Khan, M. Alazab, and G. T. Reddy, "Canin- telliids: detecting in-vehicle intrusion attacks on a controller area network using cnn and attention-based gru," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1456–1466, 2021.