Reducing Significances of Mesh Sensors Technologies through Dimensionality Reduction Algorithm

Ruhul Amin

Senior Data Entry Control Operator (IT), ED-Maintenance Office, Bangladesh Bank (Head Office), Dhaka, **BANGLADESH**

*Corresponding Contact:
Email: ruhulaminshuzon@gmail.com

ABSTRACT

In today's world, the breadth of real-time applications and networks is not limited to business and social activities. They are expanding as a field to provide improved and competitive settings for a variety of activities such as home, health, and commercial procedures. Data analytic method is used to maintain network accessibility as well as the robustness of expert services. It is necessary to clean up the data in order to reduce the computational complexity of extracting and pre-processing models. Because present approaches are sophisticated, they necessitate large computations. To this effect, the objective is to deploy a machine learning algorithm – "cuckoo search algorithm" for dimensionality reduction problems in data extraction for IoTs application. The cuckoo search-based feature extraction algorithm is a mutant algorithm that organizes itself depending on the unpredictable amount of input and generates a new and improved feature space. After the cuckoo search-based feature extraction is implemented, a few test benchmarks are provided to assess the performance of mutant cuckoo search algorithms. As a result of the low-dimensional data, classification accuracy is improved while complexity and expense are lowered.

Key words:

Dimensionality Reduction Algorithm, Cuckoo search, IoT, Classification accuracy

INTRODUCTION

As the innovative technologies continue to improve from a level to another, the next generation of computers and the internet may reduce the customary process of connection get to the end-users through the implementation of interlinking 'smart' devices called mesh sensors. This does not mean that the internet will be replaced rather it will be an advancement to the internet connection and infrastructure by integrating physical devices with transmission and processing technologies providing a vast variety of applications and services in a more rapid, consistent, and accessible manner. This kind of innovation will

lead towards pervasive computing such that each device is embedded with communication proficiency and microprocessors (Ganapathy, 2019). It will enable physical devices to 'smart' and allow them to integrate with global cyber-physical models. This development will make way for novel revolutions and openings in the ICT sector, delivering fresh applications and services by linking virtual and physical devices. It is evolving as a development in a manner that most of the surrounding devices will be connected to the internet (network) in different forms. This is a swing from the traditional internet techniques to the internet of linking physical devices that have a human interface (Vadlamudi, 2016).

The IoTs enable connection with anybody, at any habitation, and at any time. Innovative technologies are designing the world where the whole thing will be linking together through a network. With this innovation, things will be identified easily and in a more unique manner, functioning in smooth environments with the assistance of intellectual edges to link and interact within the corporeal domain. Unified smart devices are equipped and heterogeneous with smart objects such as actuators and sensors. Data captured or sensed by these devices is immense in volume and can be characterized as "big data" (Ganapathy, 2018). Data extraction and knowledge discovery approaches such as clustering, classification, and pattern analysis are recommended for IoT by different scholars to offer the right value services and environment to users. The extended data volume that is "big data" generated by smart devices with noisy, irrelevant, high dimensions, and laid off features produce an immense search space. If extraction methods are used on such fuzzy and rough data, it can increase cost, depreciate performance, and computation of extraction algorithms. However, pre-processing approaches are needed to map the raw dataset onto a new abridged feature subset, which can epitomize the original interplanetary with high precision (Patel and Patel, 2016).

Problem Statement

The scope of applications and networks in real-time is not restricted to enterprise and social activities in the present era. They are evolving as an extended field to deliver enhanced and competitive situations for different activities consisting of home, health, and business procedures. To sustain as well as preserving network accessibility and robustness of expert services, data analytic methods are essential. Data clean-up to eliminate the computation complexity of extracting and pre-processing models is needed. The current methods are complex, consequently comprise outsized computations (Yoo, 2010).

Objectives of the Study

The aim of this study is restricted to the deployment of a machine learning algorithm – "cuckoo search algorithm" for dimensionality reduction problems in data extraction for IoTs application. This study is more correlated to feature mining as likened to characteristic selection. The study will base its investigation on the transformation of current features by utilizing "Cuckoo search algorithms" for data extraction. Between open-air and indoor data services, our concentration is on open-air data services. Open-air data are of different types including images, video, text, audio, and hypertext, but our report focuses on numerical regression for the planned model (Chiroma et al., 2017).

These pieces of evidence reveal a research gap, and the planned research project is built on the following goals:

(i). Outlining unaddressed difficulties related to "big data" analysis for the internet of things, as well as behavioral analysis and the restrictions of present methodologies.

(ii) Emphasizing the significance of preparatory processing prototypes for self-organized networks in terms of reducing complication.

(iii) Presenting a methodology for reducing the dimensionality in the internet of things, built on a regressive evolutionary algorithms method, that is the cuckoo search algorithm.

(iv) Examining the performance of the proposed system.

LITERATURE REVIEW

Dimensionality Reduction

Data gathered by different sensing devices such as RFIDs and sensors in an internet of things network involves contradictory, inappropriate, and noisy data about something like a dataset. Regular extraction methods are used to increase the effectiveness and performance of internet of things services. Data is taken at a constant rate, increasing the density of extraction algorithms, considering the number of data dimensions, classifiers are required (Karkouch et al., 2016). An input dataset's data dimension represents the various characteristics or attributes that describe it.

$$L = \{f_{i=1}^n\}f \rightarrow (L^D = n)$$

Where D denote dimension, and f_i provides the n attributes that describe the dataset L.

High-dimensional datasets can impair the precision of extraction processes, as previously indicated. For the transfer of high-dimensional data into lesser dimensions, several preprocessing procedures are required as shown in Figure 1. Low-dimensional data will improve classification precision while reducing complications and cost. IoT data extraction is not the same as traditional data mining. Internet of things data mining is not the same as traditional data mining (Vadlamudi, 2019).
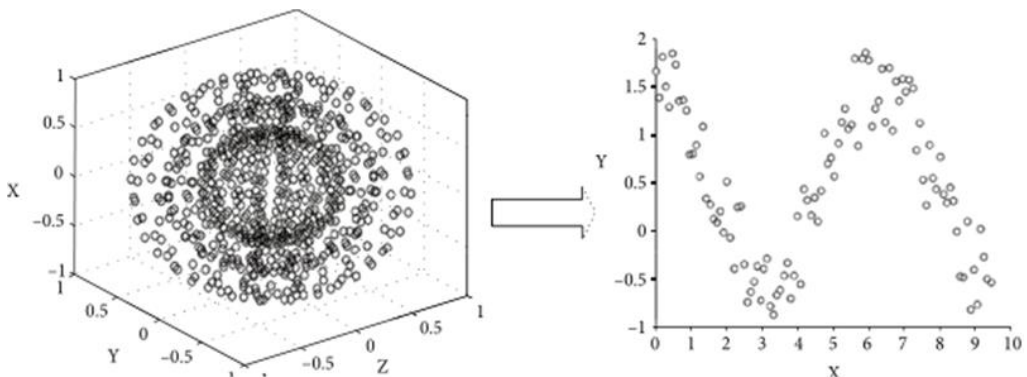

Figure 1: The transfer of high-dimensional data into lesser dimensions

Extraction issues arise in the internet of things as a result of traditional data. As a result, mining algorithms must be updated for dealing with difficulties of scalability and huge data (Cai et al., 2017). Classification many obstacles and issues in mining are represented by problems, as well as machine learning research (Paruchuri, 2018). The goal of the Internet of Things is to classify each object that falls into big data based on the characteristics displayed by its features (Tsai et al., 2015). Appropriate data representations are required to describe physical world data (Kim, 1998), and distinguishing the favorable qualities is

difficult. Data that is noisy, redundant, irrelevant, or distorted can degrade performance, lower classification accuracy, and raise the cost and calculation time of mining algorithms (Ramadan and Abdel-Kader, 2009). Data mining algorithms can benefit and function more efficiently if these factors are removed (Yeh et al., 2011). According to the research presented in Bhatt and Kankanhalli (2011), lowering data measurements (dimensionality reduction) can be used to solve this problem by only relevant and useful features were chosen.

If L is a data with dimension D, such that $L = \{f_{i=1}^n\}\epsilon\mathbb{R}^D$, then dimensionality reduction is implemented, just to ensure that no dataset M has less dimension d, D, and $M = \{f_{i=1}^n\}\epsilon\mathbb{R}^d$, where d<<D in such a manner that M denotes the dataset L with less feature dimension d.

The smaller data set allows mining algorithms to perform better and faster, resulting in better classification results and network performance. Due to the huge search space, dimensionality reduction is a difficult task, as the size of data grows exponentially in proportion to its properties. According to the features provided by its dimensions, an attribute may become useful or redundant in certain contexts. To conduct an exhaustive search, which is impossible when the search space is indefinite, optimal search strategies are required. To choose profound attributes to minimize data dimensions and represent the data in its most approved manner for categorization and other extraction tactics, a wide range of search algorithms (e.g., sequential forward/ backward selection) has been presented. Nonetheless, attribute reduction strategies are being undercut by data and algorithm issues, as well as issues for vibrant fitness function, which surpass value and computing complexity (Liu and Yu, 2005).

Dimensionality Reduction Techniques

Lowering the amount of dimensions in a large search area is a difficult and time-consuming task for today's networks and processing technology. Figure 2 illustrates a simplified knowledge discovery and dimensionality reduction procedure. The methodologies of selecting features and feature mining (Ristoski and Paulheim, 2016) which are the main for dimensionality reduction will be explored in this section.

- *Selecting Feature*

Data extraction requires pretreatment methods to reduce the difficulty, computation, storing, and cost of classification. In the Internet of Things, the raw dataset on data capturing layers is expressed by high-dimensional parameters, which makes extraction algorithms more challenging to process. The task of picking a feature subset from existing applicant features reflecting initially quantified datasets is referred to as feature selection (Ramirez-Gallego et al., 2017). Feature selection and mining are also widely employed in the fields of computer vision and image processing (Ganapathy & Neogy, 2017).

In the realm of dimensionality reduction in hyperspectral imaging, more technical work on the local manifold representation with affinity matrix usage (Hong et al., 2017). Machine learning techniques are also widely utilized for labeling graphs and learning features in the field of hyperspectral dimensionality reduction for remote sensing data (Hong et al., 2019). A set of n datasets can be used as input to a mining or training classifier. Each dataset L can be made up of N characteristics that describe the original set. As shown below, instance L is a tuple with D dimensions.

L = {F₁(L),F₂(L),F₃(L),….Fi(L),L ∈ R^D
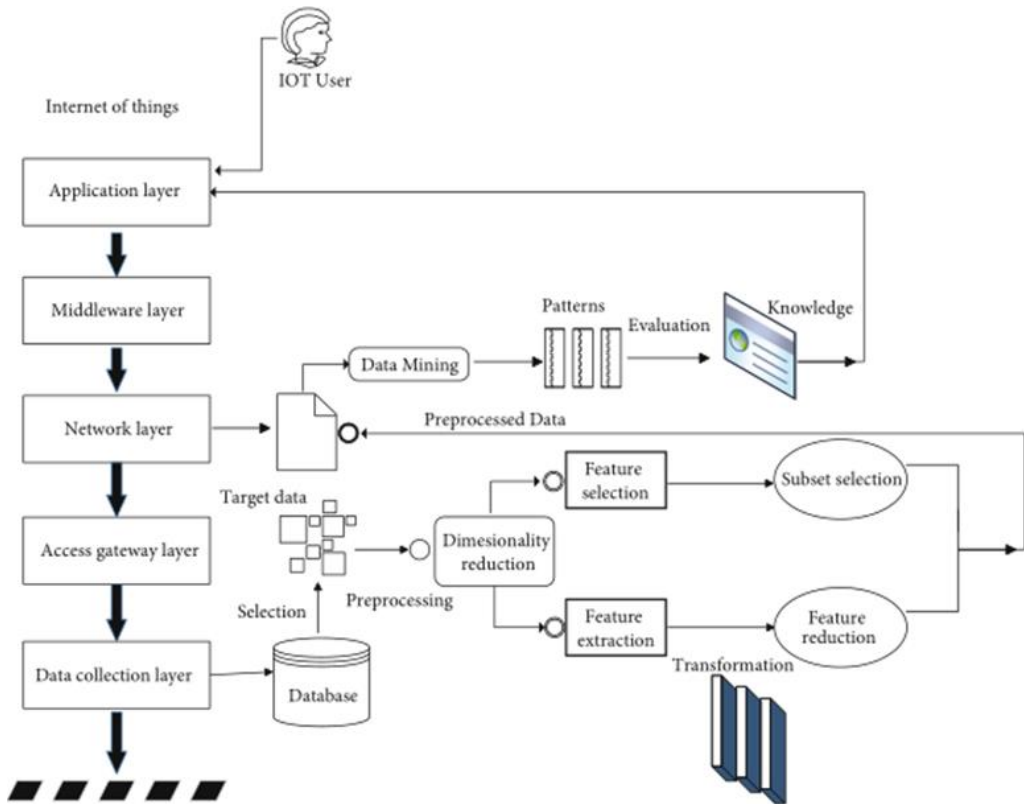
$|L| = N, \rightarrow (N = D)$

Figure 2: Simplified knowledge discovery and dimensionality reduction procedure

Where F_i is the ith feature's domain, and L's cardinality is N, with dimensions D. Let n denote the chosen dimensions d in M, $M \subseteq L$.

$$M = \{f\,i(M), f\,i+1(M)\} \rightarrow M \subseteq L \subseteq R^D,$$

$$|M| = n \rightarrow (n = d)$$

- ***Feature Mining***

Feature mining is an important data pre-processing approach that adds quality to extraction methods as an activity enhancer for the internet of things networks to modify obtainable high-dimensional databases, that is uploading high precision data depiction classic for the original feature interplanetary (Keramati et al., 2014).

$$X = \{x1, x2, x3, x4, \cdots\cdots \cdots, xn\},$$

$$Xi \in R^d \Rightarrow Yi \in R^p | (p \ll d)$$

The presence of laid-off and incorrect pieces of evidence in original datasets requires a preconditioning method (feature mining). The feature mining method dents 2 study problems – evaluation measures and searches approach. The search interplanetary comprises comprehensive and feature subset and feature mining modify existing feature to comb for ideal solution set (Ding et al., 2019).

$$\sum_{f}^{D} = O \binom{D}{f} = (1 + 1)^D = 2^D$$

Where D denotes dimension, and f represents the latest feature subset size.

Hence, dimensionality reduction is an evident-oriented issue that determines 2 fundamental reasons comprising feature space reduction and improving the exactness or precision of extraction approaches.

Cuckoo Search Optimization

Cuckoo search (CS) is a metaheuristic optimization algorithm that was first presented in 2009 (Yang and Dan, 2009) and is based on some effective qualities such as breeding the biological lifestyle of the cuckoo. The expansion of these 2 concepts is used in the algorithm: Stochastic search and randomization (random walk). The algorithm starts at the beginning to look for local optima in the local search space (Rn). Furthermore, the algorithm is not constrained by local optima; rather, as the problem grows more global, it expands to provide a globally optimal solution (Ganapathy, 2016). Even though the initial proposed method is tailored to have a small set of features and is focused on local optimization, the modified and extended CS algorithms in Huang et al. (2016), by modifying the step size and parametric values, can be modified as a dynamic and global optimization process to amend its performance boundary.

- *Cuckoo's Living Behavior*

Cuckoos are obligate brood parasite organisms that rely on other birds for the development and the growth of their progeny (Payne, 1977). The cuckoo search (CS) is triggered by the genetic actions of cuckoos, such as foraging (food seeking) (Ganapathy, 2019). Cuckoos lay their eggs in host bird nests, where the eggs hatch and the juvenile seeks host attention for food (Davies, 2010). It also mimics some of the external characteristics of host eggs. Investigation and exploitation are the two tactics used. To come up with a fresh solution, CS uses levy flights. Cuckoos may hurl the eggs of a host bird to increase the chances of their egg hatching (Gandomi et al., 2013).

- *Algorithm Constraints*

(a). Each cuckoo lays one egg at a time and places it in a nest that is picked at random

(b) The nest with the best eggs will become the next generation

(c) The number of shells is fixed. Alien eggs can be recognized by the host bird with a probability of P_a (0, 1). If this occurs, the host bird either discards the eggs or abandons the nest and constructs a new one.

- *Algorithm Formulation*

Cuckoo chooses a nest and deposits its egg there, which is held by a host bird. The random walk determines which nests are chosen. The cuckoo's foraging and flying behavior exemplify randomness. A fresh solution V^{t+1} is represented by each egg laid by a cuckoo.

$v^{t+1} = A_{CS} (v^t , P(t))$

Where A_{CS} represents a 'nonlinear cuckoo search algorithm' that maps available internet of thing d-dimensional vector v^t with variables $P(t)$ to a relatively new vector v^{t+1}.

- *Random walk*

The CS algorithm is a combination of local and global random walks that have reached an equilibrium. As a result, it can not only optimize outdoor internet of thing data but also converge to local optima when necessary. A random walk is a series of random or stochastic processes that follow each other.

Xi (t + 1) = Xi (t) + step size $\oplus$L(β),

where L specifies levy flights and $\oplus$ represents sequenced growth for a respective new step, which is then introduced to the earlier applicant solutions.

A solution is developed through levy flight in each fresh iteration, and steps for search are performed from the levy distribution bound to the heavy-tailed distribution. The heavy-tailed distribution is not exponentially bound, unlike the normal distribution, and most of the values generated to match the fitness value (objective function) criteria (Maulik and Zwart, 2006). Error! Displays a random stroll. The reference source could not be located. When the search space is exponentially unbounded and continuously expands in any dimension and size, levy flight is preferable when ten thousand steps are taken at a time, t to find a better position than the previous one. Because of its levy flights to handle network data for global optimization, Cuckoo search is highly recommended (Paruchuri & Asadullah, 2018).

When data is large in volume and numerous data analytic activities are initialized to improve network performance, quick and self-organizing algorithms are required in information and communication technology. One of the global optimization strategies developed to sort out present global optimization difficulties is metaheuristic algorithms. Cuckoo search is a newly created metaheuristics algorithm that can best fit future smart internet of thing networks and their continuous outdoor data to give valuable services with better machine learning techniques among numerous metaheuristics algorithms such as harmony search and bat algorithm (Klaine et al., 2017).

- *Parameter Tuning*

The 'Cuckoo search strategy' is optimized for local search optimization in its original form, but its algorithmic limitation can be tweaked to widen its applications to global search optimization challenges. Several factors in the cuckoo search algorithm must be modified to get the optimal performance, including the nest size, the elitism probability (probabilistic collection of appropriate candidates), and the recurrence (Paruchuri, 2015). The algorithm begins with random parameters, as do many nature-inspired algorithms. The parameters are tweaked with changing step sizes at each phase of the iteration. The choice of step size is critical for the algorithm's convergence or divergence. For quick convergence or performance requirements, the step size can be increased or lowered depending on the application.

It is worthy to note that CS is a metaheuristic algorithm that is very efficient and effective used in dimensionality reduction, with very little limitation – that is CS is compromised when is dealing with multiobjective and discrete data but it is very effective when dealing with continuous optimization difficulties. Also, CS is motivated based on the problem of dealing with enormous data volume generated from internet of thing oriented system and wireless sensor networks.

METHODS

Internet of thing Vectors and Dimensions

The in-depth search interplanetary $\mathbb{R}^d$ involving of huge dimensions D is abridged via feature extraction prior to any additional tasks pertinent to data analytics comprising data extraction methods. The resultant qualified vectors will upsurge estimate precision with the slightest difficulty and cost.

Problem Narrative

Sensor data is saved as a series of rows and columns in a database. Each column and each row indicate a different vector v_i or dataset. C and the proportions that go with it d_j. Table 1 show the internet of thing vectors included in the databases and Schema Dimensions.

If all different items on the internet of thing are represented in vector set V, where each vector v_i contains data for a single object with multiple dimensions (attributes) $d^j \epsilon D$, then the number of vectors in V equals the number of tuples in the database, resulting in $1 \leq V \leq n$: A column matrix can be used to represent a set of n vectors, as seen below:

$$V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \\ v_n \end{bmatrix}, \forall^{v_1} V \in \mathbb{R}^n, \quad ,$$

Where respective vector v_i can be denoted with its dimensions as a single row matrix.

$$v_i = [d_1 \, d_2 \, d_3 \cdots d_{m-1} d_m], \forall \in \mathbb{R}^D$$

If we substitute each vector with its dimensions, then we get the m × n matrix.

$$V_{m \times n} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \cdots & d_{1m} \\ d_{21} & d_{22} & d_{23} & \cdots & d_{2m} \\ d31 & d32 & d33 & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & d_{m3} & \cdots & d_{mn} \end{bmatrix}, \forall V,$$

With reverence to vector v_i and its dimensions, $[d_1 \, d_2 \, d_3 \cdots d_{m-1} d_m]$ converts the present features that be in the right place to the original dimension interplanetary $\mathbb{R}D$ into abridged and changed vector v_i' and interplanetary $\mathbb{R}^d$ such that

If $v_i = [d_1 \, d_2 \, d_3 \cdots d_{m-1} d_m], \forall^{v1} \in \mathbb{R}^D$:

Then $T(v_i) = T[d_1 \, d_2 \, d_3 \cdots d_{m-1} d_m]$,

$=>T(vi) = vi' = [d_1 ', d_2 ', d_3 ' \cdots\cdots\cdots, d_m ']$

Table 1: 2-dimensional data for IoTs: demonstration for datasets $v_i \in V$ and their equivalent dimensions $d^j \in D$.

$V = v_i$			$D = d^j$		
	$d^{j=1}$	$d^{j=2}$	$d^{j=3}$	...	$d^{j=m}$
$v_{i=1}$	d_{11}	d_{12}	d_{13}	...	d_{1m}
$v_{i=2}$	$\ddots$	d_{22}	d_{23}	...	d_{2m}
$v_{i=3}$	d_{31}	$\ddots$	d_{33}	...	d_{3m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$v_{i=n}$	d_{n1}	d_{n2}	d_{n3}	...	d_{nm}

Feature Extraction

V_i is a separate untransformed vector of dimension m, and V is an original vector set of dimension n, and V_i' can be represented as a converted vector of dimension m' with the property $v_i' \in \mathbb{R}^d \ll \mathbb{R}^D$. The following are the steps involved in feature extraction:

i.   Feature construction
ii. Subset feature selection
iii. Result efficiency

The algorithm for feature extraction can be created based on these goals to maintain a generic approach for meaningful and effective feature extraction. Below is an algorithm for extracting high-quality features from the original space.

Feature Construction

For various data analytics jobs, feature extraction is a technique for transforming a large number of features or dimensions into a smaller number of features. To construct new features from existing ones, several parameters and techniques are required for the transformation of original features. Depending on the type of qualities to be built, these strategies may differ. Algorithms for nominal and numeric qualities are provided in the following sections. It will reduce data dimensionality to improve search optimization for any machine learning activity, such as classification and data mining, by focusing on the most relevant data on the internet of thing networks for additional processing. A vector with a large number of well-defined properties can help achieve high prediction accuracy. The construction of nominal and numeric qualities necessitated a variety of operators and procedures to uncover hidden information that can aid data analytics for decision-making (Pittner and Kamarthi, 1999).

Algorithm for Numeric Attributes

The database of the internet of thing is made up of several dimensions, and each of these dimensions can be denoted as a numeric value or combining both strings and characters. In either mode, a set of operators are modified to construct attributes for that reason. The algorithm below signifies the steps to construct fresh attributes for numeric features according to arithmetic operators (+, /, -). The tree representation for feature construction is shown in Figure 6, along with several mathematical operations for the dimensions di. The choice of operators is based on the problem and desired outcome (Paruchuri, 2019). The algorithm begins with a set of initial input vectors $v_i^\sim$, each of which is a collection of letters

or strings. The first loop replicates each feature dj to a new vector from an untransformed vector v_i. Original attributes are destroyed after the construction of selected attributes to avoid the repetition of similar attributes. Two internal loops pick operators and properties from a new vector v_i a selection of arithmetic operators Each vector is selected in the last loop. Starting j = 1 and adds all created features one by one 2, 3, m for each vector vi, I = 1, 2, 3, n for each vector v_i.

Algorithm for Nominal Attributes

If the type of attributes is not numeric, concatenation of texts or characters is utilized to generate new features instead of arithmetic operations. The production of nominal attributes is given in the following algorithm. By forming multiple pairs among all fields in a dataset v_i^-, nominal attributes are concatenated to create new features. The first loop transfers values from the original vector to a new vector called v_i. The final two laps to create features for all vectors vi, I = 1, 2, 3, n, select each dimension from a new array vi and select each vector progressively. At the time, t, Figure 7 shows the Heaviside function pa as a representative dimension scaling factor. For a local optimum solution, the local random walk is isotropic and can be represented as follows:

$$V_i^{t+1} = v_i^t + \psi s \otimes \theta(P_a - \sigma) \otimes (v_i^j(t) - v_k^i(t))$$

where $v_i^j(t)$ and $v_k^j(t)$ denote two unique vectors IΛk with the j^{th}d dimension at time t, respectively. is a size transformation scaling factor used to control the search space on the internet of things.

Mutated Cuckoo Search=Based Feature Extraction (CSFE) Algorithm

The enhanced version of CS is used to extract old and new dimensions, as well as to reduce the overall size of current dimensions. The algorithm begins by creating new features based on the original input vectors and then selecting enhanced attributes for each vector as shown in Table 2 (Mladenic, 2005). In the following section, the technique for cuckoo search-based feature extraction is presented, with the goal function selected based on the cited problem of outdoor IoT data. In each iteration, the algorithm will generate a feature subset, and this process will be repeated until an optimal cost or performance is attained.

Table 2: Original input vectors and then selecting enhanced attributes for each vector

Parameters	Algorithmic representation
Dataset environment	V
Fitness function	$\mathscr{F}(x)$
Number of datasets	N
Feature/dimension in i^{th} vector	j
Vector in dataset environment	i
j^{th} dimension in i^{th} vector	v_i^j
Probability of discarding a vector	p_i
Levy flight	$\mathscr{L}(\lambda)$
Number of iterations	T
Time instance	T
Step size	S
Step size scaling factor	$\psi > 0$
Normal distribution	$u \, (0,1)$

RESULTS AND DISCUSSION

Dataset Generation

The data acquired by the entities that make up the internet of things are stored on the internet of things. These are computing devices (such as a scanner or a thermostat) that can talk with one another over the internet to share information and services. Because these smart objects are distinct, the EPC of each one may be used as a primary key in the database to keep track of them.
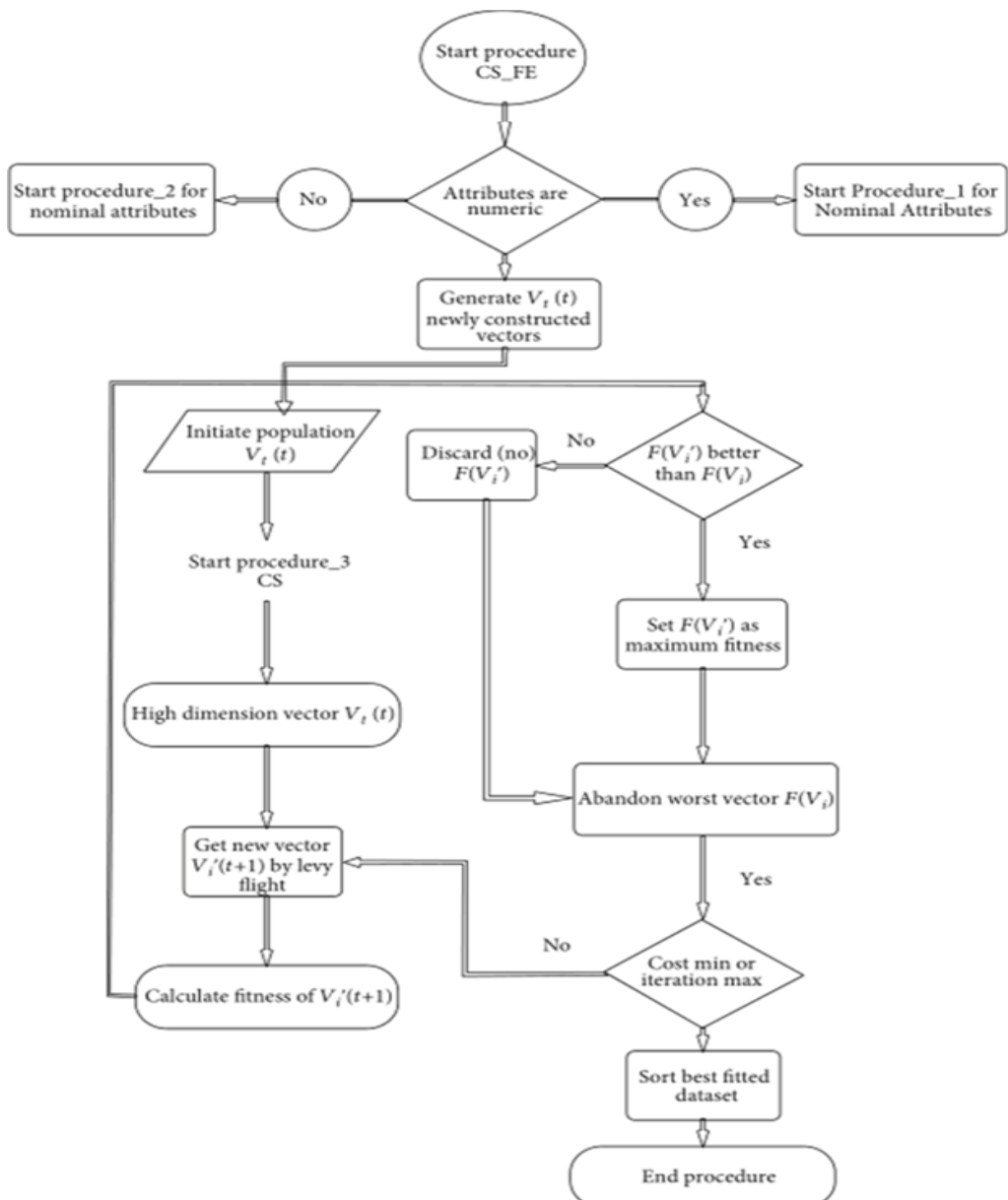

Figure 3: A flowchart for mutant cuckoo search-based feature extraction

For security and confidentiality reasons, firms that are advancing in IoT are unwilling to publish their private data openly. As a result, datasets relevant to IoT networks are scarce in the literature and on the internet. To use provided solutions for dimensionality (CSFE). As a result, IoT network datasets are hard to come by in research and on the internet. An acceptable and relevant dataset is required to apply proposed techniques (CSFE) for dimensionality reduction to outdoor IoT data. It will make it easier to get the right results by using the proposed technique in MATLAB to reduce the overall broad search space R^n and feature space R^D.

IoTify is a web-based simulation environment for developing IoT applications with virtual hardware components such as sensors. It enables the design and construction of virtual IoT devices in JavaScript and facilitates the formation of a virtual lab. JavaScript object name (JSON) with extension JSON is used to create the IoTify database. Table 3 shows the dataset that was utilized to generate the results and analyze the proposed methods (CSFE). When an urgent situation emerges, IoT-based devices can extract particular and required information from a patient's blood and send the resulting report to the doctor. Figure 3 shows a flowchart for mutant cuckoo search-based feature extraction, which includes methods for nominal and numerical attributes as well as the development of a reduced feature subset.

Table 3: Number of vectors and dimensions for each vector in the dataset: CBC results from internet of things patients.

Dataset name	IoT-oriented patients' CBC results
Number of vectors	498
Number of dimensions	59

Analyzing and Visualizing Algorithm Results

The suggested approach, mutant cuckoo search-based feature extraction, is implemented in MATLAB, and the results are displayed using graphs, and plots. The graphs for the newly formed space V' with n rows and m' decreased dimensions are shown in Figure 4.
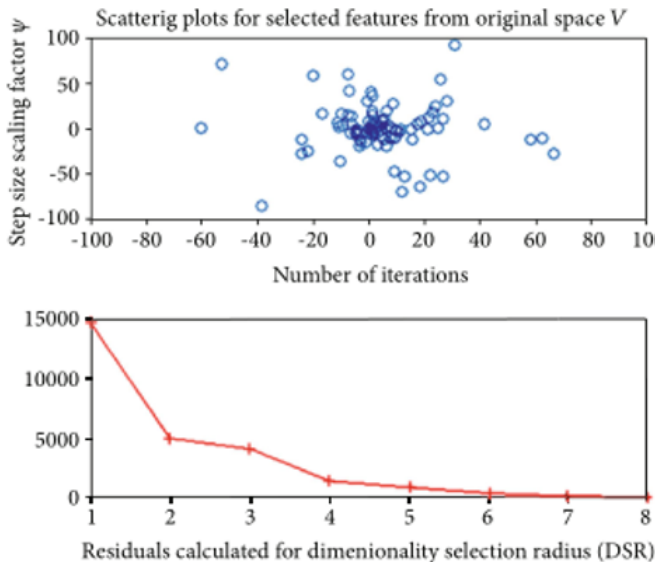

Figure 4: The graphs for the newly formed space V' with n rows and m' decreased dimensions

Plots are spread out and at a distance, indicating that dimensions were chosen from a large search space. The number of iterations for random generations is shown in the subplot, and residuals for each dimension are calculated to assess its weightage for selection. The levy flight (λ) and step size scaling factor ψ, are used to create the dimension selection, and the steps are picked at random from the levy distribution.

CSFE explores the wide search space R^n and R^d for candidates. Because data generated by IoT is continuous and requires a global optimization solution to boost network performance, search is continually rising as new things enter the IoT, necessitating an algorithm that can modify itself to respond to instant changes. The heavy-tailed distribution for CSFE derived for the input dataset of IoT-based Patients' CBC results is shown in Figure 5. Because the tail is exponentially expanding and constantly gives the best fitness, the highest peak is the global optimal solution for IoT data. The optimal choice is determined using DS and the coherence estimation factor. Because the area under the curve is not exponentially constrained, as data becomes extinct, the number of fitness values closer to the best optimum grows. The magnitude of the bars in Figure 6 reveals that the majority of the fitness values were globally best.
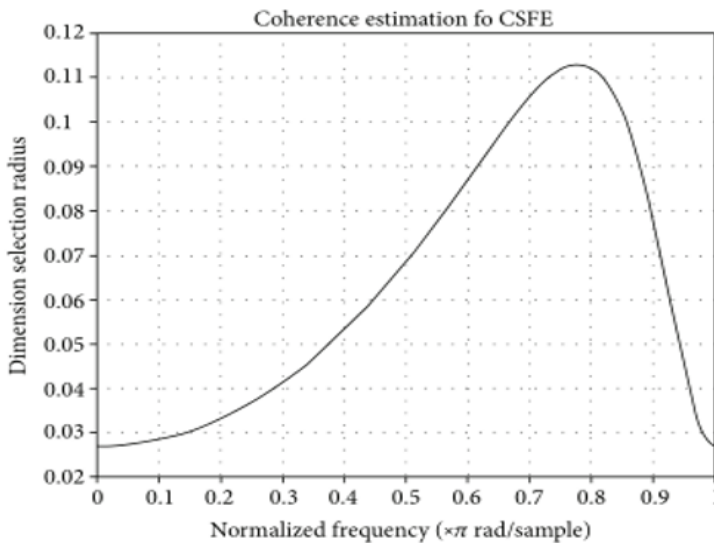

Figure 5: The heavy-tailed distribution for CSFE derived for the input dataset of IoT-based Patients' CBC results
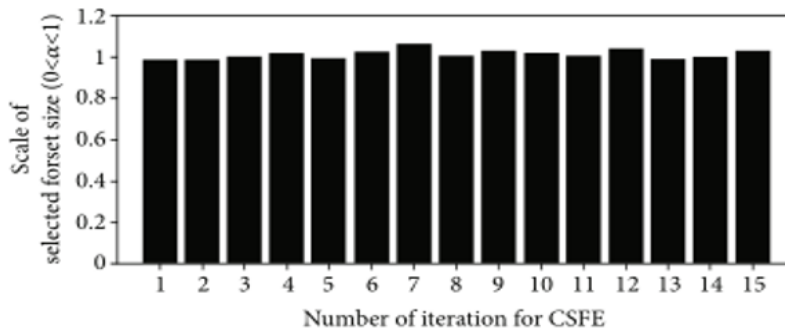

Figure 6: The magnitude of the bars

The term "Internet of Things" is used to describe technical advancement that is based on both the resulting network created by the integration of smart items and developed internet technologies, as well as a variety of supporting devices, equipment, and machinery. Applications and services are being created to take advantage of these technologies to support new business trends and provide daily conveniences. IoT is a networked smart device and integrated network architecture that supplements internet services by providing availability of all types of services to anybody, anytime and anywhere.

It is becoming a trend that most of the objects in our environment will be connected in some way. This marks a change away from traditional internet tactics in favor of using the internet to connect physical items that interact with one another and humans. These technologies generate massive volumes of data, which becomes crucial when analytics and machine learning approaches are applied. These types of technologies generate massive volumes of data, which becomes even more important when analytics and machine learning techniques are used to make them intelligent and self-organizing.

CONCLUSION

Although IoT networks have emerged and are capable of completing demanding tasks, they must be improved to meet future problems. It is conceivable when networks are not just fast and have top-tier servers, but they are also smart enough to deal with unforeseen scenarios. This goal can be met by introducing efficient and global optimization techniques. A metaheuristic global optimization technique is developed in this study to minimize the dimensions of outside data for IoT. The cuckoo search-based feature extraction algorithm is a mutant algorithm that organizes itself depending on the unpredictable amount of input and generates a new and improved feature space.

REFERENCES

Bhatt, C. A. and Kankanhalli, M. S. (2011). Multimedia data mining: state of the art and challenges. *Multimedia Tools and Applications, 51*(1), 35–76.

Cai, H., Xu, B., Jiang, L. and Vasilakos, A. V. (2017). IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges. *IEEE Internet of Things Journal*, *4*(1), 75–87.

Chiroma, H., Herawan, T. and Fister Jr, I. (2017). Bio-inspired computation: Recent development on the modifications of the cuckoo search algorithm. *Applied Soft Computing*, 61, 149–173.

Davies, N. (2010). Cuckoos, Cowbirds and Other Cheats, A&C Black.

Ding, W., Jing, X., Yan, Z. and Yang, L. T. (2019). A survey on data fusion in internet of things: towards secure and privacy preserving fusion. *Information Fusion*, 51, 129–144.

Ganapathy, A. (2016). Virtual Reality and Augmented Reality Driven Real Estate World to Buy Properties. *Asian Journal of Humanity, Art and Literature*, *3*(2), 137-146. https://doi.org/10.18034/ajhal.v3i2.567

Ganapathy, A. (2018). Cascading Cache Layer in Content Management System. *Asian Business Review*, *8*(3), 177-182. https://doi.org/10.18034/abr.v8i3.542

Ganapathy, A. (2019). Image Association to URLs across CMS Websites with Unique Watermark Signatures to Identify Who Owns the Camera. *American Journal of Trade and Policy*, *6*(3), 101-106. https://doi.org/10.18034/ajtp.v6i3.543

Ganapathy, A. (2019). Mobile Remote Content Feed Editing in Content Management System. *Engineering International*, *7*(2), 85-94. https://doi.org/10.18034/ei.v7i2.545

Ganapathy, A., & Neogy, T. K. (2017). Artificial Intelligence Price Emulator: A Study on Cryptocurrency. *Global Disclosure of Economics and Business*, *6*(2), 115-122. https://doi.org/10.18034/gdeb.v6i2.558

Gandomi, A. H., Yang, X. S. and Alavi, A. H. (2013). Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. *Engineering with Computers, 29*(1), 17–35.

Hong, D., Yokoya, N. and Zhu, X. X. (2017). Learning a robust local manifold representation for hyperspectral dimensionality reduction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(6), 2960–2975.

Hong, D., Yokoya, N., Chanussot, J., Xu, J. and Zhu, X. X. (2019). Learning to propagate labels on graphs: an iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction. *ISPRS Journal of Photogrammetry and Remote Sensing, 158*, 35–49.

Huang, L., Ding, S., Yu, S., Wang, J. and Lu, K. (2016). Chaos-enhanced cuckoo search optimization algorithms for global optimization. *Applied Mathematical Modelling, 40*(5–6), 3860–3875.

Karkouch, A., Mousannif, H., Al Moatassime, H. and Noel, T. (2016). Data quality in internet of things: a state-of-the-art survey. *Journal of Network and Computer Applications*, *73*, 57–81.

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M, and Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. *Applied Soft Computing, 24*, 994–1012.

Kim, J. (1998). Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation, MIT press.

Klaine, P. V., Imran, M. A., Onireti, O. and Souza, R. D. (2017). A survey of machine learning techniques applied to self-organizing cellular networks. *IEEE Communication Surveys and Tutorials, 19*(4), 2392–2431.

Liu, H. and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering, 17*(4), 491–502.

Maulik, K. and Zwart, B. (2006). Tail asymptotics for exponential functionals of Lévy processes. *Stochastic Processes and their Applications, 116*(2), 156–177.

Mladenić, D. (2005). Feature selection for dimensionality reduction, in International Statistical and Optimization Perspectives Workshop," in Subspace, Latent Structure and Feature Selection, pp. 84–102, Berlin, Heidelberg, Springer.

Paruchuri, H. (2015). Application of Artificial Neural Network to ANPR: An Overview. *ABC Journal of Advanced Research*, *4*(2), 143-152. https://doi.org/10.18034/abcjar.v4i2.549

Paruchuri, H. (2018). AI Health Check Monitoring and Managing Content Up and Data in CMS World. *Malaysian Journal of Medical and Biological Research*, *5*(2), 141-146. https://doi.org/10.18034/mjmbr.v5i2.554

Paruchuri, H. (2019). Market Segmentation, Targeting, and Positioning Using Machine Learning. *Asian Journal of Applied Science and Engineering*, *8*(1), 7-14.

Paruchuri, H., & Asadullah, A. (2018). The Effect of Emotional Intelligence on the Diversity Climate and Innovation Capabilities. *Asia Pacific Journal of Energy and Environment*, *5*(2), 91-96. https://doi.org/10.18034/apjee.v5i2.561

Patel, K. K. and Patel, S. M. (2016). Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. *International Journal of Engineering in Computer Science, 6*(5), 1-20.

Payne, R. B. (1977). The Ecology of Brood Parasitism in Birds. *Annual Review of Ecology and Systematics, 8*(1), 1–28.

Pittner, S. and Kamarthi, S. V. (1999). Feature extraction from wavelet coefficients for pattern recognition tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 21*(1), 83–88.

Ramadan, R. M. and Abdel-Kader, R. F. (2009). Face recognition using particle swarm optimization-based selected features. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 2*(2), 51–65.

Ramirez-Gallego, S., Krawczyk, B., García, S., Woźniak, M. and Herrera, F. (2017). A survey on data preprocessing for data stream mining: current status and future directions. *Neurocomputing, 239*, 39–57.

Ristoski, P. and Paulheim, H. (2016). Semantic Web in data mining and knowledge discovery: A comprehensive survey. *Journal of Web Semantics, 36*, 1–22.

Tsai, C. W., Lai, C. F., Chao, H. C. and Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big Data, 2*(1), 1-15.

Vadlamudi, S. (2016). What Impact does Internet of Things have on Project Management in Project based Firms?. *Asian Business Review*, *6*(3), 179-186. https://doi.org/10.18034/abr.v6i3.520

Vadlamudi, S. (2019). How Artificial Intelligence Improves Agricultural Productivity and Sustainability: A Global Thematic Analysis. *Asia Pacific Journal of Energy and Environment*, *6*(2), 91-100. https://doi.org/10.18034/apjee.v6i2.542

Yang, X. Y. and Dan, W. U. (2009). Atomic simulations for surface-initiated melting of Nb(111). *Transactions of the Nonferrous Metals Society of China, 19*(1), 210– 214.

Yeh, J. Y., Wu, T. H. and Tsao, C. W. (2011). Using data mining techniques to predict hospitalization of hemodialysis patients. *Decision Support Systems*, *50*(2), 439–448.

Yoo, Y. (2010). Computing in Everyday Life: A Call for Research on Experiential Computing. *MIS Quarterly*, *34*(2), 213–231.

--0--