Wireless Communications and Mobile Computing

Smart Antennas and Intelligent Sensors Based Systems: Enabling Technologies and Applications 2020

Lead Guest Editor: Fawad Zaman
Guest Editors: Hing Cheung So, Daehan Kwak, Farman Ullah, and Sungchang Lee

Smart Antennas and Intelligent Sensors Based Systems: Enabling Technologies and Applications 2020

Wireless Communications and Mobile Computing

Smart Antennas and Intelligent Sensors Based Systems: Enabling Technologies and Applications 2020

Lead Guest Editor: Fawad Zaman
Guest Editors: Hing Cheung So, Daehan Kwak, Farman Ullah, and Sungchang Lee

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9820571, 3 pages
https://doi.org/10.1155/2022/9820571

Editorial

Smart Antennas and Intelligent Sensors Based Systems: Enabling Technologies and Applications, 2020

Fawad Zaman,[1] **Hing Cheung So**,[2] **Daehan Kwak**,[3] **Farman Ullah**,[4] **and Sungchang Lee**[5]

[1]*COMSATS University Islamabad, Islamabad, Pakistan*
[2]*City University of Hong Kong, Hong Kong*
[3]*Kean University, New Jersey, USA*
[4]*COMSATS University Islamabad, Attock, Pakistan*
[5]*Korea Aerospace University, Gyeonggi-do, Republic of Korea*

Correspondence should be addressed to Fawad Zaman; fawad.zaman@comsats.edu.pk

Received 22 April 2022; Accepted 22 April 2022; Published 19 May 2022

The present Industry 4.0 revolution, as well as advances in IoT and Sensors Technologies, has made human life easier and more convenient. Controlling, monitoring, and analyzing real-time systems for a range of applications, including the COVID-19 pandemic, health monitoring, and smart homes, have become possible because of smart antennas, intelligent sensor communication, and computing-based networks. Smart antennas and intelligent sensor-based systems offer a variety of applications in multiuser communication, particularly in unpredictable and unforeseen conditions [1]. Smart antennas have been widely employed in adaptive beamforming, in which the main beam must be focused in a certain direction and nulls must be handled in unwanted directions [2]. The integration of smart antennas and intelligent-based sensor networks allows for the development of various algorithms that may be utilized for environmental decision-making, sensing, and monitoring [3–5]. For contactless epidemic illness monitoring, breathing difficulties detection, and COVID-19 forbidden activities recognition, smart antenna and intelligent sensors are a key area of research.

The main focus of this special issue is to give a broad perspective on current research in the fields of smart antennas and intelligent-based sensor networks in order to enable new technologies and applications.

We strongly encouraged specific academics to submit studies on intelligent sensors and smart antenna-based devices. As a result, this special issue highlights the most up-to-date research in this field.

The paper "A Novel Deceptive Jamming Approach for Hiding Actual Target and Generating False Targets" describes a new method for hiding the actual target while simultaneously producing several false targets against FDA radar. The modified FDA radar is expected to be mounted aboard the actual aircraft for this purpose. To take advantage of FDA radar's range-dependent pattern nulling capabilities, it intercepts the opponent's radar signals and broadcasts back to place nulls in the radiation pattern at the required range and direction. To deceive the opponent's radar system, the proposed deceptive jammer creates delayed versions of received signals to create bogus targets with varied ranges.

The paper "On the Performance of Self-Concatenated Coding for Wireless Mobile Video Transmission Using DSTS-SP-Assisted Smart Antenna System" describes a novel approach to the concept of self-concatenated convolutional coding (SECCC) with sphere packing (SP) modulation using DSTS-based smart antennas. For the Rayleigh fading channel, the suggested DSTS-SP SECCC scheme is tested. With the help of an interleaver, the SECCC structure is created using the recursive systematic convolutional (RSC) code. Extrinsic Information Transfer (EXIT) curves are used to study the proposed system's convergence behavior. The suggested system's performance is measured using the H.264

standard video codec. The DSTS-SP SECCC's perceived video quality is determined to be much superior than the DSTS-SP RSC's.

The study "On the Performance of Wireless Video Communication Using Iterative Joint Source Channel Decoding and Transmitter Diversity Gain Technique" described a wireless video communication system based on iterative joint source channel decoding (IJSCD). The sphere packing (SP) modulation assisted differential space-time spreading (DSTS) multiple input-multiple output (MIMO) technique is used in the projected transmission system. By maintaining the maximum possible Euclidean distance between the modulated symbols, the SP modulation-aided DSTS transmission mechanism achieves substantial diversity gain. Furthermore, because no channel estimation mechanism is used, the suggested DSTS system results in a low-complexity MIMO scheme. Various combinations of IJSCD error protection schemes helped by source bit coding (SBC) have been utilized, all with the same total bit rate budget.

The paper "Dimensionality Reduction for the Internet of Things Using the Cuckoo Search Algorithm: Reduced Implications of Mesh Sensor Technologies" highlights a problem in the Internet of Things network and presents a unique cuckoo search-based outdoor data management system. The feature extraction approach is used to extract useful information from unstructured and high-dimensional data. 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.

The study "A Comparative Analysis of Different Outlier Detection Techniques in Cognitive Radio Networks with Malicious Users" presents a new type of malicious user, the lazy malicious user (LMU), which has two phases of operation: awake and asleep. Statistical analysis is used to detect anomalous user behavior and mitigate its negative consequences. In the presence of the LMU and opposing types of malevolent users, results for various hard combination techniques are obtained. The results of simulations for error probability, detection probability, and false alarm at various levels of SNRs and varying contributions of the LMUs and OMUs indicated that the median test outperforms the other outlier detection techniques in MU detection.

The study "Nonorthogonal Multiple Access for Next-Generation Mobile Networks: A Technical Aspect for Research Direction" reviews and compares the basic principle of NOMA with other orthogonal multiple access technologies (OMA). In the most recent NOMA plan, a complete survey is offered. The design principles of NOMA schemes are covered, as well as recent deployments. Furthermore, the bit error rate, system capacity, and energy efficiency of the systems are compared. NOMA can meet the required goals in terms of user data rate, system capacity, interference cancellation technique, and reception complexity, according to the performance findings.

The study "Bodacious-Instance Coverage Mechanism for Wireless Sensor Network" suggested a Bodacious-instance Coverage Mechanism (BiCM) based on instance (node) redeployment. In the coverage region, the suggested technique creates new instance positions. It has two stages: in the first, it uses the Dissimilitude Enhancement Scheme (DES) to locate the intended instance position and move the instance to a new location, and in the second, it uses the depuration to reduce the moving distance between the initial and intended instance positions. Furthermore, the optimal parameters for several BiCM characteristics such as loudness, pulse emission rate, maximum frequency, grid points, and sensing radius have been discovered.

A new notion of CLC to IBC heterogeneous generalized signcryption is presented in the publication "A Lightweight Nature Heterogeneous Generalized Signcryption (HGSC) Scheme for Named Data Networking-Enabled Internet of Things." The proposed method delivers security features based on situational requirements while being compliant with NDN's structural policy. Given the resource constraints of IoT, a lightweight elliptic curve cryptosystem called the hyperelliptic curve cryptosystem is utilized, which provides the same level of security as bilinear pairing and an elliptic curve cryptosystem with a small key size.

Game theory was applied to develop the performance of intrusion detection systems in the study "Intrusion Detection into Cloud-Fog-Based IoT Networks Using Game Theory." The infiltration mode of the attacker and the behavior of the intrusion detection system are examined as a two-player nonparticipatory dynamic game, with Nash equilibrium solutions employed to generate specific subgames. Various parameters were investigated during the simulations utilizing game theory and Nash equilibrium definitions to extract the parameters with the most accurate detection findings. The results of the suggested method's simulation demonstrated that using intrusion detection systems based on cloud-fog in the Internet of Things can be extremely effective in recognizing attacks with the least number of errors in this network.

An intelligent moth flame optimization-based clustering (IMOC) for a drone-assisted vehicular network is presented in the paper "IMOC: Optimization Technique for Drone-Assisted VANET (DAV) Based on Moth Flame Optimization." This method is utilized to give maximum coverage for the vehicular node while using the fewest cluster heads (CHs) possible. The key topic addressed in this article is delivering optimal route by offering end-to-end connectivity with minimal overhead. The performance indicators used for comparison study are node density, grid size, and transmission ranges. These parameters were adjusted for each algorithm during simulations, and the results were recorded. Ant colony optimization, comprehensive learning particle swarm optimization, and gray wolf optimization were used to compare state-of-the-art clustering techniques for routing.

The study "IoT-Based Healthcare Support System for Alzheimer's Patients" employed a variety of communication protocols between sensors and smartwatch, including Message Queue Telemetry Transport (MQTT) and WebSocket (with authentication and auto closing of connection). Doctors, patients, and ambulances may all be tracked using the

secure backend admin panel. These protocols are established with security in mind to preserve patients' privacy.

In the work "A New Computing Paradigm for Off-Grid Direction of Arrival Estimation Using Compressive Sensing," a method for solving grid mismatch or off-grid target for direction of arrival (DOA) estimation using compressive sensing (CS) methodology is offered. The sources are located at a few angles in comparison to the full angle domain, i.e., they are spatially sparse sources, and their position can be estimated using CS approaches that can achieve super resolution and estimation with a lesser amount of samples. The source energy is spread among the adjacent grids due to grid mismatch in CS approaches, and a fitness function based on the difference of the source energy among the adjacent grids is introduced.

The paper "Management of Load-Balancing Data Stream in Interposer-Based Network-on-Chip Using Specific Virtual Channels" describes a strategy for avoiding the aforementioned interference by designing two different virtual channels and several links that control which memory block is accessed. When employing the interposer layer, our method uses the destination address to determine which channel and link should be used. When compared to typical load-balancing and unbalanced systems, simulation results demonstrate that the suggested mechanism reduces latency by 32% and 14%, respectively.

The goal of the study "Presenting an Effective Method to Detect and Track the Broken Path in VANET Utilizing UAVs" is to simulate a VANET in an urban area using cloud computing infrastructure and unmanned aerial vehicles (UAV) to prevent the negative impact of packet delivery and routing barriers. To assess the proposed method, it is compared to the ClouDiV basic protocol. The proposed technique outperforms other methods with varying densities and variable times in terms of efficiency and performance, according to Ns-2 simulation findings.

The study "Support Vector Machine-Based Classification of Malicious Users in Cognitive Radio Networks" proposes a support vector machine (SVM)-based machine learning approach to categorize valid SUs and MUs in the CRN. Both classification and regression are accomplished using the proposed SVM-based approach. By drawing a hyperplane on the base of greatest margin, it clearly classifies authentic SUs and MUs. The sensing data from the valid SUs are integrated at the FC using Dempster-Shafer (DS) evidence theory after successful classification. Simulations are used to demonstrate the effectiveness of the proposed SVM-based classification technique when compared to existing systems.

Conflicts of Interest

There is no conflict of interest.

Fawad Zaman
Hing Cheung So
Daehan Kwak
Farman Ullah
Sungchang Lee

References

[1] F. Zaman, S. Lee, M. K. A. Rahim, and S. Khan, "Smart antennas and intelligent sensors based systems: enabling technologies and applications," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 6475832, 3 pages, 2019.

[2] A. Manikas, Ed., *Beamforming: Sensor Signal Processing for Defence Applications*, vol. 5, World Scientific, 2015.

[3] F. Zaman, "Joint angle-amplitude estimation for multiple signals with L-structured arrays using bioinspired computing," *Wireless Communications and Mobile Computing*, vol. 2017, Article ID 9428196, 12 pages, 2017.

[4] Y. Xiao, J. Xie, L. Huang, and H. C. So, "Multiantenna assisted source detection in toeplitz noise covariance," *IEEE Signal Processing Letters*, vol. 26, no. 6, pp. 813–817, 2019.

[5] K. Luo and A. Manikas, "Joint transmitter–receiver optimization in multitarget MIMO radar," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6292–6302, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8844630, 20 pages
https://doi.org/10.1155/2021/8844630

Research Article

A Novel Deceptive Jamming Approach for Hiding Actual Target and Generating False Targets

Shahid Mehmood,[1] Aqdas Naveed Malik,[1] Ijaz Manssor Qureshi,[2] Muhammad Zafar Ullah Khan,[1] and Fawad Zaman [3]

[1]*Department of Electrical Engineering, International Islamic University, Islamabad, Pakistan*
[2]*Department of Electrical Engineering, Air University Islamabad, Pakistan*
[3]*Department of Electrical & Computer Engineering, COMSATS University Islamabad, Pakistan*

Correspondence should be addressed to Fawad Zaman; fawad.zaman@comsats.edu.pk

Received 6 August 2020; Revised 2 March 2021; Accepted 24 March 2021; Published 7 April 2021

Academic Editor: Cong Pu

Deceptive jamming is a popular electronic countermeasure (ECM) technique that generates false targets to confuse opponent surveillance radars. This work presents a novel approach for hiding the actual target while producing multiple false targets at the same time against frequency diverse array (FDA) radar. For this purpose, the modified FDA radar is assumed to be mounted on the actual aircraft. It intercepts the opponent's radar signals and transmits back to place nulls in the radiation pattern at the desired range and direction to exploit FDA radar's range-dependent pattern nulling capability. The proposed deceptive jammer produces delayed versions of the intercepted signals to create false targets with multiple ranges to confuse the opponent's radar system. The novel mathematical model is proposed whose effectiveness is verified through several simulation results for different numbers of ranges, directions, and antenna elements.

1. Introduction

Modern warfare is information and electronic based, which is evidently replacing the conventional platforms [1]. It requires no further confrontation between soldier to soldier, trench to trench, platoon to platoon, and platform to platform; rather, it relies on nonlinear and nonsymmetric war between system to system [2]. As the 21st century unfolds, the concept of electronic information warfare has become center of the gravity in which the radar system constitutes the key components that provide early warning capabilities [3]. Radar works as an early warning system and bestows extra time space to react against imminent threat. The principle of the radar to detect the desired target is transmitting radio waves towards the target and calculates round-trip time of the reflected waves after striking with the target [3]. To counter enemy radars, ECM techniques (also known as radar jamming techniques) were introduced [4]. These techniques are used to deny the important information about the desired aircrafts (direction of arrival, range, velocity, etc.) that any foe radar seeks [4]. In the presence of strong electronic counter measure (radar jammers) systems, it is difficult to detect the target aircraft; but there are many electronic counter-countermeasure (ECCM) techniques available in literature to counter radar jammer and to locate correctly the target [5, 6]. Some ECCM (which are also known as antijamming) systems steer nulls towards strong interfering signals to secure own functioning.

Mainly, there are two types of radar jamming techniques: mechanical methods (passive jamming) and electrical methods (active jamming) [3]. In mechanical methods, physical means like chaffs, decoys, corner reflectors, and stealth are used in securing the desired aircrafts' flights and to deceive enemy radars [4]. These physical means of radar jamming are traditional techniques which are not so effective. Electrical methods of radar jamming are effective and still in use [1, 2]. These methods can be categorized in two types: barrage jamming (also known as noise jamming) and deceptive jamming

[4]. Barrage jamming methods use radar jammers to put huge powers across the desired spectrum of the frequencies which blankets the radar's display to interrupt its normal functioning [7]. There are two main reasons of the ineffectiveness of this type which is huge power losses for longer periods of times, and even it cannot cover the entire frequency spectrum at the same time [8]. The deceptive jamming mode is used to deceive the enemy radars by showing them multiple very similar fake targets with different aircraft attitudes (range, direction, velocity, acceleration, etc.) [9].

This paper focuses only on the deception jamming, because this is the most effective way to secure the flight of the desired aircraft from the enemy radars by showing congruent false targets [10]. Therefore, implementing effective and efficient deceptive jamming (DJ) techniques has become a hotspot area of research in radar electronic countermeasures [11, 12]. Many methods have been developed in the modern literature [13–15]. In pursuit of deceptive radar jamming, the simplest way to generate multiple false targets is to hold enemy radar signals and after doing time-modulation transmit those signals back to the enemy radar [16]. When the enemy radar receives these signals, it will perceive multiple false targets with different ranges, but with the same direction along with the actual target. A modest contribution has been made using target pose and motion information to generate false targets in [17]. To deceive the opponent radar with a number of fake targets, another effort has been made in [18] using micromotion characteristics, but both have the same issue of computation complexity.

Multiple false targets also have been achieved using electromagnetic properties (EM model modulation) and translation modulation in [19]. By exploiting the concept of sub-Nyquist sampling theorem, a series of multiple fake targets have proposed in [20, 21]. Another remarkable approach has been defended to produce multiple false targets using product modulation in which an offline deceptive signal template is produced and then multiplied with enemy radar signal in [22]. Interrupted-sampling repeater jamming (ISRJ) establishes a novel approach to generate deception jamming by allowing the single radar antenna jammer to sample periodically and iterating a fraction of the intercepted enemy radar signal [23]. Inappropriately, high complexity and large computation is the main drawback of inefficiency of the above deceptive techniques in the field of electronic countermeasures. Further, these techniques are also unable to hide real target.

A recent effort has been exercised to achieve the goal by adding escort-free flight jammer drone ahead of the actual aircraft based on periodic the 0-π phase modulation which neutralizes the effectiveness of the enemy radar by displaying its multiple verisimilar false targets [24]. The escort-free flight jammer intercepts enemy radar signals and after doing phase modulation in the periodic 0-π sequence, these signals are retransmitted towards the actual target, whereby these are scattered towards the enemy radar and present multiple false targets with different ranges [24], but it considers the scenario where a separate escorting drone jammer is required which is not feasible in most practical situations.

Against to the only angle-dependent beam scanning techniques, FDA is an efficient beam scanning array used for phased array radars which has recently got tremendous attention in literature due to its greater achievement of wide angle coverage [25–28]. The radiation pattern of the phase array radar (PAR) depends only upon the direction while FDA radiation pattern depends upon the direction and range, due to its diversity in frequency across the array elements. Hence, FDA radiation pattern is capable of null steering to the particular range and direction. FDA is implemented by applying small increment in frequencies across array elements to achieve range angle-dependent beam scanning transmission [29, 30]. Hence, it enables beam scanning without need of any phase shifters and physical steering [31–33].

A good effort in field of deceptive jamming is explored in [34] which utilizes frequency diverse array. It produces multiple fake targets at different distances across the slant range but at the same azimuthal range aligned with the actual target. The technique is unable to draw false targets at different azimuthal ranges other than azimuthal range of the actual target. A novel approach in the field of deceptive jamming has been introduced in [35] with wave scattering using FDA for space-borne synthetic aperture radar (SAR). This approach considers the scenario where we offer deceptive jamming to the opponent radar for securing our valuable targets in its own territory with the help of placing deceptive reflectors. But this technique provides no solution to tackle the opponent radar operating from its own territory without the help of any ground-based situated wave scattering reflectors. Further, this method is also unable to hide its own target from the enemy foresight.

Although a modest effort has been made in [36] to introduce the deceptive jamming approach through frequency diversity, there are a few shortcomings in the proposed technique. These include (a) that deceptive jammer should be synchronized/attached or working in collaboration with a friendly GPS satellite system, (b) deceptive jammer must have prior knowledge of the location in space of the opponent radar, (c) the method is slow because it is using FFT and IFFT to convert signals from time domain to frequency domain and then back to previous domain, which makes it performance slow, and (d) it does not give any solution if the foe radar is also an FDA radar.

Paper in the reference [37] presented a deception jamming method which generates multiple scenes (multiple false targets) using FDA radar antenna, where number of false targets depend on the number of antenna array elements. The technique in reference [38] uses the simplest way to generate nulls towards the desired direction and range in order to suppress the offered range-angle dependent interference jamming, but this technique is unable to offer deceptive jamming to confuse the opponent radar, and it also does not hide its own target from the vision of the opponent radar. Further, as we know that FDA is time, angle, and range-dependent, but in this technique, the time-dependency factor of the FDA radar is diminished by considering it as zero ($t = 0$), which is not an appropriate for the practical scenarios [38].

Until now, best to the authors' knowledge, available deceptive jamming techniques in the present literature are not dealing with the hiding of the actual target along the generation of false targets using FDA radars. Present literature

also does not offer deceptive jamming for the opponent radars which work on FDA radar principle. The proposed study would investigate these limitations in depth and subsequently would present a probable solution against it. The main contribution of the work is summarized below.

(i) This research produces a novel deceptive jamming approach in the field of ECM

(ii) The algorithm works against the opponent FDA radar and hides the actual target from it

(iii) For this purpose, enemy radar pulse is captured by the target FDA radar, and null is placed at the radar range to hide its own target alongside after its time modulation

(iv) It is equally effective to tackle ground-based opponent FDA radar in its own territory

(v) The proposed technique efficiently works without help of ground-based wave-scattering reflectors or advance escort-free-drone jammers

(vi) The proposed algorithm also confuses the opponent FDA radar by multiple false targets at different user-defined ranges

The remaining part of the paper is organized in the following way. Sections 2 and 3 introduce mathematical background of the FDA radar and comparison with existing techniques, respectively. Section 4 depicts the proposed method to secure the flight of the actual aircraft in the enemy territory by neutralizing the dangers of enemy radars, while Section 5 shows the effectiveness and correctness of the proposed techniques via simulations in three dimensions and in two dimensions for four different cases. Finally, conclusion of the paper has been presented in Section 6.

2. Data Model for the FDA Radar

FDA radar uses small increment in frequency of each element over the antenna array. Radiation pattern of the FDA radar is a function of range, angle, and time [25–28]. FDA radar implements waveform diversity among the radiating elements which brings more functionality [29, 30]. Figure 1 depicts an FDA that consists of uniform linear array (ULA) having n-isotropic radiating elements. The distance d between any two adjacent elements is taken same with the uniform current distribution. The carrier frequency of each element is incremented by a small constant frequency offset [31]. The simplest monochromatic signal is assumed to be transmitted from the nth element of the array, and it can be mathematically expressed as [32]

$$s_n(t) = \exp\ (j2\pi f_n t), \tag{1}$$

where f_n is the frequency of the nth element as $f_n = f_0 + (n-1)\Delta f$ for $n = 1, \cdots, N$. Similarly, f_0, Δf, and N represent carrier frequency, a small constant frequency increment, and total number of elements, respectively, in the FDA array. When signal of the nth element reaches at a far-field location after time t_0 with range R_1 (reference to the first element in the array) and azimuth direction θ, its radiating beam can be represented as [31]

$$s_n(t - t_0) = \exp\ \{j2\pi f_n(t - t_0)\}. \tag{2}$$

For $t_0 = R_n/c$, (2) can be expressed as

$$s_n\left(t - \frac{R_n}{c}\right) = \exp\ \left\{j2\pi f_n\left(t - \frac{R_n}{c}\right)\right\}, \tag{3}$$

where c stands for the speed of light and $R_n = R_1 - (n-1)dsin(\theta)$ shows the distance from the nth element of the array to the target location. The array factor (AF) for FDA can be written as [32]

$$AF = \sum_{n=0}^{N-1} \exp\left\{j2\pi f_n\left(t - \frac{R_n}{c}\right)\right\}. \tag{4}$$

After placing values of f_n and R_n, (4) becomes as

$$AF = \exp\ \{j\psi_0\} \sum_{n=0}^{N-1} \exp\left\{j\frac{2\pi n}{c}\Phi\right\}, \tag{5}$$

where $\Phi = c\Delta f t - \Delta f R + df_0 sin\theta + n\Delta f d sin\theta$.

$$AF \cong \exp\ \{j\psi_1\} \frac{\sin\ [(N\pi/c)\Phi]}{\sin\ [(\pi/c)\Phi]}, \tag{6}$$

where ψ_0 stands for $2\pi f_0\ (t - R_1/c)$ and ψ_1 stands for $\psi_0 + \pi(N-1)[\Delta f R_1/c - (df_0\ sin\theta)/c - \Delta f d sin\theta/c]$, while n^2 of the fourth term $(n^2(\Delta f d \sin\theta/c))$ has been replaced by n in the fourth term $(n(\Delta f d \sin\theta/c))$ of the last expression of the AF to achieve closed form expression. So, the problem in hand is how to hide actual aircraft target from the effectiveness of the opponent FDA radar along with generation of multiple fake deceptive targets at different ranges in the direction of actual aircraft using the ULA-based FDA radar.

3. Comparison with Existing Techniques

The three deceptive jamming techniques [34–36] are selected for comparison. All of these techniques use frequency diverse arrays to generate multiple false targets. A deceptive jamming technique which is explored in [34] utilizes frequency diverse array. It generates multiple false targets at various distances across the slant range but at the same azimuthal range which aligned with the actual target as shown in the simulation (Figure 2). A good effort [35] has been made in the field of deception jamming using FDA which generates multiple fake targets at different positions in the slant range and azimuthal range to confuse the opponent radar.

Four elements array with adjacent distance of half wavelength were used in FDA with frequency offset 500 kHz, and its simulation of [35] is shown in Figure 3, which reflects actual target at slant range 7500 m and azimuthal range 0 m along with four false targets which are situated at slant ranges

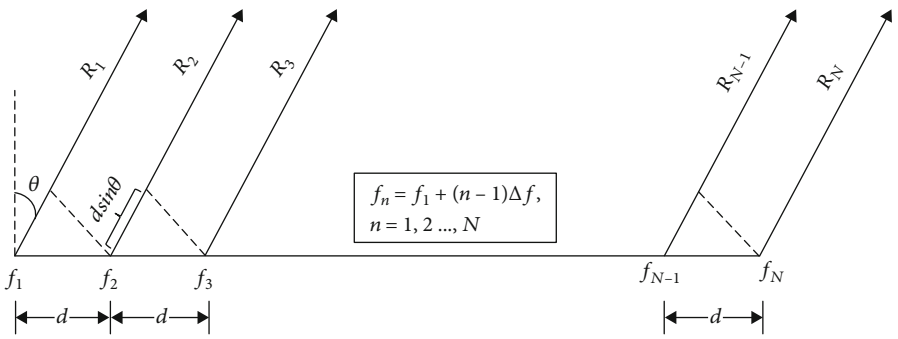


Figure 1: Geometry of the FDA radar uniformly linearly polarized with n elements.

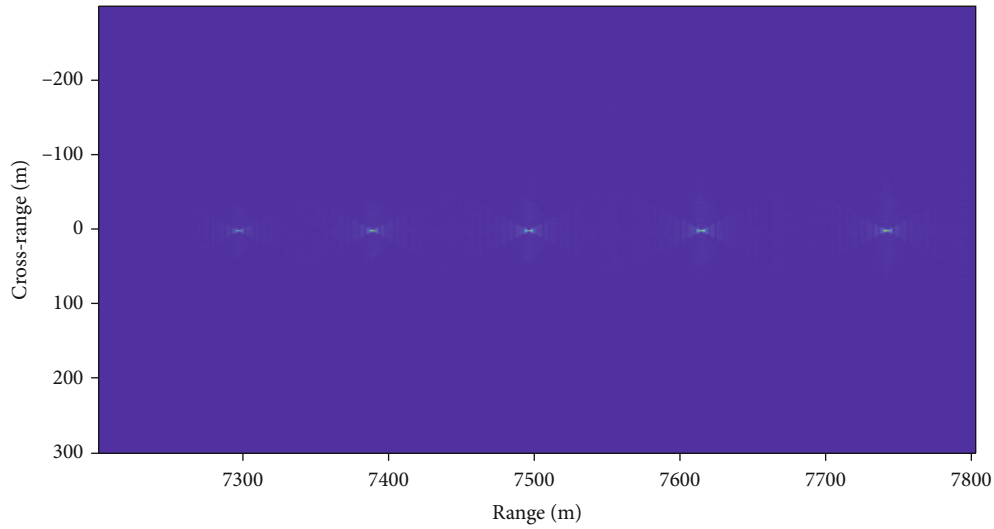


Figure 2: Actual target at (7500, 0) and false targets at {(7300, 0), (7395, 0), (7615, 0), and (7750, 0)} [34].

7300 m, 7395 m, 7615 m, and 7750 m with azimuthal range at 100 m. Figure 4 shows the effectiveness of the technique [36] by considering the jammer at the middle of the scene (7500 m, 0 m); this algorithm generates four false targets at different locations. Algorithm assumes FDA array of eight elements with frequency offset 300 kHz. It is evident from Figures 2–4 that although all three techniques [34–36] are big achievement in the field of deception jamming using frequency diverse array, none is able to hide actual target alongside generating multiple false targets.

4. Proposed Method

The key idea behind this research is to interrupt enemy radar signals and then using these signals (after proposed modifications), we hide our object (aircraft) along with generating multiple fake targets. We assumed that the proposed (modified FDA) radar is mounted on the target aircraft, and the opponent radar is placed on the ground as shown in scenario Figure 5. In purpose of hiding its own target aircraft, the interrupted signal of the opponent radar will be transmitted back after desired changes to place null at the range and the direction of the foe radar receiver. In the current scenario, it is assumed that the opponent radar is capable of transmitting and receiving FDA radiation patterns. In order to deceive the enemy radar with multiple fake targets, time-delayed replicas of the received signal will be sent towards the enemy radar. The graphical abstract of the proposed method is shown in Figure 6

4.1. Actual Target Hiding. In this section, the mathematical model is formulated to hide the actual target. Let the radiated signal from the *nth* element of the enemy FDA radar which is given as

$$s_n(t) = \exp\{j2\pi f_n t\},\ 0 \le t \le T, \tag{7}$$

where t is the time indexing within the radar pulse width T. We assume that the enemy radar is situated at distance r and at direction θ from the target. The enemy radar transmitted signal from *nth* element is received at *mth* element of the target FDA radar that can be expressed as

$$y_{m,n}(t - \tau_{m,n}) = \exp\left\{j2\pi f_n(t - \tau_{m,n})\right\}, \quad m = 1, 2 \cdots N, \tag{8}$$

where $\tau_{m,n} = [r - (n-1)d \sin\theta + d(m-1) \sin\theta]\, 1/c$. Generally, (8) can be given as

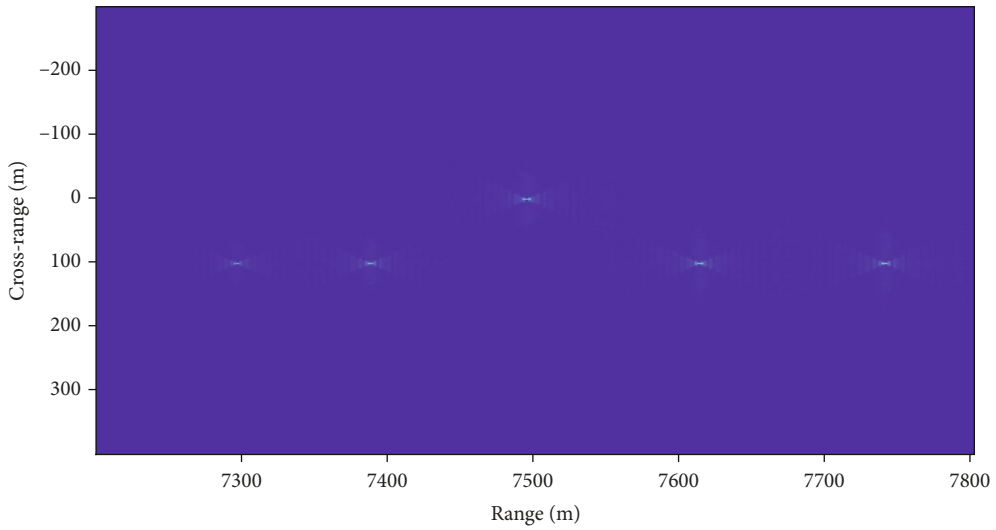


Figure 3: Actual target at (7500, 0) and false targets at {(7300, 100),(7395, 100),(7615, 100),and (7750, 100)} [35].

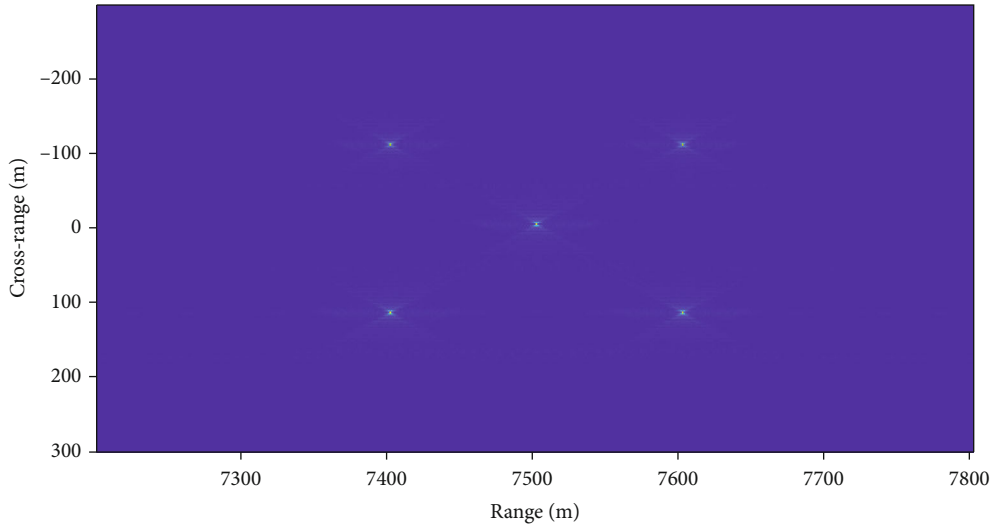


Figure 4: Actual target at (7500, 0) and false targets at {(7400, -100),(7400, 100),(7600, -100),and (7600, 100)} [36].

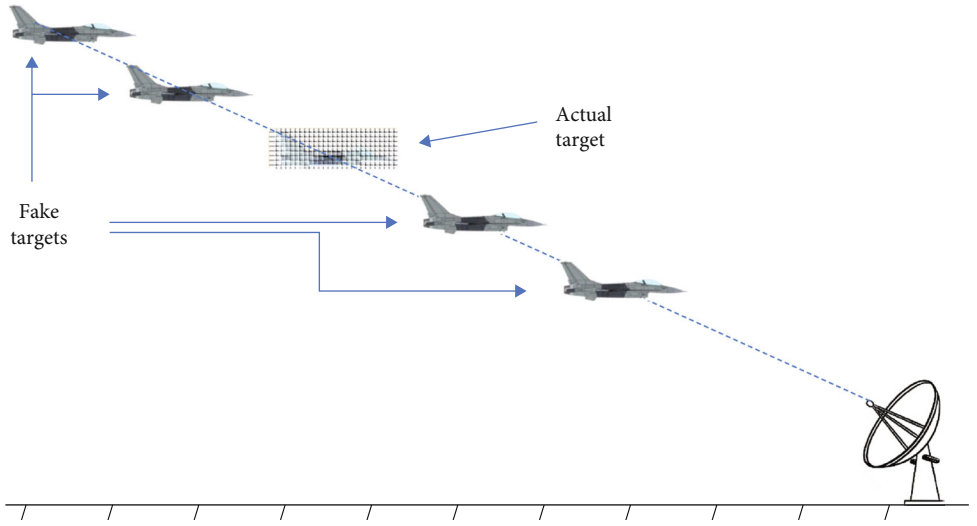


Figure 5: Proposed and adopted scenario.

$$y_m = \sum_{n=1}^{N} \exp\{j2\pi f_n(t-\tau_{m,n})\}. \quad (9)$$

Hence the received signal can be decomposed after processing through matched filtering with $\exp^*\{j2\pi f_n t\}$ as processed below.

$$\begin{aligned} y'_{m,n} &= \exp\{j2\pi f_n(t-\tau_{m,n})\} \times \exp^*\{j2\pi f_n t\}, \\ &= \exp\{j2\pi f_n t - j2\pi f_n \tau_{m,n} - j2\pi f_n t\}, \\ &= \exp\{-j2\pi f_n \tau_{m,n}\}. \end{aligned} \quad (10)$$

After placing values of $\tau_{m,n}$ and f_n, the expression becomes

$$y'_{m,n} = \exp\left\{-j2\pi (f_0 + (n-1)\Delta f)\left([r-(n-1)d\sin\theta + d(m-1)\sin\theta]\frac{1}{c}\right)\right\}, \quad (11)$$

$$= \exp\left\{-j2\pi\left[\frac{f_0 r}{c} - \frac{f_0(n-1)d\sin\theta}{c} + \frac{f_0(m-1)d\sin\theta}{c} + \frac{(n-1)\Delta f r}{c} - \frac{(n-1)^2\Delta f d\sin\theta}{c} + \frac{(n-1)(m-1)\Delta f d\sin\theta}{c}\right]\right\}. \quad (12)$$

Last two terms are insignificant; so these, terms are ignored.

$$\begin{aligned} y''_{m,n} &= \exp\left\{-j2\pi\frac{f_0 r}{c}\right\} \times \exp\left\{j2\pi\frac{f_0(n-1)d\sin\theta}{c}\right\} \\ &\quad \times \exp\left\{-j2\pi\frac{f_0(m-1)d\sin\theta}{c}\right\} \\ &\quad \times \exp\left\{-j2\pi\frac{(n-1)\Delta f r}{c}\right\}, \\ &= \exp\left\{-j2\pi\frac{r}{\lambda_0}\right\} \times \exp\left\{j2\pi\frac{(n-1)d\sin\theta}{\lambda_0}\right\} \\ &\quad \times \exp\left\{-j2\pi\frac{(m-1)d\sin\theta}{\lambda_0}\right\} \\ &\quad \times \exp\left\{-j2\pi\frac{(n-1)\Delta f r}{c}\right\}, \\ &= \exp\left\{-j2\pi\frac{r}{\lambda_0}\right\} \sum_{n=1}^{N} \exp\left\{-j2\pi\frac{\Delta f}{c}(n-1)r\right\} \exp \\ &\quad \cdot\left\{j2\pi\frac{d}{\lambda_0}(n-1)\sin\theta\right\} \\ &\quad \exp\left\{-j2\pi\frac{d}{\lambda_0}(m-1)\sin\theta\right\}. \end{aligned} \quad (13)$$

For simplicity, outside factor is eliminated.

$$y'_{m,n} = \exp\left\{-j2\pi\frac{\Delta f}{c}(n-1)r\right\} \exp\left\{j2\pi\frac{d}{\lambda_0}(n-1)\sin\theta\right\} \exp \cdot\left\{-j2\pi\frac{d}{\lambda_0}(m-1)\sin\theta\right\}. \quad (14)$$

Alternatively,

$$y'_m = \sum_{n=1}^{N} \exp\left\{-j2\pi\frac{\Delta f}{c}(n-1)r\right\} \exp\left\{j2\pi\frac{d}{\lambda_0}(n-1)\sin\theta\right\} \exp \cdot\left\{-j2\pi\frac{d}{\lambda_0}(m-1)\sin\theta\right\}, \quad (15)$$

where the first factor of the above expression dictates phase shift caused by the target range, and rest factors reflect phase shifts caused by the direction and different wave paths due to physical displacement diversity of the target and the source array elements. Now, the expression $y'_{m,n}$ can be recomposed into these factors in this form.

$$y_{m,n} = \mathrm{a_n}(\mathrm{r})\mathrm{a_n}(\theta)\mathrm{b_m}(\theta), \quad (16)$$

where $a_n(r) = \exp\{-j2\pi(\Delta f/c)(n-1)r\}$,

$$a_n(\theta) = \exp\left\{j2\pi\frac{d}{\lambda_0}(n-1)\sin\theta\right\},$$

$$b_m(\theta) = \exp\left\{-j2\pi\frac{d}{\lambda_0}(m-1)\sin\theta\right\}. \quad (17)$$

Now, we can transform the above received snapshot expression into vector form as

$$\begin{aligned} \mathbf{Y}_s &= [y_{11}, y_{12}, \cdots, y_{1N}, y_{21}, y_{22}, \cdots, y_{2N}, \cdots\cdots\cdots, y_{N1}, y_{N2}\cdots, y_{NN}]^T, \\ &= \mathbf{b}(\theta)\otimes\mathbf{a}(r,\theta), \end{aligned} \quad (18)$$

where $\mathbf{Y}_s \in \mathbb{C}^{N^2\times 1}$, $\mathbf{b}(\theta) \in \mathbb{C}^{N\times 1}$, and $\mathbf{a}(r,\theta) \in \mathbb{C}^{N\times 1}$. The superscript T and $\otimes$ reflect transpose and Kronecker product operators, respectively. The vector $\mathbf{a}(r,\theta)$ can be decomposed further in this way

$$\mathbf{a}(r,\theta) = \mathbf{a_r}(r)\odot\mathbf{a_\theta}(\theta), \quad (19)$$

where $\mathbf{a_r}(r) \in \mathbb{C}^{N\times 1}$ and $\mathbf{a_\theta}(\theta) \in \mathbb{C}^{N\times 1}$ are the range and angular steering vectors, respectively. The $\odot$ is called Hadamard product operator which reflects element-wise product between vectors. After applying DOA (direction of arrival) and range algorithms, one can find direction and range of the enemy radar, but this is beyond the scope of our research.

Now, in order to hide our target from the enemy radar, we have to transmit back the processed received signal towards the enemy radar carrying the desired information of range and direction of the enemy radar, but without considering the 3rd factor, which is not part of the deceptive jamming transmission propagation. Hence, the desired signal structure will become as follows:

$$y_n = \exp\left\{-j2\pi\frac{\Delta f}{c}(n-1)R_0\right\} \exp\left\{j2\pi\frac{d}{\lambda_0}(n-1)\sin\theta_0\right\}, \quad (20)$$

where $\theta_0 \& R_0$ are direction and range of the enemy radar, respectively. We can simplify it further in this way.

$$\boldsymbol{Y}_t = \sum_{n=1}^{N} \exp\left\{ j(n-1)\left[\frac{2\pi}{c}(\mathrm{f}_0 \mathrm{dsin}\theta_0 - \Delta f R_0)\right]\right\}. \quad (21)$$

Now, first we will verify the above result using simpler way, and then we will place null in the desired radiation pattern at a certain direction and range to cover our target. The generalized phase difference between any two elements of the FDA radar is found as

$$\Delta\varphi_{n-1,n} = \frac{2\pi d \sin\theta}{\lambda_0} - \frac{2\pi\Delta f R_1}{c} + \frac{(2n-3)2\pi\Delta f d \sin\theta}{c}. \quad (22)$$

In (22) the 3rd term which is insignificant, it can be ignored.

$$\Delta\varphi_{n-1,n}(R_1,\theta) = \frac{2\pi d \sin\theta}{\lambda_0} - \frac{2\pi\Delta f R_1}{c} = \frac{2\pi}{c}(f_0 d \sin\theta - \Delta f R_1). \quad (23)$$

The AF at range R_1 and direction θ from the radar can be expressed as

$$AF = \sum_{n=1}^{N} \exp\left\{ \mathrm{j(n-1)}\left[\frac{2\pi}{\mathrm{c}}(\mathrm{f}_0 \mathrm{dsin}\theta - \Delta f R_1)\right]\right\}. \quad (24)$$

It is considered that the enemy radar is situated at range R_0 and angle θ_0 from the target. Then, we assume that the interrupted waveform with carrier frequency f_0 can be expressed as below.

$$\mathrm{AF}(R_0,\theta_0) = \sum_{n=1}^{N} \exp\left\{ \mathrm{j(n-1)}\left[\frac{2\pi}{\mathrm{c}}(\mathrm{f}_0 \mathrm{dsin}\theta_0 - \Delta f R_0)\right]\right\}. \quad (25)$$

Now, this is the desired expression, and it is the same equation that we have concluded before. Hence, the expression of the derived AF is same to the signal which is meant for transmission towards the enemy radar. Using this AF, a null will be placed at desired range and direction to hide its own target from the vision of the enemy radar. At the end of this section, the above AF will be factorized in order to find weights for placing null at $R_0 \& \theta_0$. Now, we will explain how to generate null at R_0 by considering the simplest case. From the above expression, it is evident that the interelement phase difference is the same for the whole array between all adjacent elements; so, we can rewrite the above expression of the phase difference in the following way. After getting knowledge of above expressions, AF of the diverse frequency array can be modeled in the following way.

$$\mathrm{AF} = \sum_{n=1}^{N} z^{n-1}, \quad (26)$$

where $z = \exp(j\psi)$, $\psi = \alpha + \beta + \gamma$, $\alpha = 2\pi f_0 d \sin\theta_0/c$, and $\gamma = -2\pi\Delta f R_0/c$. Extra term β was added into the AF to get scanning capabilities. In standard practices, usually, β is added in the desired AF to steer the radiation patter of the desired communication.

The above expression is the simplest form of the AF for the FDA antenna arrays, which has fix the main beam direction and null directions. Steering of the beam pattern for such an array is not possible. In other words, to get control over the nulls of the radiation pattern, we need to plug in and update weights of the expression. Now, to steer the beam pattern (main beam and nulls) of the frequency diverse array, appropriate weights are necessary. Further, above AF must be put to equal to zero for the calculation of the appropriate weights to steer the beam pattern towards the desired directions.

$$AF = \sum_{n=1}^{N} A_{n-1} z^{n-1} = 0. \quad (27)$$

In determining of the weights, we considered the simplest case to avoid complexity of the proposed method. To find weights of the desired radiation pattern, we will follow this procedure.

$$AF = (z - r_0)\left(\sum_{n=1}^{N-1} z^{n-1}\right) = 0, \quad (28)$$

where $r_0 = e^{j(\alpha_0+\beta_s+\gamma_0)}$ is the desired null. Here, $\alpha_0 = (2\pi/c) f_1 d \sin\theta_0$, $\gamma_0 = -\Delta f R_0$, $\beta_s = (2\pi/c) f_1 d \sin\theta_s$, and θ_s are directions of the main beam. Then, the desired expression with updated weights will be as follows.

$$AF = \sum_{n=1}^{N} A_{n-1} z^{n-1}, \quad (29)$$

where $A_0 = -r_0$, $A_i = 1 - r_0$, $i = 1, 2, \cdots, N-2$, and $A_{N-1} = 1$. After applying these weights into the AF, we will be able to steer the null towards its desired direction θ_0 and range R_0. By following the above mechanism, our proposed method is capable of camouflaging its own target from the lethalness of the enemy radar.

4.2. Displaying Multiple Fake Targets. In second part of the proposed method, we will display multiple false targets to the enemy radar. For the purpose of multiple fake deceptive targets, we will process the received signal with two steps: first, the received signal will be time modulated with appropriate delays, and then power will be maximized and transmitted towards the enemy radar. We assume that the signal transmitted from the *nth* element of the enemy radar and received at the *mth* element of the false target generator (FTG) which is located at target that is

$$y_{j,m,n}(t) = \exp\left\{ j2\pi f_n \left(t - \tau_{j,m,n}\right)\right\}, \quad (30)$$

where

$$\begin{aligned}\tau_{j,m,n} &= \tau_j - \tau_{j,n} + \tau_{j,m},\\ &= \left[r_j - (n-1)d\sin\theta_j + (m-1)d\sin\theta_j\right]1/c\\ &= \left[r_j + (m-n)d\sin\theta_j\right]1/c.\end{aligned} \tag{31}$$

We suppose that the locally generated adjustable oscillator frequency is f_{FTG}; the receiving and processing time delay of the enemy radar signal is τ_{FTG}. Thus, the received signal can be down converted as follows.

$$y_{j,m,n} = \exp\left\{j2\pi f_n\left(t-\tau_{j,m,n}\right)\right\}\times\exp\left\{-j2\pi f_{\mathrm{FTG}}(t-\tau_{\mathrm{FTG}})\right\}. \tag{32}$$

As we know that our FTG is working on the principle of the frequency diverse array radar, so we can reflect that $f_{\mathrm{FTG}} = f_n$. After this change, the expression can be simplified in this way.

$$\begin{aligned}y_{j,m,n} &= \exp\left\{j2\pi f_n\left(t-\tau_{j,m,n}\right)\right\}\times\exp\left\{-j2\pi f_n(t-\tau_{\mathrm{FTG}})\right\},\\ &= \exp\left\{j2\pi f_n\left(\tau_{FTG}-\tau_{j,m,n}\right)\right\}.\end{aligned} \tag{33}$$

In order to generate multiple fake targets, signal will pass through the process of time modulation. Different time delays are made as follows.

$$\Delta\tau_{\mathrm{ft},i} = \frac{2\Delta r_{\mathrm{ft},i}}{c},\quad i = 1, 2, \cdots\cdots, k = \text{number of false targets}, \tag{34}$$

where $\Delta r_{\mathrm{ft},i} = r_{\mathrm{ft},i} - r_j$ or $r_{\mathrm{ft},i} = r_j + \Delta r_{\mathrm{ft},i}$, while $\Delta\tau_{\mathrm{ft},i}$, $\Delta r_{\mathrm{ft},i}$, $r_{\mathrm{ft},i}$, and r_j represent new deceptive time delay increment, deceptive range increment relative to FTG, range of false targets relative to enemy radar, and reference range between enemy radar and jammer, respectively. Hence, the new updated signal will become as follows.

$$y_{\mathrm{ft},i} = \exp\left\{j2\pi f_n\left(\tau_{\mathrm{FTG}}-\tau_{j,m,n}-\Delta\tau_{\mathrm{ft},i}\right)\right\}. \tag{35}$$

Before the transmission of the signal towards the enemy radar, it must undergo carrier modulation with locally generated adjustable carrier frequency. After modulation, signal will be thrown to the enemy radar.

$$\begin{aligned}y_{\mathrm{ft},i}(\mathrm{t}) &= \exp\left\{j2\pi f_n\left(\tau_{\mathrm{FTG}}-\tau_{j,m,n}-\Delta\tau_{\mathrm{ft},i}\right)\right\}\times\exp\left\{j2\pi f_n(t-\Delta\tau_{\mathrm{ft},i})\right\},\\ &= \exp\left\{j2\pi f_n\left(\tau_{FTG}-\tau_{j,m,n}-\Delta\tau_{\mathrm{ft},i}+t-\Delta\tau_{\mathrm{ft},i}\right)\right\},\\ &= \exp\left\{j2\pi f_n\left(t+\tau_{FTG}-\tau_{j,m,n}-2\Delta\tau_{\mathrm{ft},i}\right)\right\}.\end{aligned} \tag{36}$$

The signal received by the enemy radar is

$$\begin{aligned}y_{\mathrm{ft},i,m,n}\left(\mathrm{t}-\tau_{\mathrm{j,m,n}}\right) &= \exp\left\{j2\pi f_n\left(t+\tau_{\mathrm{FTG}}-\tau_{j,m,n}-2\Delta\tau_{\mathrm{ft},i}-\tau_{j,m,n}\right)\right\},\\ &= \exp\left\{j2\pi f_n\left(t+\tau_{\mathrm{FTG}}-2\tau_{j,m,n}-2\Delta\tau_{\mathrm{ft},i}\right)\right\}.\end{aligned} \tag{37}$$

After passing this signal through the matched filter, we get

$$\begin{aligned}y_{\mathrm{ft},i,m,n} &= \exp\left\{j2\pi f_n\left(t+\tau_{\mathrm{FTG}}-2\tau_{j,m,n}-2\Delta\tau_{\mathrm{ft},i}\right)\right\}\times\exp\left\{-\mathrm{j}2\pi\mathrm{f_n}t\right\},\\ &= \exp\left\{j2\pi f_n\left(\tau_{\mathrm{FTG}}-2\tau_{j,m,n}-2\Delta\tau_{\mathrm{ft},i}\right)\right\}.\end{aligned} \tag{38}$$

Now, place values of $\tau_{j,m,n}$ and $\Delta\tau_{\mathrm{ft},i}$, and the received signal will become as

$$\begin{aligned}y_{\mathrm{ft},i,m,n} &= \exp\left\{j2\pi f_n\left(\tau_{\mathrm{FTG}}-2\left[\tau_j-\tau_{j,n}+\tau_{j,m}\right]-2\frac{2\Delta r_{\mathrm{ft},i}}{c}\right)\right\},\\ &= \exp\left\{j2\pi f_n\left(\tau_{\mathrm{FTG}}-\frac{2}{\mathrm{c}}\left[r_j-(n-1)d\sin\theta_j\right.\right.\right.\\ &\qquad\left.\left.\left.+(m-1)d\sin\theta_j\right]-\frac{4}{c}\left[r_{\mathrm{ft},i}-r_j\right]\right)\right\}.\end{aligned} \tag{39}$$

As we know that $f_n = f_0 + (n-1)\Delta f$,

$$\begin{aligned}y_{\mathrm{ft},i,m,n} &= \exp\left\{j2\pi(f_0+(n-1)\Delta f)\begin{bmatrix}\tau_{\mathrm{FTG}}-\frac{2\mathrm{r_j}}{\mathrm{c}}+\frac{2(n-1)d\sin\theta_j}{c}\\ -\frac{2(m-1)d\sin\theta_j}{c}-\frac{4r_{\mathrm{ft},i}}{c}+\frac{4r_j}{c}\end{bmatrix}\right\},\\ &= \exp\left\{j2\pi(f_0+(n-1)\Delta f)\left[\tau_{\mathrm{FTG}}+\frac{2\mathrm{r_j}}{\mathrm{c}}+\frac{2(n-1)dsin\theta_j}{c}-\frac{2(m-1)dsin\theta_j}{c}-\frac{4r_{\mathrm{ft},i}}{c}\right]\right\},\\ &= \exp\left\{j2\pi\begin{bmatrix}f_0\tau_{\mathrm{FTG}}+f_0\frac{2\mathrm{r_j}}{\mathrm{c}}+\frac{2(n-1)f_0d\sin\theta_j}{c}-\frac{2(m-1)f_0d\sin\theta_j}{c}\\ -\frac{4f_0r_{\mathrm{ft},i}}{c}+(n-1)\Delta f\tau_{\mathrm{FTG}}+\frac{2(n-1)\Delta f\mathrm{r_j}}{\mathrm{c}}+\frac{2(n-1)^2\Delta fd\sin\theta_j}{c}\\ -\frac{2(n-1)(m-1)\Delta fd\sin\theta_j}{c}-\frac{4(n-1)\Delta fr_{\mathrm{ft},i}}{c}\end{bmatrix}\right\}.\end{aligned} \tag{40}$$

Terms I and II will be taken outside, while terms VI, VIII, and IX will be ignored due to their insignificance.

$$y_{\mathrm{ft},i,m} = \exp\{\mathrm{j}2\pi f_0 \tau_{\mathrm{FTG}}\} \times \exp\left\{\mathrm{j}4\pi \frac{f_0 \mathrm{r_j}}{\mathrm{c}}\right\} \sum_{n=1}^{N} \exp\left\{ j2\pi \left[\begin{array}{l} \frac{2(n-1)f_0 d \sin\theta_j}{c} - \\ \frac{2(m-1)f_0 d \sin\theta_j}{c} - \frac{4f_0 r_{\mathrm{ft},i}}{c} + \\ \frac{2(n-1)\Delta f \mathrm{r_j}}{\mathrm{c}} - \frac{4(n-1)\Delta f r_{\mathrm{ft},i}}{c} \end{array} \right] \right\},$$
$$= \mathrm{A}_\varphi \sum_{n=1}^{N} \exp\left\{ \frac{j4\pi}{c} \left[\begin{array}{l} (n-1)f_0 d \sin\theta_j - (m-1)f_0 d \sin\theta_j \\ -2f_0 r_{\mathrm{ft},i} + (n-1)\Delta f \mathrm{r_j} - 2(n-1)\Delta f r_{\mathrm{ft},i} \end{array} \right] \right\}, \tag{41}$$

where $\mathrm{A}_\varphi = \exp\{\mathrm{j}2\pi f_0 \tau_{\mathrm{FTG}}\} \times \exp\{\mathrm{j}4\pi(f_0 \mathrm{r_j}/\mathrm{c})\}$ represents the phase change due to the enemy radar to jammer reference distance and time due to signal interception and processing time delay. For simplicity, we can ignore A_φ.

$$y_{\mathrm{ft},i,m,n} = \exp\left\{ j4\pi \left[\begin{array}{l} \frac{(n-1)f_0 d \sin\theta_j}{c} - \frac{(m-1)f_0 d \sin\theta_j}{c} \\ -\frac{2f_0 r_{\mathrm{ft},i}}{\mathrm{c}} + \frac{(n-1)\Delta f \mathrm{r_j}}{c} - \frac{2(n-1)\Delta f r_{\mathrm{ft},i}}{c} \end{array} \right] \right\}. \tag{42}$$

Now, we can deform above result into two factors: direction and range.

$$y_{\mathrm{ft},i,m,n} = \exp\left\{ -j4\pi \left[\begin{array}{l} \left(\frac{(n-1)f_0 d \sin\theta_j}{c} - \frac{(m-1)f_0 d \sin\theta_j}{c}\right) \\ \times \left(\frac{2f_0 r_{\mathrm{ft},i}}{\mathrm{c}} + \frac{2(n-1)\Delta f r_{\mathrm{ft},i}}{c} - \frac{(n-1)\Delta f \mathrm{r_j}}{c}\right) \end{array} \right] \right\}. \tag{43}$$

For the sake of simplicity, we can ignore the 1st factor which represents radiation pattern due to direction, and we consider only the 2nd factor here which is our current area of discussion.

$$y_{\mathrm{ft},i,m,n} = \exp\left\{ -j4\pi \left[\frac{2f_0 r_{\mathrm{ft},i}}{\mathrm{c}} + \frac{2(n-1)\Delta f r_{\mathrm{ft},i}}{c} - \frac{(n-1)\Delta f \mathrm{r_j}}{c} \right] \right\}. \tag{44}$$

Now, we have these three terms which are range-dependent and playing important role of determining the radiation pattern of the desired signal. It is evident that first and second terms are representing fake target ranges while third term is representing actual target range. Third term is the most insignificant term as relative to other contributing terms, and due to this reason, its effect will be overcome by other terms, or it can be simply ignored.

$$y_{\mathrm{ft},i,m,n} = \exp\left\{ -j4\pi \left[\frac{2f_0 r_{\mathrm{ft},i}}{\mathrm{c}} + \frac{2(n-1)\Delta f r_{\mathrm{ft},i}}{c} \right] \right\}. \tag{45}$$

Hence, the enemy radar will observe the fake targets rather than observing the actual target.

Table 1: Parameters of actual and false targets.

Case#	Parameter	Actual target	FT 1	FT 2	FT 3	FT 4
1	Range (km)	50	30	40	60	70
	Angle (degrees)	50	50	50	50	50
2	Range (km)	40	30	50	60	70
	Angle (degrees)	10	10	10	10	10
3	Range (km)	70	40	50	60	80
	Angle (degrees)	40	40	40	40	40
4	Range (km)	60	40	50	70	80
	Angle (degrees)	20	20	20	20	20

5. Simulations

Our objective in this simulation is to offer deception jamming to the enemy radars. Consider the proposed scenario as shown in Figure 5 which is based on the surface to the air signal model. The actual aircraft is situated in air far-zone field while the opponent FDA radar is located at the surface. The proposed deceptive jammer is mounted on the actual aircraft. Both rivals are working on W-band ($f_0 = 100\ GHz$) of the FDA radar having N number of isotropic antenna elements with equal interelement spacing $\lambda/2$ and uniform current distribution along the whole linear array geometry configuration. The frequency increment of the FDA radars is kept$\Delta f = 0.3$ KHz.

Our proposed deceptive jammer works in passive searching mode that means it does not transmit and receive own signals to scan opponent radar. It instead utilizes the opponent radar signals to trace the desired parameters. But as quick as the computational system at the target finds direction of arrival (DOA), range, and pulse repetition interval (PRI) of the opponent radar, our proposed method can send deceptive echoes towards the opponent radar. These deceptive-echoes carry radiation pattern with appropriate null linked to the desired range and direction. They will be transmitted back in synchronous with actual target echoes to hide the actual target and delayed versions of false echoes to generate multiple false targets. Afterwards, in result of the proposed algorithm, the opponent radar will not be able to navigate the actual aircraft. So, track mode of the opponent radar will not work here.

To avoid any waveform time-dependent periodicity and mutual interference between the array element pulse width of the echo that is kept $\leq 30\ \mu s$ and for the sake of simplicity, the PRI is taken 0.5 ms. Four different simulation cases have been considered. In each case, one true target is assumed, and the proposed deceptive jammer is mounted on it. The technique hides the actual aircraft along with generating four false targets in each test case with the parameters given in Table 1. For 3-D simulations, ten radiating antenna elements are considered, but for the 2-D simulations, we have taken $N = 10, 20, 30, 40, 50$ isotropic antenna elements, in the ULA-based FDA radar for each case.

6. Case-I

In first case, we assume that the actual target aircraft is situated at distance 50 km and direction 50 degrees. Frequency

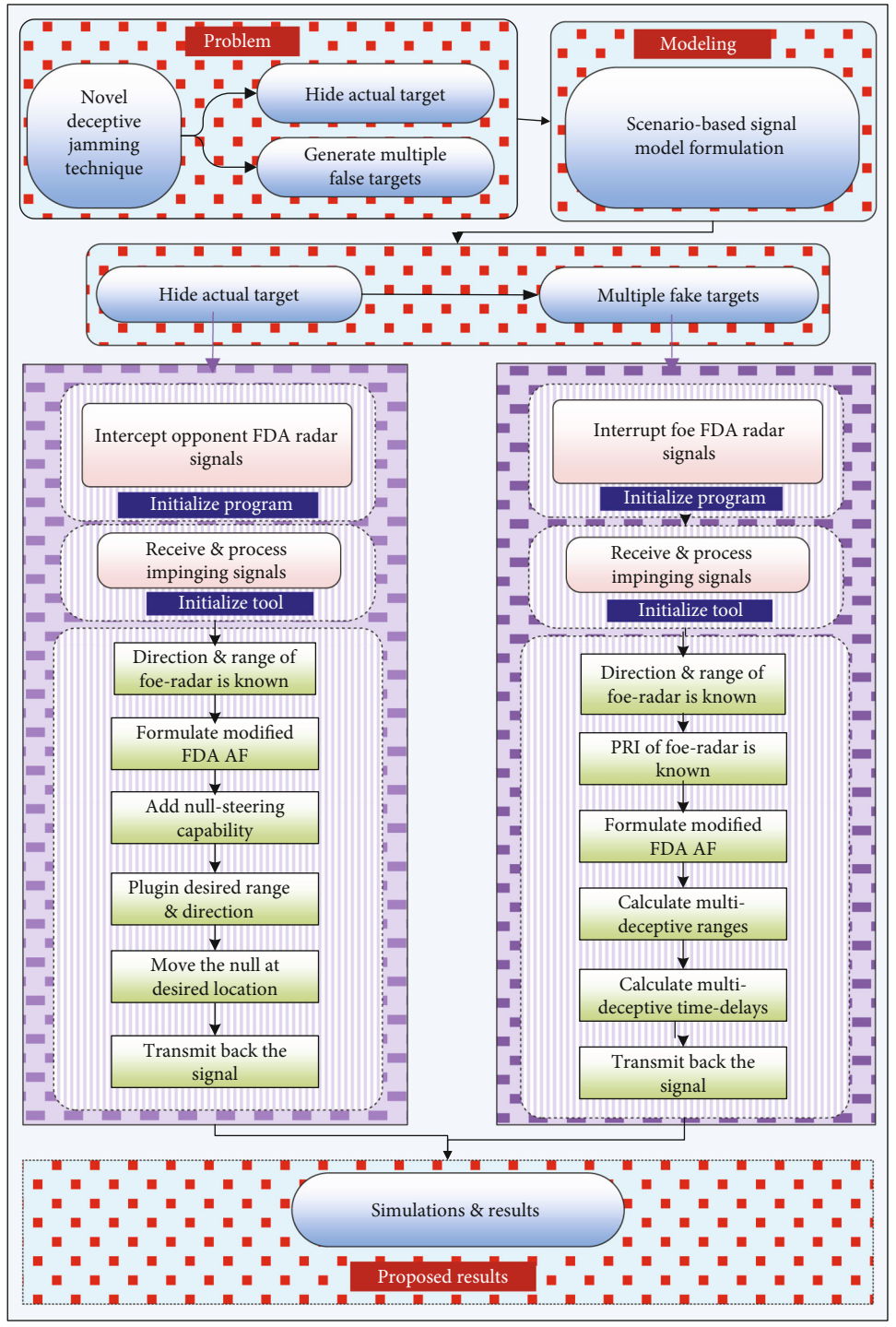


Figure 6: Proposed model.

of the first element in the array is taken 100 GHz, while incremental frequency is assumed 0.3 kHz. Figure 7 shows 3-D simulation of the case where null has been placed at the real target aircraft location in order to hide it from the enemy radar. We have considered different numbers of antenna-elements in the array of ULA-based FDA radar for 2-D simulations.

Instead of showing results like Figures 7, 11, 15, and 19 in 3-D, we have plotted the results in more simplified way using only output power in Figures 10, 14, 18, and 22, respectively. These figures (10, 14, 18, and 22) simply further elaborate the results of Figures 7, 11, 15, and 19, respectively. These figures show two scenarios. In the first scenario, it hides the actual target by simulating equation (29), whereas eq. (29) places null in the received signal at the opponent radar's location by means when the opponent receives this signals, he will perceive min. power at the target location. In the second scenario, it generates false targets at certain ranges and angles by

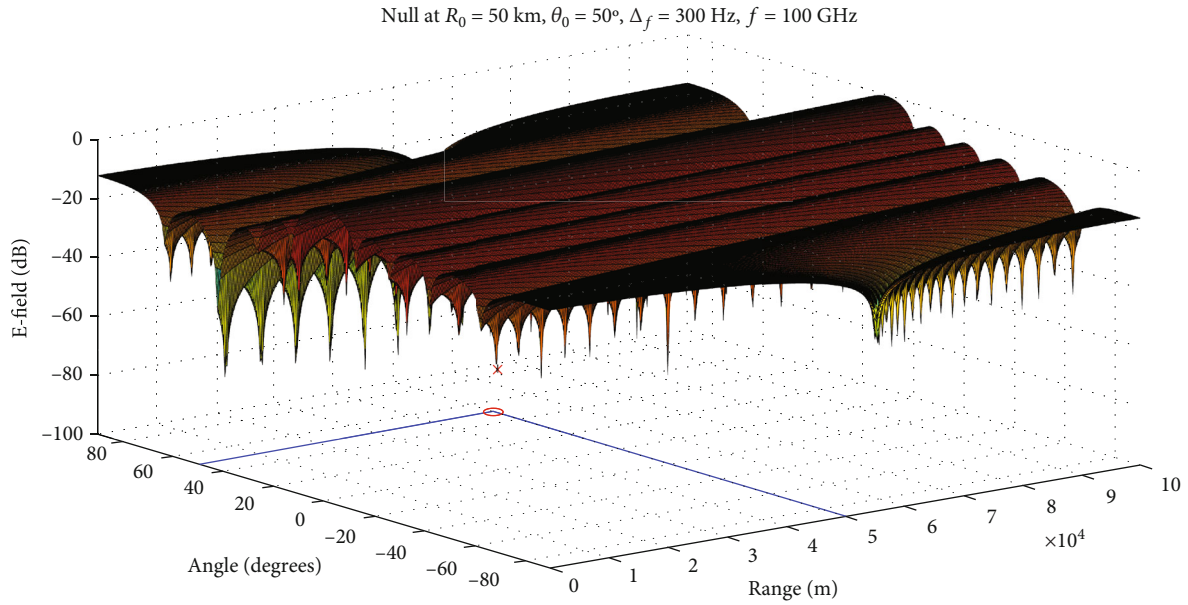


Figure 7: Actual target hides at distance 50 km and direction 50°.

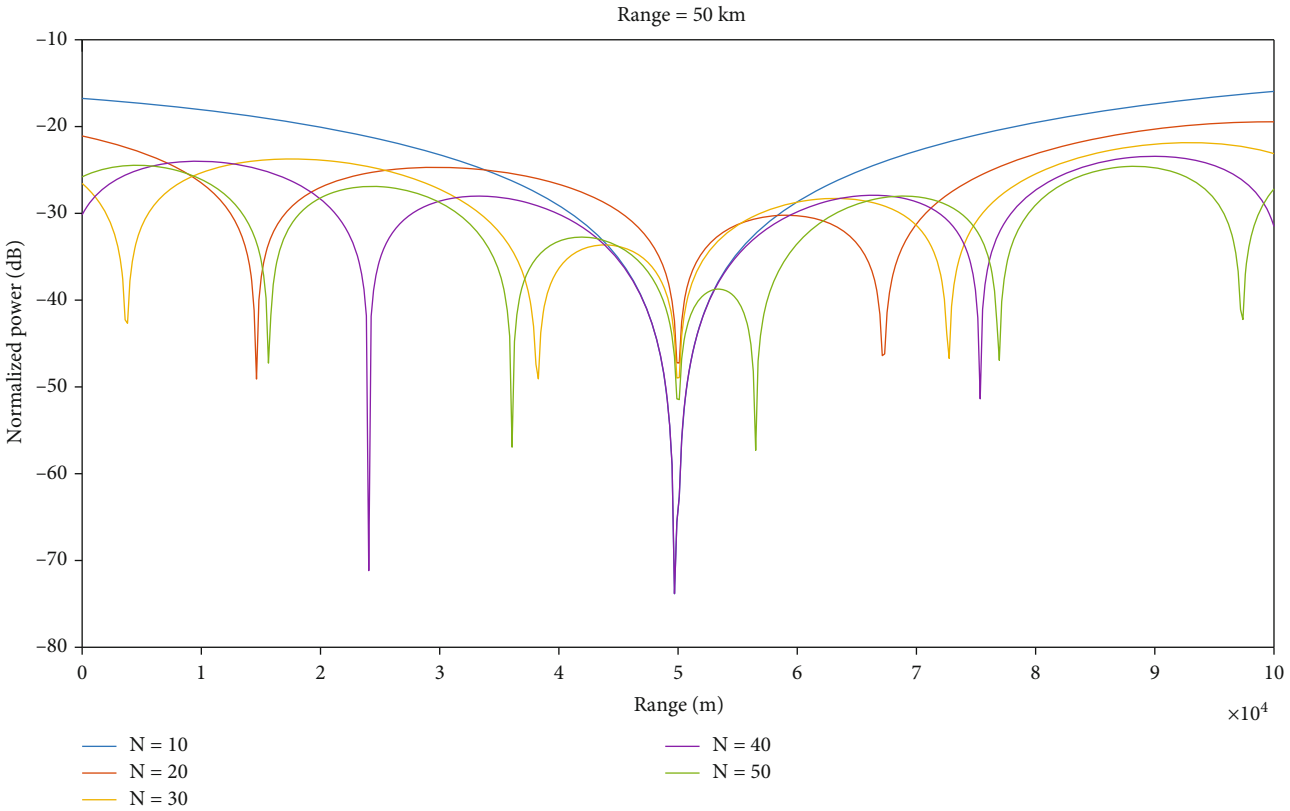


Figure 8: Null's placement at opponent radar's range with different numbers of antenna elements.

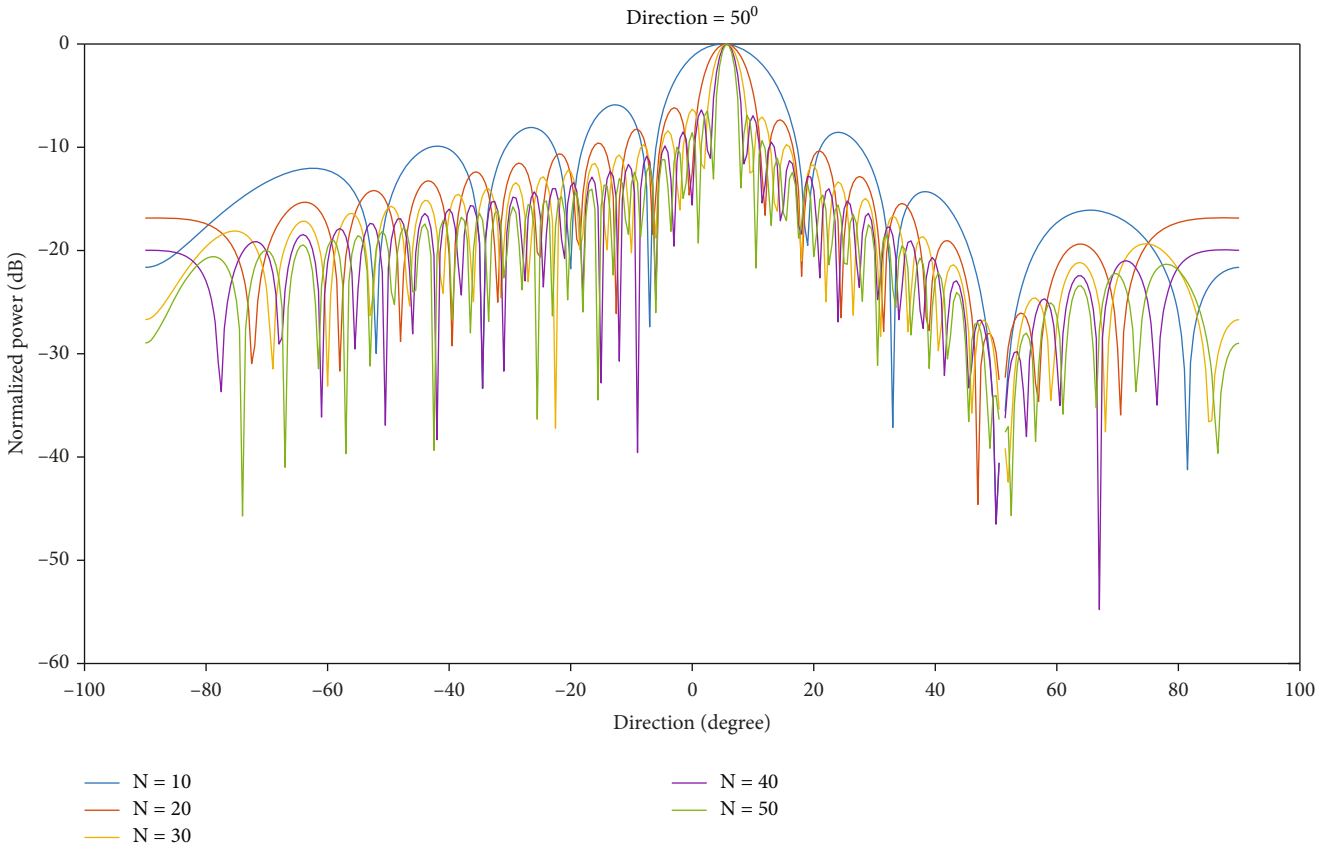


Figure 9: Null's placement at opponent radar's direction with different numbers of antenna elements.

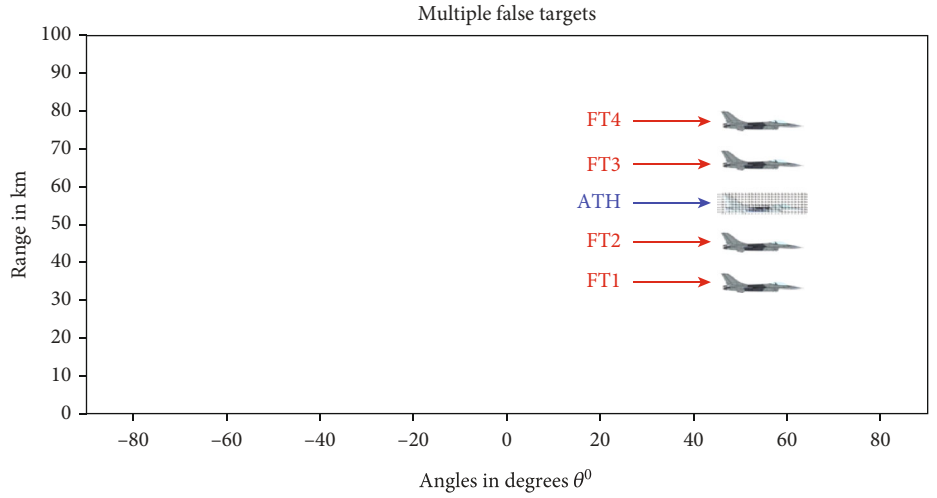


Figure 10: Actual target hides at 50 km, and false targets appear at {30, 40, 60, 70} km.

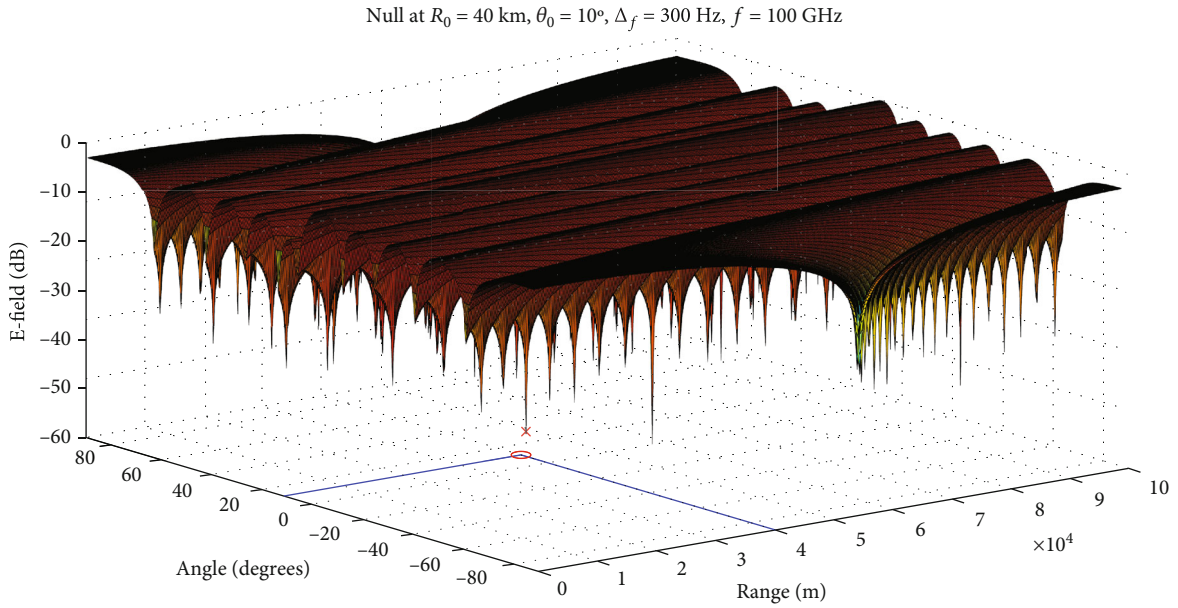


Figure 11: Actual target hides at distance 40 km and direction 10°.

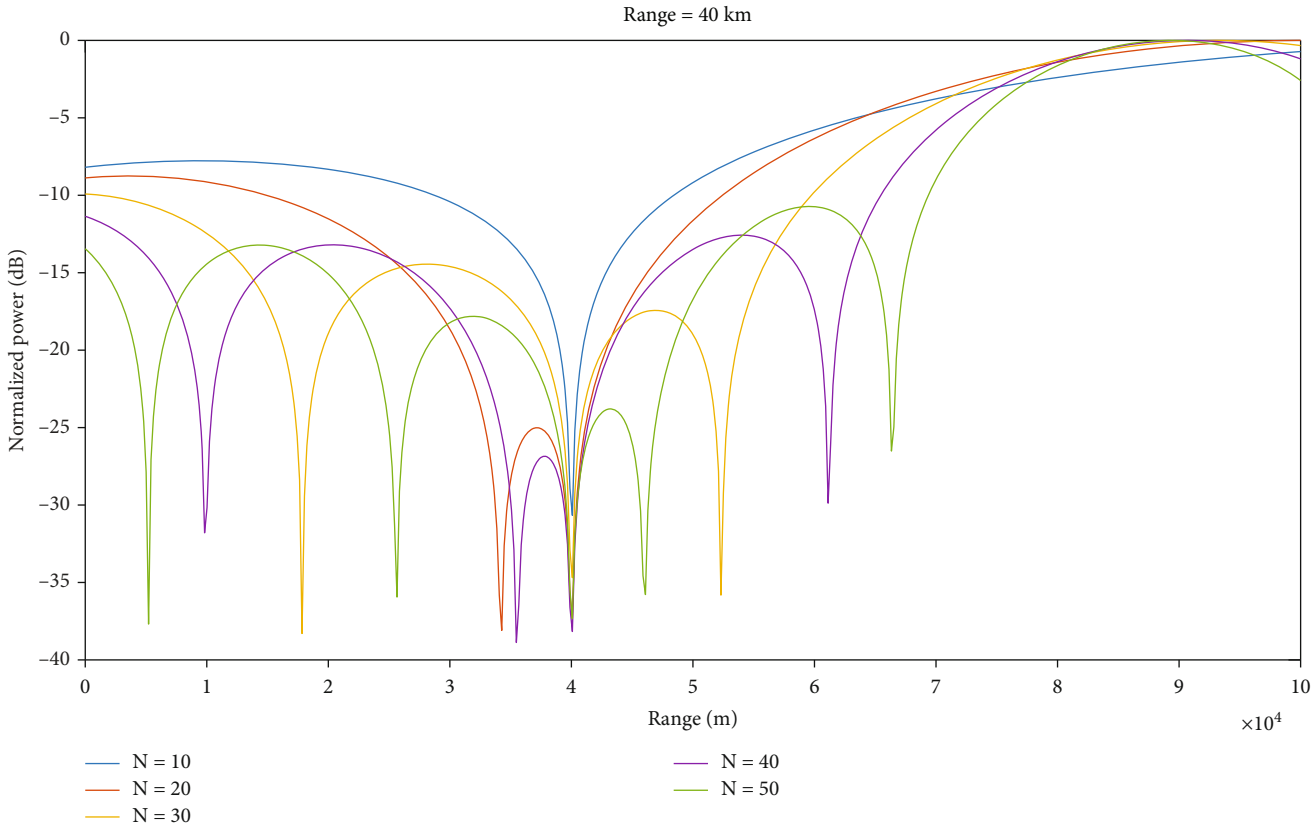


Figure 12: Null's placement at opponent radar's range with different numbers of antenna elements.

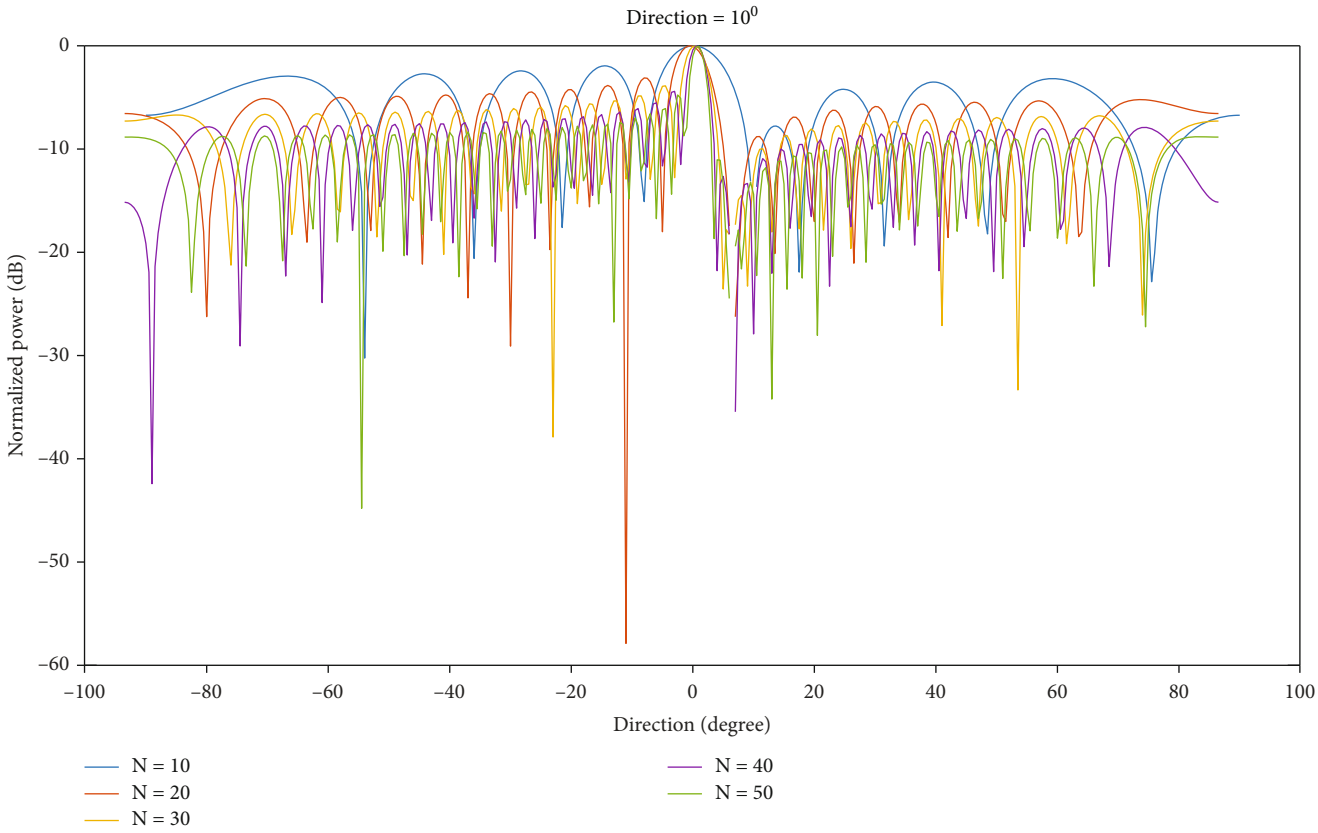


Figure 13: Null's placement at opponent radar's direction with different numbers of antenna elements.

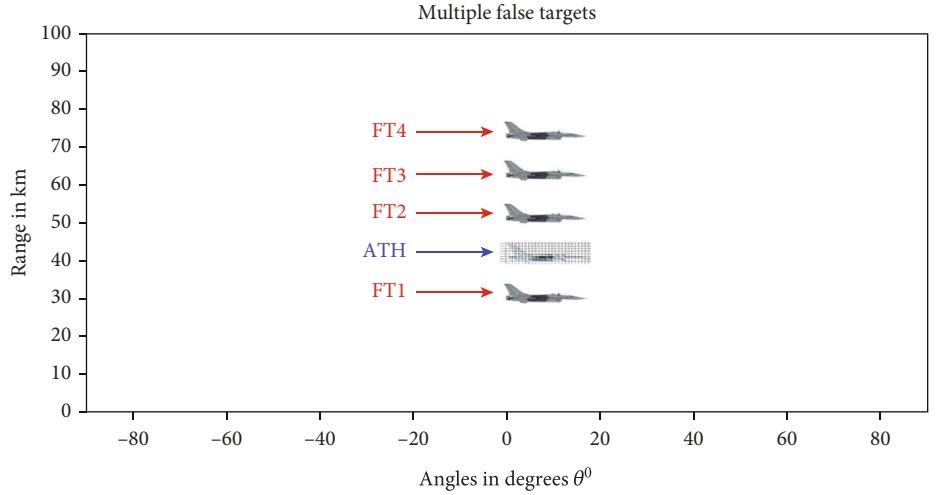


Figure 14: Actual target hides at 40 km, and false targets appear at {30, 50, 60, 70} km.

evaluating equation (45) which will offer maximum power at different ranges.

Figure 8 reflects 2-D results of the null position at the range of the actual aircraft to camouflage it, while Figure 9 also verifies our results of null position at the desired direction of the real target to cover it. The proposed method also generates four fake targets along the same direction of the actual target and at different ranges of 30 km, 40 km, 60 km, and 70 km. This has been proved in Figure 10 where the desired target is hidden along with fake targets at their respective ranges.

7. Case-II

For this case, we have assumed that the real-target aircraft is located at range 40 km and direction 10 degree. Carrier frequency and incremental frequency, f_0 and Δf, are considered 100 GHz and 0.3 kHz, respectively. Figure 11 represents a 3-

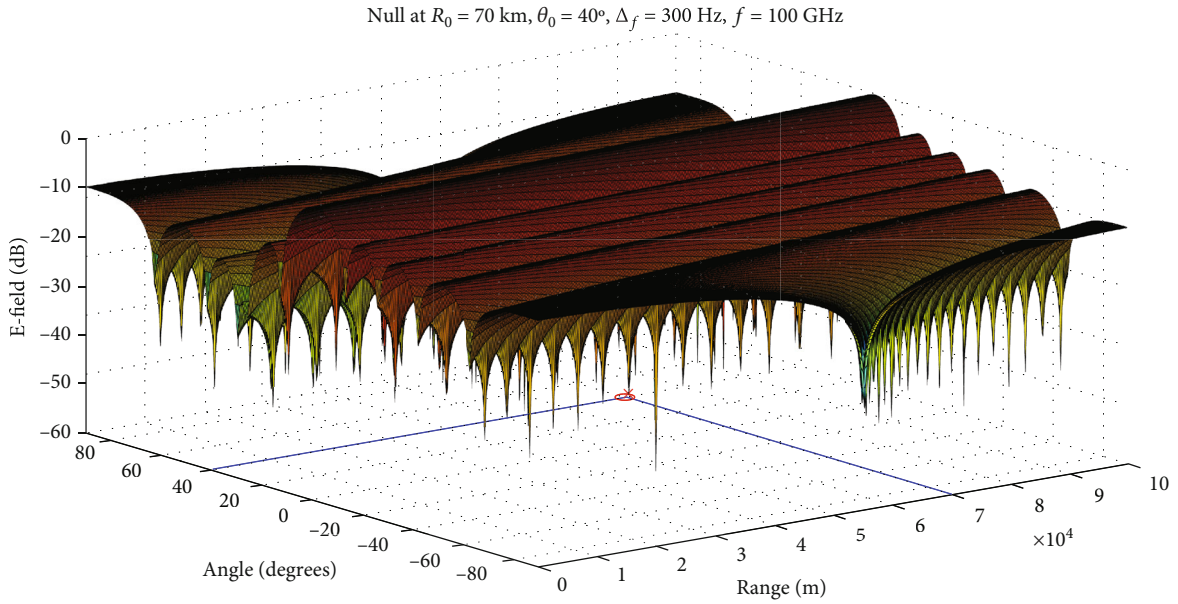


Figure 15: Actual target hides at distance 70 km and direction 40°.

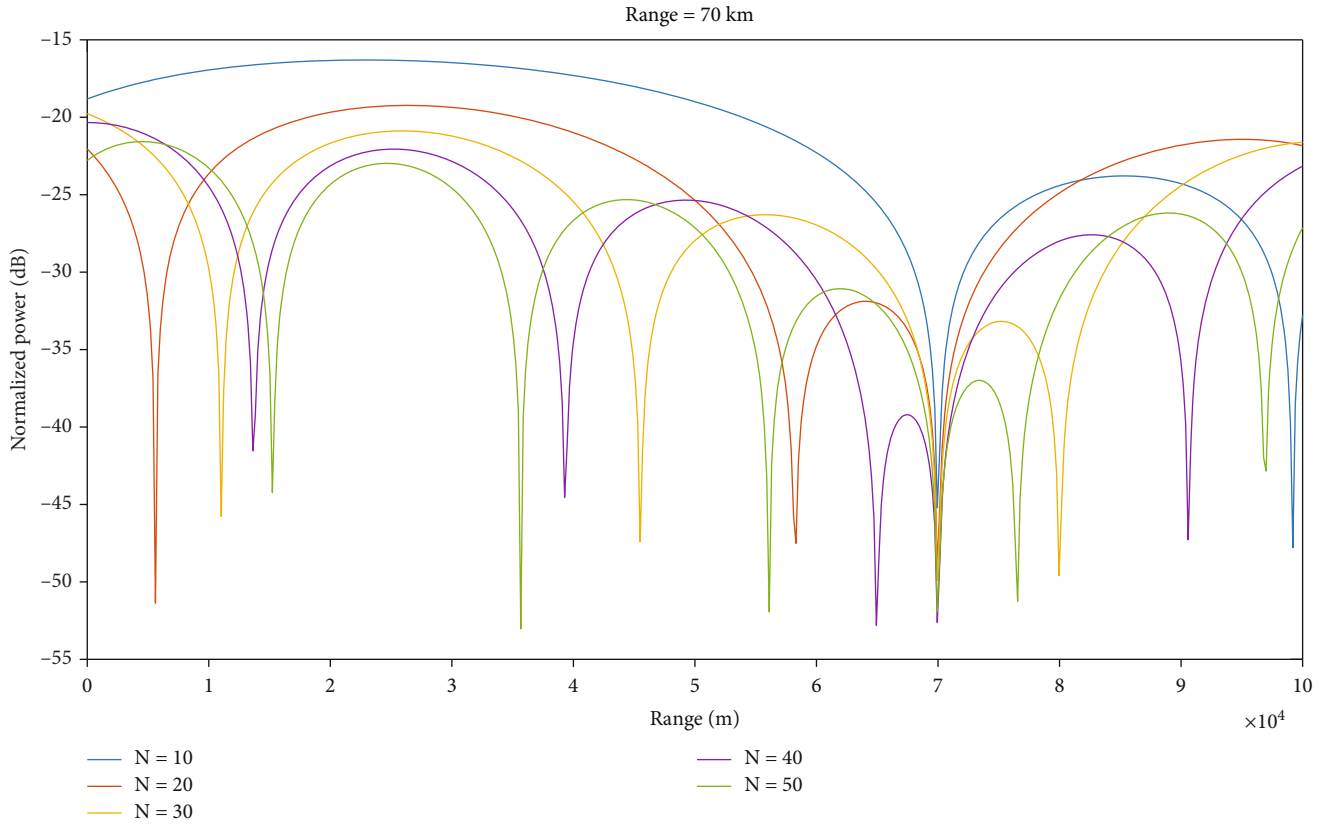


Figure 16: Null's placement at opponent radar's range with different numbers of antenna elements.

D simulation where the desired null has been placed at the real target aircraft location to hide it from the opponent radar. Two-dimensional results of the null position at the desired range (40 km) of the actual-aircraft are shown in Figure 12. In another graph, we have proved that the null has been placed at the desired direction of the actual target aircraft reflected in Figure 13. Further, our technique also generates four false targets along the direction 10 degrees but at ranges of 30 km, 50 km, 60 km, and 70 km as shown in Figure 14.

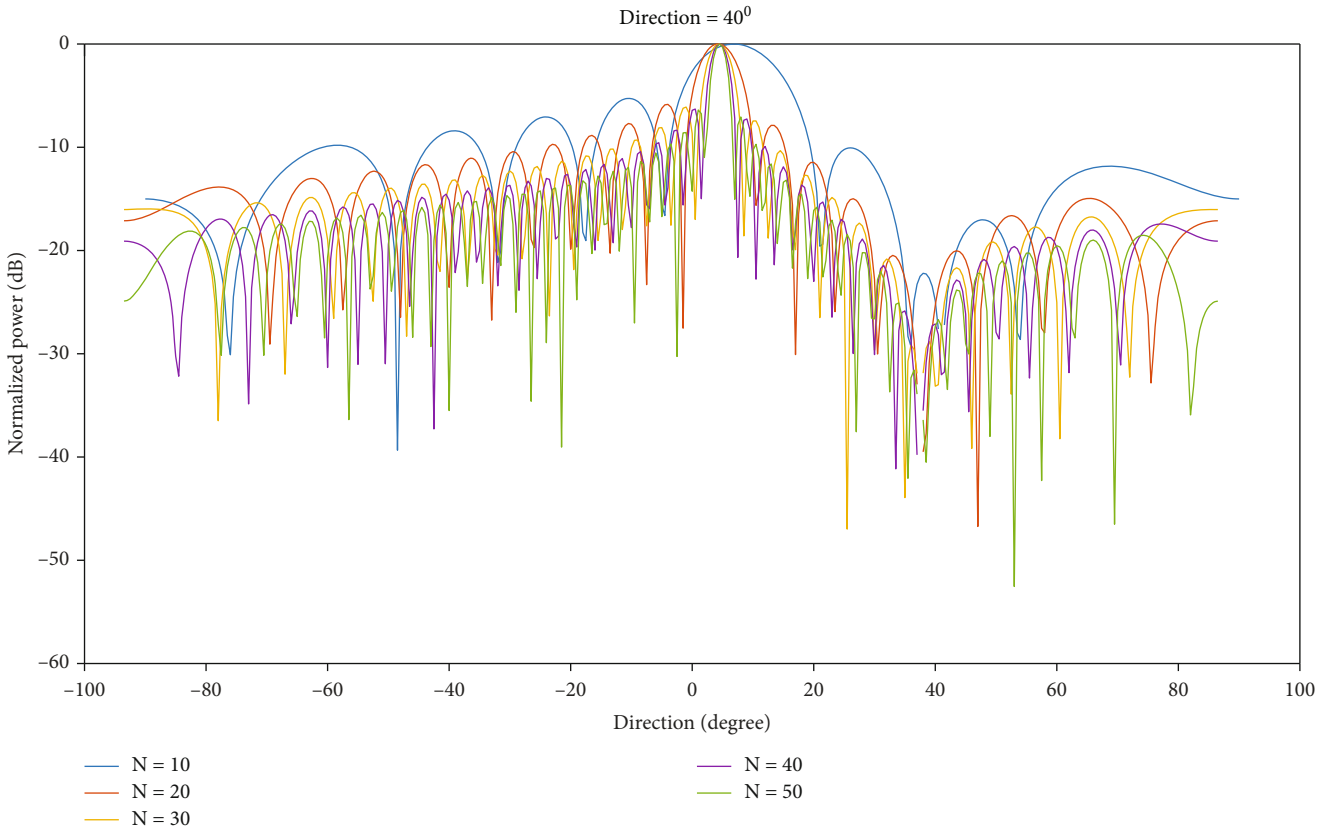


Figure 17: Null's placement at opponent radar's direction with different numbers of antenna elements.

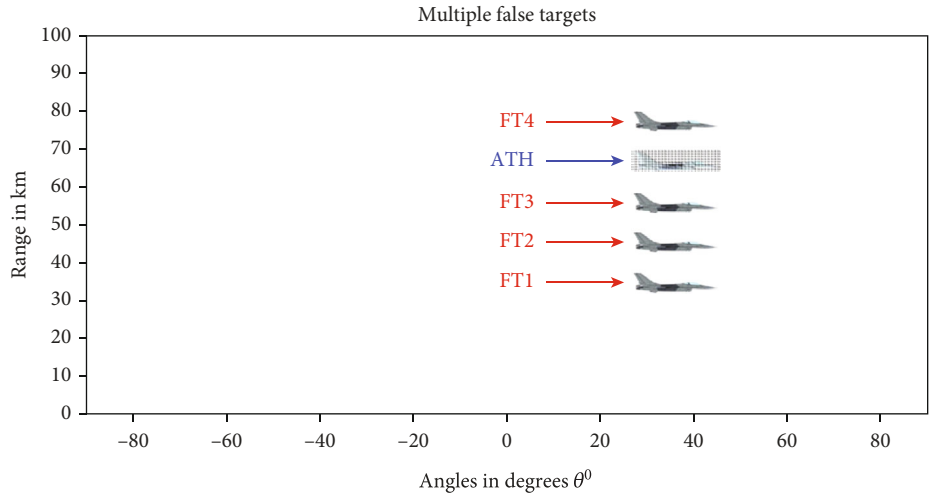


Figure 18: Actual target hides at 70 km, and false targets appear at {40, 50, 60, 80} km.

8. Case-III

In 3rd case, we have assumed that the real target is positioned at range of 70 km from the opponent radar along the direction 40 degrees. Our proposed technique can verify the results in Figure 15 by showing that the desired null has been place effectively at the actual target location. Its two-dimensional counterpart graphs are shown in Figures 16 and 17 to validate its correctness in the desired range and direction, respectively. Figure 18 shows multiple fake targets in the direction of the real target but at distances 40 km, 50 km, 60 km, and 80 km away from the opponent radar.

9. Case-IV

In the last case, we have taken the actual target aircraft at distance 60 km away from the foe radar and at direction 20 degrees. The proposed model draws a null at its location to

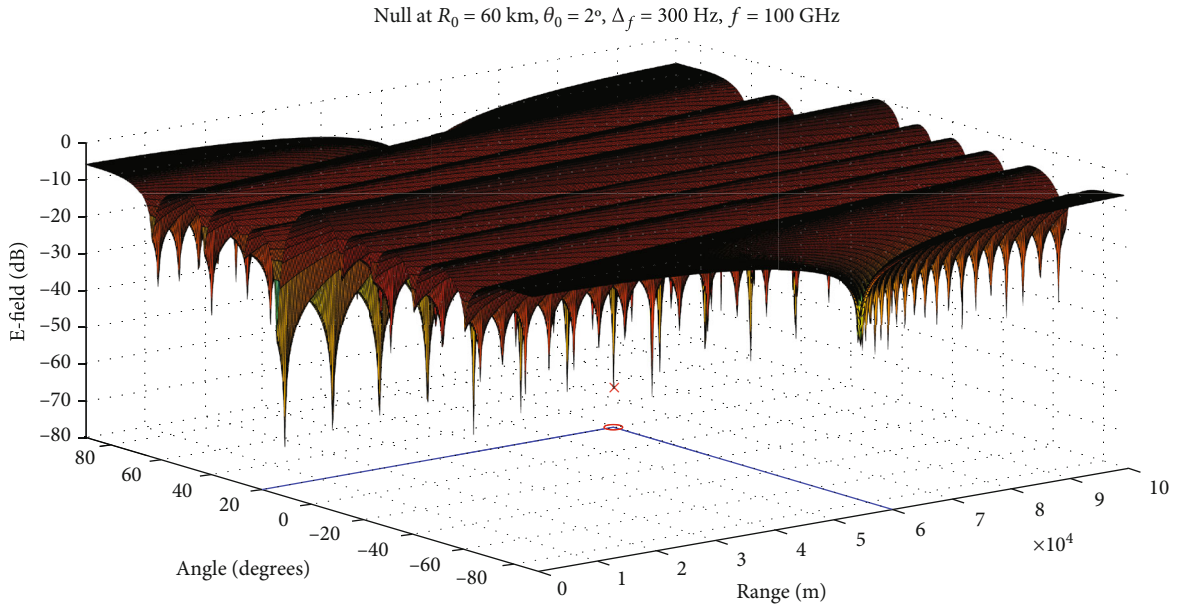


Figure 19: Actual target hides at distance 60 km and direction 20°.

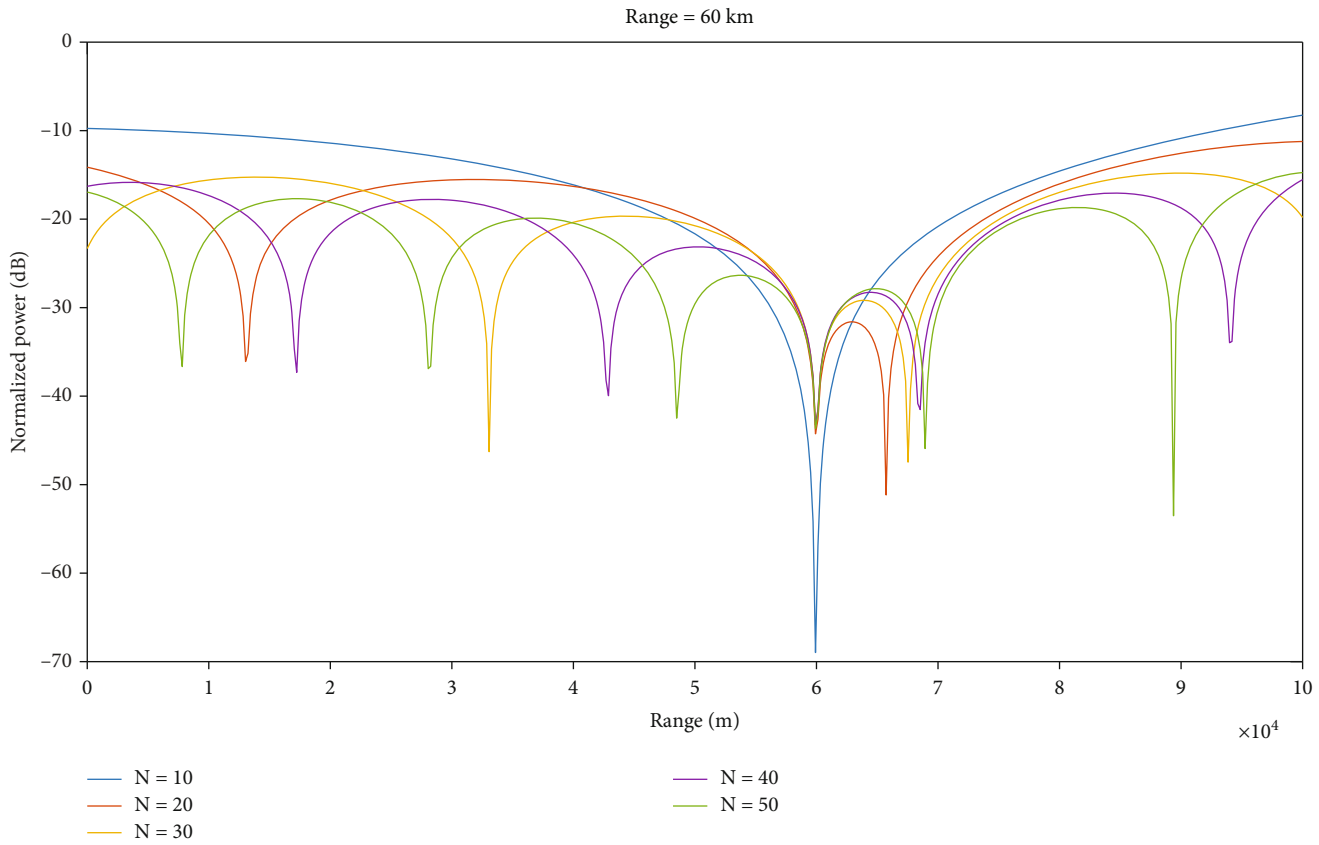


Figure 20: Null's placement at opponent radar's range with different numbers of antenna elements.

hide the aircraft from the vision of the opponent radar, and it is evident from its 3-D simulation in Figure 19. Further exploration of the method has been disclosed in Figures 20 and 21, whereby 2-dimensional graphs have been drawn to show the placement of the null at the desired range and direction, respectively. Vertical axis shows power in dB for different numbers of radiating antenna elements in the array. Lastly, to generate multiple fake targets at range 40 km, 50 km, 70 km, and 80 km along the same direction of the target aircraft, the proposed method calculates appropriate time

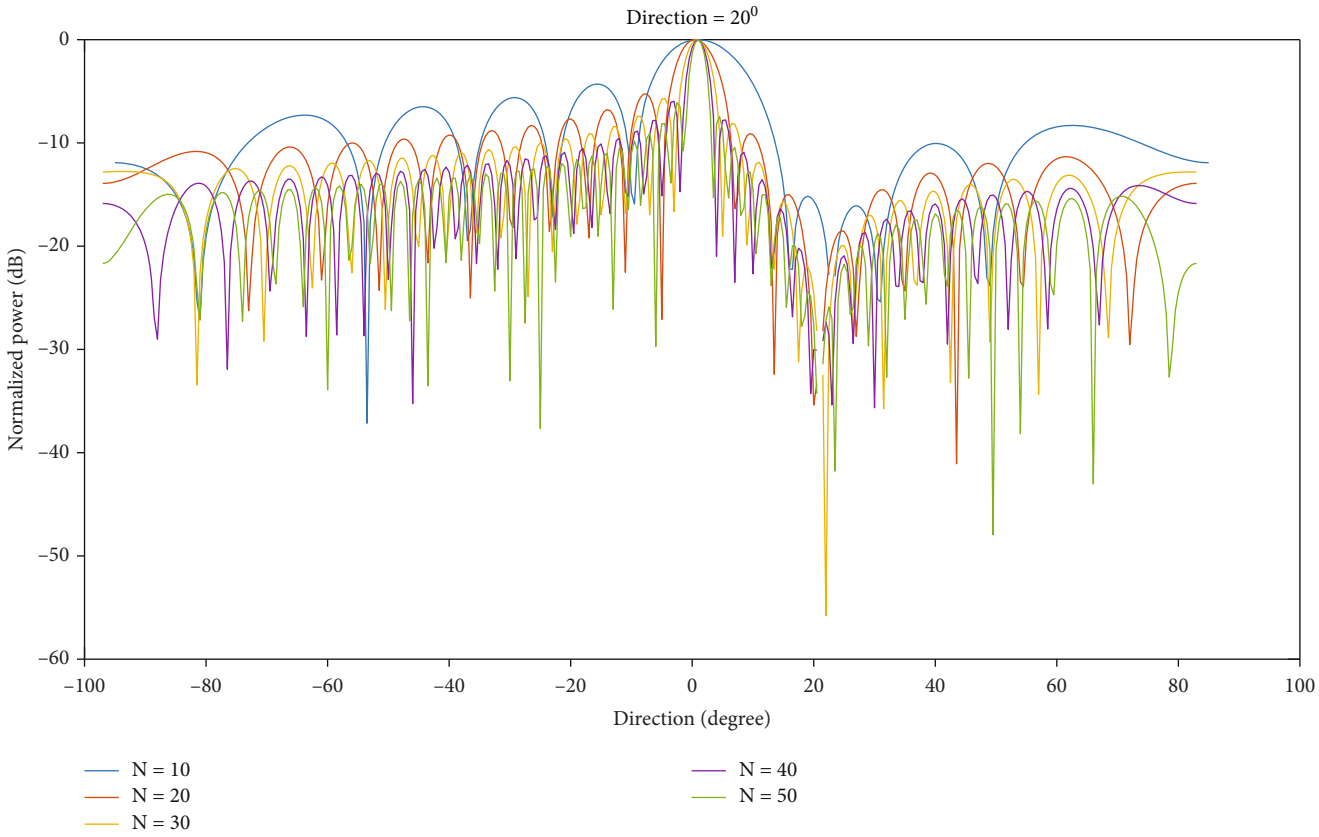


Figure 21: Null's placement at opponent radar's direction with different numbers of antenna elements.

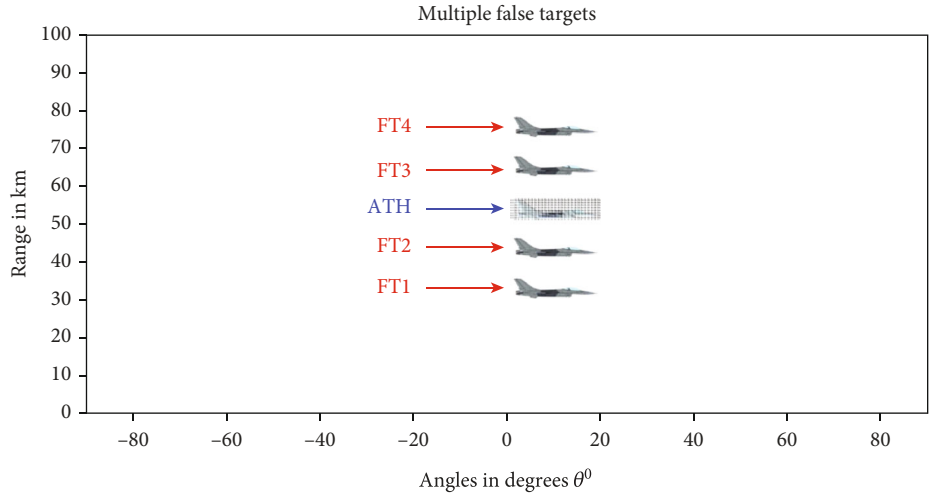


Figure 22: Actual target hides at 60 km, and false targets appear at {40, 50, 70, 80} km.

delays. The deceptive jammer sends echoes back with these time delays to confuse the enemy radar by showing him multiple false targets at desired ranges as shown in Figure 22.

10. Conclusion and Future Directions

A novel approach in the field of deception jamming has been developed in this research which hides actual target and displays multiple fake targets at different arbitrary ranges along the same direction to the enemy radar. This method has been developed to neutralize the effectiveness and dangers of the enemy radar which ultimately guarantees the safe penetration of the actual aircraft into the enemy territory. Moreover, a number of time modulations are performed for intercepted signal to display multiple deceptive jammer false targets at different ranges but along the same direction. It is assumed that the enemy radar has the capability of range angle-dependent radiation pattern characteristics. FDA radar's

radiation pattern is also time-dependent for larger pulse width. This effect has been covered by considering narrow radar pulse width. Hence, in this way, time dependency will not affect the radiation pattern of the FDA radar.

One of the other emerging area of research is how to counter deception jamming (anti-jamming) using FDA-MIMO radars in the field of ECCM techniques, because in the FDA-MIMO radar, we synthesize features of both radars (FDA and MIMO), and we achieve frequency diversity due to the FDA radar as well as waveform diversity due to the MIMO radar. We will try to investigate their relation with the presented method as well. We will also explore that how we can use FDA-MIMO radars to hide the actual aircraft target from the ground-based FDA opponent radar without the help of wave-scattering reflectors or advance escort-free-drone jammer. This research can give notion towards production of airborne deceptive jammers in the future.

Finally, the effectiveness and correctness of the proposed research have been verified by doing theoretical analysis and simulations by considering different cases with a number of distinct ranges of the actual target and fake deceptive targets. To put the research in the simplest form, the FDA radar is considered with the uniform linear array configuration. Moreover, the passage of time hardware computational and accuracy capabilities increases tremendously. Hence, limitations in implementing the algorithms due to hardware will be overcome; one can implement the proposed work through hardware with the collaboration of any national/international research organization.

Data Availability

All the data used in this work is based on simulations in MATLAB and is available for any research work.

Conflicts of Interest

All the authors declare that there is no conflict of interest.

Acknowledgments

Prof. Dr. Ijaz Mansoor Qureshi, a prolific Pakistani scholar, engineer, and scientist, suddenly died on the 10th of January 2021 at the age of 67. A renowned expert of signal array processing and evolutionary computing techniques, Professor Qureshi worked in different national universities of Pakistan for decades. He authored more than 200 publications and mentored more than 50 PhD scholars, including myself. Knowledge, wisdom, and experience in the field were aspects of his life; moreover, he was a man of great hospitality, friendship, and kindness. The authors found him warm, smiling, and engaging. Simplicity and morality were the salient features of his personality. His scholarship, penetrating mind, and truly and lovely personality will be long remembered in our minds and hearts.

References

[1] B. Zohuri, "Electronic countermeasure and electronic counter-countermeasure," in *Radar Energy Warfare and the Challenges of Stealth Technology*, pp. 111–145, Springer, Cham, 2020.

[2] Y. Yan-Juan, Z. Feng, A. Xiao-Feng, L. Xiao-Bin, and Z.-F. Xu, "A study on effectiveness modeling of multi-false-target jamming," in *2016 CIE International Conference on Radar (RADAR)*, pp. 1–5, Guangzhou, China, 2016.

[3] A. Farina and M. Skolnik, "Electronic counter-countermeasures," *Radar handbook 2*, McGraw-Hill Education, 2008.

[4] L. Neng-Jing and Z. Yi-Ting, "A survey of radar ECM and ECCM," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 3, pp. 1110–1120, 1995.

[5] A. Ahmed, W. Q. Wang, Z. Yuan, S. Mohamed, and T. Bin, "Subarray-based FDA radar to counteract deceptive ECM signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, 2016.

[6] S. Y. Nusenu, A. Basit, and E. Asare, "FDA transmit beamforming synthesis using Chebyshev window function technique to counteract deceptive electronic countermeasures signals," *Progress in Electromagnetics Research Letters*, vol. 90, pp. 53–60, 2020.

[7] F. H. Zhao, P. Zhang, and Y. S. Wang, "The study of barrage-type jamming for SAR," *Wireless Communication Technology*, vol. 16, no. 3, pp. 53–57, 2007.

[8] J. Schuerger and D. Garmatyuk, "Deception jamming modeling in radar sensor networks," in *MILCOM 2008 - 2008 IEEE Military Communications Conference*, pp. 1–7, San Diego, CA, USA, 2008.

[9] Y. Li, L. Gaohuan, and C. Huilian, "The study of multi-false targets deception against stepped-frequency waveform inverse synthetic aperture radar," in *2008 9th International Conference on Signal Processing*, pp. 2481–2484, Beijing, China, 2008.

[10] Y. J. Lee, J. R. Park, W. H. Shin, K. I. Lee, and H. C. Kang, "A study on jamming performance evaluation of noise and deception jammer against SAR satellite," in *2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, pp. 1–3, Seoul, Korea (South), 2011.

[11] B. Rao, G. Zhaoyu, and Y. Nie, "Deception approach to track-to-track radar fusion using noncoherent dual-source jamming," *IEEE Access*, vol. 8, pp. 50843–50858, 2020.

[12] S. ZHU, Q. Luo, and C. Tong, "The analysis of moving targets deceptive jamming model to SAR," *Electronic Information Warfare Technology*, vol. 1, p. 13, 2012.

[13] F. Zhou, B. Zhao, M. Tao, X. Bai, B. Chen, and G. Sun, "A large scene deceptive jamming method for space-borne SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 8, pp. 4486–4495, 2013.

[14] B. Zhao, F. Zhou, M. Tao, Z. Zhang, and B. Zheng, "Improved method for synthetic aperture radar scattered wave deception jamming," *IET Radar, Sonar & Navigation*, vol. 8, no. 8, pp. 971–976, 2014.

[15] Y. Liu, W. Wang, X. Pan, D. Dai, and D. Feng, "A frequency-domain three-stage algorithm for active deception jamming against synthetic aperture radar," *IET Radar, Sonar & Navigation*, vol. 8, no. 6, pp. 639–646, 2014.

[16] X. Yan, Y. Li, P. Li, and J. Wang, "Multiple time-delay smart deception jamming to pseudo-random code phase modulation fuze," in *2012 International Conference on Computer*

Distributed Control and Intelligent Environmental Monitoring, pp. 428–432, Zhangjiajie, China, 2012.

[17] P. Shi-rui, L. Yong-chun, L. Xin, and W.-f. Dong, "Study on target pose and deception jamming to ISAR," in *2009 2nd Asian-Pacific Conference on Synthetic Aperture Radar*, pp. 526–530, Xi'an, China, 2009.

[18] Z. Zong, L. Huang, H. Wang, L. Huang, and Z. Shu, "Micro-motion deception jamming on Sar using frequency diverse array," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2391–2394, Yokohama, Japan, 2019.

[19] Z. Bo, F. Zhou, X. Shi, Q. Wu, and Z. Bao, "Multiple targets deception jamming against ISAR using electromagnetic properties," *IEEE Sensors Journal*, vol. 15, no. 4, pp. 2031–2038, 2014.

[20] W. Wang, X.-Y. Pan, Y.-C. Liu, D.-J. Feng, and F. Qi-Xiang, "Sub-Nyquist sampling jamming against ISAR with compressive sensing," *IEEE Sensors Journal*, vol. 14, no. 9, pp. 3131–3136, 2014.

[21] X.-Y. Pan, W. Wang, and G.-Y. Wang, "Sub-Nyquist sampling jamming against ISAR with CS-based HRRP reconstruction," *IEEE Sensors Journal*, vol. 16, no. 6, pp. 1597–1602, 2015.

[22] L. X. D. Feng, Q. Liu, and X. Wang, "ISAR decoy generation by utilizing coherent multiplication modulated jamming," *Acta Electronica Sinica*, vol. 42, no. 12, pp. 2501–2508, 2014.

[23] D. Feng, L. Xu, X. Pan, and X. Wang, "Jamming wideband radar using interrupted-sampling repeater," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1341–1354, 2017.

[24] Q. Shi, C. Wang, J. Huang, and N. Yuan, "Multiple targets deception jamming against ISAR based on periodic 0- π phase modulation," *IEEE Access*, vol. 6, pp. 3539–3548, 2018.

[25] P. Antonik, M. C. Wicks, H. D. Griffiths, and C. J. Baker, "Frequency diverse array radars," in *2006 IEEE Conference on Radar*, p. 3, Verona, NY, USA, 2006.

[26] P. Antonik, M. C. Wicks, H. D. Griffiths, and C. J. Baker, "Multi-mission multi-mode waveform diversity," in *2006 IEEE Conference on Radar*, p. 3, Verona, NY, USA, 2006.

[27] P. Antonik and M. C. Wicks, "Method and apparatus for simultaneous synthetic aperture radar and moving target indication," US Patent 8,803,732, 2014.

[28] P. Antonik, M. C. Wicks, H. D. Griffiths, and C. J. Baker, "Range-dependent beamforming using element level waveform diversity," in *2006 International Waveform Diversity & Design Conference*, pp. 1–6, Lihue, HI, USA, 2006.

[29] M. Secmen, S. Demir, A. Hizal, and T. Eker, "Frequency diverse array antenna with periodic time modulated pattern in range and angle," in *2007 IEEE Radar Conference*, pp. 427–430, Waltham, MA, USA, 2007.

[30] P. F. Sammartino, C. J. Baker, and H. D. Griffiths, "Frequency diverse MIMO techniques for radarIEEE Transactions on Aerospace and Electronic Systems," vol. 49, no. 1, pp. 201–222, 2013.

[31] W.-Q. Wang, "Range-angle dependent transmit beampattern synthesis for linear frequency diverse arrays," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 8, pp. 4073–4081, 2013.

[32] W.-Q. Wang, "Frequency diverse array antenna: new opportunities," *IEEE Antennas and Propagation Magazine*, vol. 57, no. 2, pp. 145–152, 2015.

[33] W.-Q. Wang, "Overview of frequency diverse array in radar and navigation applications," *IET Radar, Sonar & Navigation*, vol. 10, no. 6, pp. 1001–1012, 2016.

[34] Y. Zhu, H. Wang, S. Zhang, Z. Zheng, and W. Wang, "Deceptive jamming on space-borne SAR using frequency diverse array," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 605–608, Valencia, Spain, 2018.

[35] B. Huang, W.-Q. Wang, S. Zhang, H. Wang, R. Gui, and L. Zheng, "A novel approach for spaceborne SAR scattered-wave deception jamming using frequency diverse array," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 9, pp. 1568–1572, 2020.

[36] W. Mao, H. Wang, S. Zhang, and X. Liu, "A novel deceptive jamming method via frequency diverse array," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2369–2372, Yokohama, Japan, 2019.

[37] H. Wang, S. Zhang, W.-Q. Wang, B. Huang, Z. Zheng, and L. Zheng, "Multi-scene deception jamming on SAR imaging with FDA antenna," *IEEE Access*, vol. 8, pp. 7058–7069, 2019.

[38] Y. Liao, T. Hu, X. Chen et al., "Antenna Beampattern with range null control using weighted frequency diverse array," *IEEE Access*, vol. 8, pp. 50107–50117, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 2396575, 1 page
https://doi.org/10.1155/2021/2396575

Corrigendum

Corrigendum to "IoT-Based Healthcare Support System for Alzheimer's Patients"

Rozita Jamili Oskouei,[1] **Zahra MousaviLou**,[2] **Zohreh Bakhtiari**,[1] **and Khuda Bux Jalbani**[3]

[1]*Department of Computer Science and Information Technology, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran*
[2]*Vali-e-Asr Hospital, School of Medicine, Zanjan University of Medical Science, Zanjan, Iran*
[3]*Riphah Institute of System Engineering, Riphah International University, Islamabad, Pakistan*

Correspondence should be addressed to Rozita Jamili Oskouei; rozita2010r@gmail.com

Received 27 February 2021; Accepted 27 February 2021; Published 16 March 2021

In the article titled "IoT-Based Healthcare Support System for Alzheimer's Patients" [1], author Zohreh Bakhtiari was affiliated to "Riphah Institue of System Engineering, Riphah International University, Islamabad, Pakistan" which is incorrect. The correct affiliation for this author is as follows:

"Department of Computer Science and Information Technology, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran"

The contact detail for Zohreh Bakhtiari was incorrect. The correct contact information is shown as follows:

Zohreh Bakhtiari: z_bakhtiari82@yahoo.com

The corrected list of affiliations and contact details are shown in the author information above.

References

[1] R. J. Oskouei, Z. MousaviLou, Z. Bakhtiari, and K. B. Jalbani, "IoT-Based Healthcare Support System for Alzheimer's Patients," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8822598, 15 pages, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8836808, 10 pages
https://doi.org/10.1155/2021/8836808

Research Article

On the Performance of Self-Concatenated Coding for Wireless Mobile Video Transmission Using DSTS-SP-Assisted Smart Antenna System

Nasru Minallah,[1] **Ishtiaque Ahmed**,[2] **Muhammad Ijaz**,[3] **Atif Sardar Khan**,[1] **Laiq Hasan**,[1] **and Atiqur Rehman**[3]

[1]*Department of Computer Systems Engineering, University of Engineering and Technology Peshawar, Peshawar 25000, Pakistan*
[2]*National Centre in Big Data and Cloud Computing, University of Engineering and Technology Peshawar (NCBC-UETP), Peshawar 25000, Pakistan*
[3]*Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar*

Correspondence should be addressed to Muhammad Ijaz; mijaz@hbku.edu.qa

Received 6 August 2020; Revised 6 November 2020; Accepted 19 December 2020; Published 15 January 2021

Academic Editor: Daehan Kwak

In the current age of advanced technologies, there is an escalating demand for reliable wireless systems, catering to the high data rates of mobile multimedia applications. This article presents a novel approach to the concept of Self-Concatenated Convolutional Coding (SECCC) with Sphere Packing (SP) modulation via Differential Space-Time Spreading- (DSTS-) based smart antennas. The two transmitters provide transmit diversity which is capable of recuperating the signal from the effects of fading, even with a single receiving antenna. The proposed DSTS-SP SECCC scheme is probed for the Rayleigh fading channel. The SECCC structure is developed using the Recursive Systematic Convolutional (RSC) code with the aid of an interleaver. Interleaving generates randomness in exchange for extrinsic information between the constituent decoders. Iterative decoding is invoked at the receiving side to enhance the output performance by attaining fruitful convergence. The convergence behaviour of the proposed system is investigated using EXtrinsic Information Transfer (EXIT) curves. The performance of the proposed system is ascertained with the H.264 standard video codec. The perceived video quality of DSTS-SP SECCC is found to be significantly better than that of the DSTS-SP RSC. To be more precise, the proposed DSTS-SP SECCC system exhibits an E_b/N_0 gain of 8 dB at the PSNR degradation point of 1 dB, relative to the equivalent rate DSTS-SP RSC. Similarly, an E_b/N_0 gain of 10 dB exists for the DSTS-SP SECCC system at 1 dB degradation point when compared with the SECCC scheme dispensing with the DSTS-SP approach.

1. Introduction

The recent developments in wireless technologies have resulted in the expansion of higher data rates of cellular systems with diverse applications, severely limiting the available bandwidth [1]. The existing wireless communication systems provide a backbone for the hugely utilized internet in the present era. It is estimated that the evolution of next-generation applications and the advancements in the Internet of Things (IoTs) will remarkably add to the increasing data capacity needs by 30-40% per year [2]. The fifth-generation (5G) wireless technology has very specific aims of further increasing the data rate and catering for the ascents in wireless services, by efficient utilization of the available bandwidth [3].

Ever since the pioneering work of Shannon in 1948 [4], researchers started investing efforts to design fast, efficient,

and high-quality transceivers attaining high bit-rate communication with the least Bit-Error Rate (BER). Standing on the shoulders of giants, researchers and scientists paved the way for wireless and digital communication. In a typical wireless communication model, transmission of voice, image, video, or any other type of multimedia content is assisted by the source and channel coding techniques. In order to successfully transmit information, a reliable and an efficient communication system is needed. The information from any source is forwarded to the transmitter for its operation and conversion into the transmission signal. The signal is then disseminated with the aid of a transmitting antenna, such that it propagates through the channel. It should be noted that the effects of the channel are mostly not beneficial for the signal. The channel offers several impingements, such as addition of noise, interference, and fading, leading to error formation and reduction in the system's performance. These effects are mitigated at the receiver side, resulting in reliable data transfer to the destination for end users.

In the recent proliferation of multimedia services, source coding has become an important topic for researchers to delve into. The main purpose of source coding is to compress the original data meant to be transmitted via wireless technology. Multimedia data are extensively compressed with the help of source coding techniques. Hence, several video coding techniques were put forward by scientists, for encoding different types of multimedia content. Most of the video coding techniques deploy the hybrid coding mechanism [5]. Hybrid coding makes use of the transform coding with motion-compensated prediction. In this treatise, we will be considering the Advanced Video Coding (AVC) standard, also referred to as H.264 video codec, as our source coding technique. H.264 is considered to be the best standard for achieving the requirements of fast next-generation efficient and ubiquitous communication [6]. The reliable transmission of multimedia contents becomes a very demanding task when dealing with high-compression and efficient multimedia standards. Due to the employment of hybrid and compression efficient techniques, error propagation occurs in the standard video stream [7]. Therefore, there must be a way forward to tackle such errors. One popular channel coding technique, known as Forward Error Correction (FEC), is very much beneficial to protect the compressed stream from errors. Hence, it is plausible that for reliable transmission of multimedia contents over nonideal channels, it is necessary to perform source coding and channel coding [7–9]. With this, the terminology of Joint Source-Channel Decoding (JSCD) is coined, as explained for H.264 in [10]. The reason for the large use of the H.264 standard lies in the fact that some of its intriguing features are so favourable that they are retained in the modern standards as well [11].

It is very much clear from the discussion above that channel coding is an integral part of the wireless communication paradigm. The inimical effects of the addition of noise, signal distortion, interferences, and fading render wireless communication more prone to unreliability than wireline communication [12]. To overcome the errors and hence to decrease the BER, the idea of redundancy was introduced [4]. The operation of redundancy incorporates additional (parity) bits before the data is transmitted, making it possible to detect and tackle any arising error and mismatch at the receiver. This recovery process greatly improves the performance of a wireless system by reducing the BER [13]. The overall process of incorporating redundancy is termed as Shannon Coding, which is elucidated in [4]. Different channel coding techniques have been presented and implemented in the recent years. Some of them include optimal code design for enhanced performance and security [14] and optical orthogonal coding for cable communication [15] and for the wireless scenario including Hamming codes [16–18], Polar codes [19], and Bose-Chaudhuri-Hocquenghem (BCH) and convolutional codes [20]. There is a capacity associated with each type of channel. The maximum throughput at which any channel can reliably and correctly transmit information to the receiver is called channel capacity. The popular Shannon capacity is attributed to the great work presented by the father of information theory, Claude Shannon, in [4]. Many researchers persistently continued to hone their findings, resulting in the feasible designs of systems approaching Shannon's capacity limit [21–23].

Forney presented the concept of concatenated coding [24], though not given much attention by researchers in the early days. Soon after the proposition of turbo codes, based on the concatenated convolutional philosophy [25], scientists avidly started to peruse the concepts of concatenated coding. In [26], the three main categories of concatenated coding are discussed. These include the Parallel Concatenated Convolutional (PCC), Serial Concatenated Convolutional (SCC), and Self-Concatenated Convolutional (SeCC) codes. The constituent encoders in a PCC coding are linked in a parallel fashion, mutually sharing information via an interleaver (Π). In SCC coding, the N number of component encoders is serially interconnected with the help of $N-1$ interleavers. Finally, the only encoder of SeCC coding requires the functionality of an interleaver to convert the input data to an interleaved version and simultaneously feeding to the encoder. For the convolutional codes, it is worthy to mention that they are vastly adopted in modern wireless standards and satellite communications. These intuitive codes are covered in detail in [27], and several decoding approaches for such codes are mathematically discussed in [28, 29].

Moving to the advanced topic of Differential Space-Time Spreading (DSTS), we briefly discuss how they evolved. Space-Time Block Codes (STBCs) constitute an important family of Multiple-Input Multiple-Output (MIMO) systems. STBC offers a simpler approach to encoding and spreading with reasonably good performance [30, 31]. Inspired by the concept of STBC, authors in [32] proposed Space-Time Spreading (STS). The technique of coherent detection was common in all of the STBCs and STS systems. For coherent detection to be possible, there is a stringent requirement of accurate and complete channel knowledge at the receiving side. This Channel State Information (CSI) renders the system more complex and expensive. In quest of efficient systems, yet mitigating the complex requirements of CSI, Tarokh et al. proposed Differential Space-Time Block Coding (DSTBC) using two transmitters, later demonstrated for a larger number of transmitters as well [33, 34]. The aim of

low-complexity system design was achieved, although the only snag was a little performance loss. The DSTS technique can be integrated with several modulation schemes like Phase-Shift Keying (PSK), Quadrature Amplitude Modulation (QAM), and Sphere Packing (SP), requiring no channel estimation [35]. SP modulation is widely becoming popular in the construction of error correction codes. Su et al. introduced the merger of transmit diversity techniques with SP modulation, evincing that the SP-aided STBC surpassed in performance the conventional STBC counterpart [35]. Minimum Euclidean distance was deemed to be an appropriate metric for evaluating the attainable gain of orthogonal transmit diversity schemes [35]. SP modulation assures the best possible minimum Euclidean distance between the modulated symbols and enhances the error resilience property of the system. Regarding Euclidean distance, it is simply the length of a straight line between any two points in the Euclidean space. The DSTS-SP approach was adopted by several authors to attain prolific performance in turbo detection [35], cooperative communication [36], adaptive multirate wideband speech coding [37], and iteratively decoded irregular variable length coding [38]. Until now, the literature has been silent on the incorporation of DSTS-SP in SeCC codes. As SeCC coding offers significant performance with little complexity, hence, further exploration needs to be carried out in this regard.

Keeping in view the above background, we aim to introduce a novel methodology of SElf-Concatenated Convolutional Coding (SECCC), with iteratively detected SP modulation-assisted DSTS-based smart antennas. This article is somehow an extension of the work presented in [8]. We build upon the system proposed in [8] such that the beneficial feature of DSTS-SP is incorporated for an enhanced performance. The iteratively decoded DSTS-SP SECCC proposition will be analyzed via the EXtrinsic Information Transfer (EXIT) chart curves. The rationale of the research work presented in this article is summarized as follows:

(i) Introduction to the novel concept of DSTS-SP SECCC and its performance comparison with other schemes

(ii) Utilization of the double antennas for the sake of attaining a rich transmit diversity gain, supporting profitable performance using a simple receiving antenna requiring no CSI

(iii) Provision to the understanding of video performance of the proposed system using the H.264/AVC standard

The rest of the paper is organized as follows. The proposed DSTS-SP SECCC system is made plain in Section 2. Further description of the proposed system and parameter settings follow in Section 3. Section 4 presents the EXIT chart analysis and highlights some of the linked terminologies used in this expedition. Section 5 articulately covers the simulation results. Finally, Section 6 succinctly concludes the paper with future research description.

2. DSTS-SP SECCC System Overview

The block diagram of the proposed system is depicted in Figure 1. It can be seen that the information bits are initially passed through the block of source coding, employing H.264/AVC. The technique of slice structuring is adopted, resulting in the partitioning of each video frame into independently coded multiple slices. Furthermore, we have also subsumed the approach of Data Partitioning (DP). DP helps in generating different streams of the source coded video per slice, on the basis of important coding elements and parameters. We arrange each type of stream with several occurrences in each frame and concatenate the three resultant partitions, represented as A, B, and C into a single stream x_i. The overall source coding operation compresses the data and might put a contrary effect on the reliability of the original data. For this, channel coding block serves the purpose. The SECCC channel encoder is used in our designed system. Modulation is done with the aid of SP to shift the spectrum of channel coded signal to a form suitable for wireless transmission over the Rayleigh channel. Transmission of a signal over wireless or radio channel is often associated with reflection, diffraction, and scattering experiences. One of the effects produced as a result of these experiences is the multipath phenomenon [39]. Due to the multipath effect, the transmitted signal splits into multiple versions based on the power and fading distributions. Therefore, it becomes crucial to precisely predict the channel model for wireless systems [39]. The Rayleigh fading channel is deemed to be a useful propagation channel in the scenario of a multipath environment for wireless systems [40]. The Rayleigh probability density function is given by

$$p(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0. \tag{1}$$

Here, $\sigma^2 = E[r^2]$ is the variance of the circularly symmetric complex random variable r, having real and imaginary parts. The term $E[*]$ denotes statistical averaging. The SP modulated signal is passed through the DSTS block for incorporating the beneficial feature of transmit diversity gain and transmitted via two transmit antennas (Txs). At the receiving side, we consider the use of one receive antenna (Rx1) for the sake of simplicity and without any loss of generality. The received signal is passed through the DSTS decoder, and then, it is demodulated back to its original form. It is notable that the DSTS decoder is a suboptimum one, operating with a simpler way of accepting successively received interdependent signals. The potential discrepancy in the receiving data is detected and overcome by the SECCC decoder. The schematics of the SECCC decoder are explained below. The resultant $L(y_i)$ signal is processed by the SECCC decoder to yield the overall reconstructed signal, which is deconcatenated and then MUltipleXed (MUX). Eventually, the source decoder outputs the bits to end users.

The architectural design of the DSTS encoder is illustrated in Figure 2. Its major subcomponent blocks are the differential encoder and STS encoder. The modulated signal is differentially encoded till the refined output q is ready to be

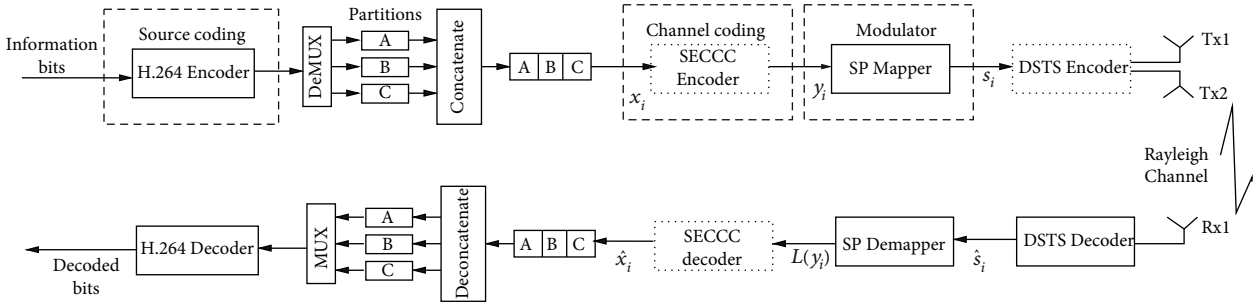


Figure 1: Block diagram of the DSTS-SP SECCC system.

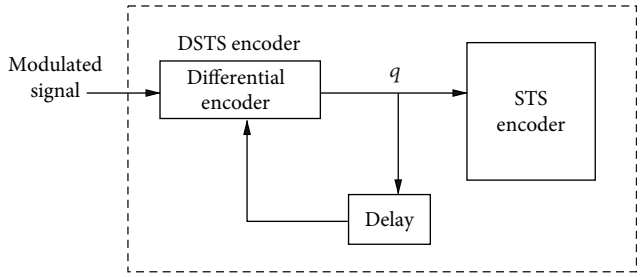


Figure 2: The DSTS encoder structure.

provided to the STS encoder. There is a delay component between the differential and STS encoders. This delay is usually provided by the interleaver. The reason for introducing this delay component is to make the output q from the differential encoder highly uncorrelated and differential. The simultaneous feedback to the differential encoder is invoked till the set value of delay, attaining enhanced q. The STS encoder spreads the data via the technique of Walsh codes. Walsh coding renders the overall process of encoding and resultant code longer, providing lower throughput per antenna. The differentially spread data is divided into two substreams, each transmitted via a separate antenna. These antennas with certain transmit power values assist the final signal to propel over the channel.

Moving to the details about SECCC, the various stages involved are highlighted in Figure 3. SECCC is somehow similar to PCC coding in the sense that constituent encoders of PCC are replaced with a single code, involving an even-odd number of interleavers, as specified in [41]. The intriguing feature of SECCC is the simplicity of its structure. SECCC systems essentially contain a single encoder and decoder as shown in Figure 3. The SECCC encoder maneuvers by accepting the source coded bits and simultaneously converts them to interleaved bits. The interleaver makes the bits profusely uncorrelated. The Parallel-to-Serial (P/S) converter receives both the direct source coded stream x_i and its interleaved version x_i'. The serial output is fed to rate R_1 RSC encoder via the Generator Polynomial (GP) of $(G_0, G_1, G_2 = 13,15,17)_8$, represented in octal format, where the first term G_0 represents the feedback polynomial [42]. Generally, the number of bits get increased due to the RSC encoding. After the RSC encoding stage, there is an interleaver to randomize the encoded bits. The next stage is of puncturing at rate R_2. In order to maximize the bandwidth efficiency, puncturer obliterates some bits from transmission. For instance, a rate a/b puncturer will stop $b - a$ bits from transmission. Resultantly, the number of bits gets reduced after the puncturer. It is obvious that by a mere alteration in the values of R_1 and R_2, various rates of SECCC could be invoked. For the present case, we will be using the values of R_1 and R_2 to be 1/2 each. Finally, the overall rate R of the SECCC encoder is as given in [43] and computes to 1/2 for the specified values of R_1 and R_2:

$$R = \frac{R_1}{2 * R_2}. \tag{2}$$

The SECCC decoder consists of a single SISO Maximum *A Posteriori* (MAP) decoder. The MAP decoder is hypothetically divided into two component decoders, for better understanding of the exchange of extrinsic information. The output from the SP demapper is fed to the depuncturer, inserting zeros (if required) in the places of bits which were punctured. The two component decoders mutually share soft extrinsic information until the specified number of iterations. With the help of interleavers and deinteleavers, the output knowledge of the one component decoder is presented as *a priori* input to the other component decoder. The two hypothetical component decoders iterate until there is no further improvement attainable after feedback. Upon reaching this point, the maximum iterative performance gain is achieved, and it is known as the point of convergence. In view of the two component decoders, the single SECCC decoder can be viewed as a PCC decoder. This way, the iterations are also self-iterations of the MAP decoder. Lastly, the output bits from the SECCC decoder are Serial-to-Parallel (S/P) converted and forwarded to the source decoder accordingly.

3. Proposed System Description

The proposed design is simulated using the IT++ signal processing and communication library. The results were generated using the parametric setting as stated in Table 1. The constituent inner and outer rates of the error protection schemes are stated in Table 2. It should be noted that in the precursor paper [8], we utilized the RSC-coded

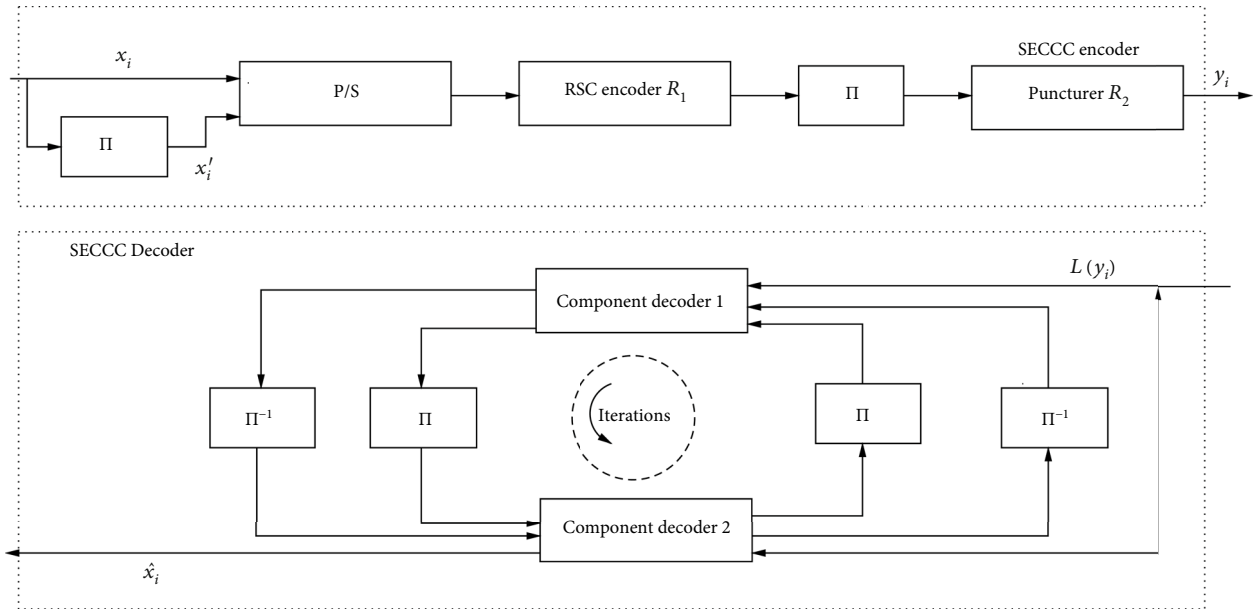


Figure 3: Block diagram of SECCC using iterative decoding.

Table 1: DSTS-SP SECCC system parameters.

Parameters	Value	Parameters	Value
Source code	H.264/AVC	Channel	Rayleigh fading
Source bit-rate	64 kbps	Tx antennas	2
Video sequence	QCIF *Akiyo*	Rx antennas	1
Frame rate	15 fps	RSC generator	$(G_0, G_1, G_2 = 13{,}15{,}17)_8$
Slices per frame	9	Interleaving bits	10000
Number of MBs per slice	11	Normalised Doppler frequency	0.01
Intraframe MB update	3	Modulation scheme	SP
MIMO scheme	DSTS	Spreading code	Walsh code

Table 2: Inner and outer rates of the error protection schemes.

Error protection scheme	Rate: Outer code	Rate: Inner code	Rate: Overall
DSTS-SP SECCC	$R_1 = 1/2$ RSC	$R_2 = 1/2$ puncturer	$R = \frac{1}{2}$
DSTS-SP RSC	RSC = $\frac{1}{4}$	Puncturer = $\frac{1}{2}$	$R = \frac{1}{2}$
SECCC scheme of [8]	$R_1 = 1/2$ RSC	$R_2 = 1/2$ puncturer	$R = \frac{1}{2}$

benchmarker. But for the work presented here, we will be employing the DSTS-SP RSC and SECCC scheme dispensing with the DSTS-SP for comparison. The SECCC scheme of [8] employs QPSK modulation whereas the other schemes of Table 2 use the DSTS-SP transmission mechanism. The Quarter Common Intermediate Format (QCIF) *Akiyo* sequence in (176 × 144) pixels or 99 Macroblocks (MBs) each of size (16 × 16)-pixel resolution using the H.264 encoder at 64 kbps and 15 frames-per-second (fps) is deployed. The 99 MBs per frame improve the efficiency of the iterative decoding and reinforce the use of longer interleavers, without causing any unnecessary delay in the system. To counterbalance the effect of error propagation, the approach of intraframe coded MB updates and predicted "*P*" frames is employed. Hence, there are 3 intraframe MBs per QCIF frame and 44 "*P*" frames after each intra "*I*" frame, with a time lag of 3 seconds between two successive "*I*" frames. Furthermore, the complexity of the source encoder is kept realistic by avoiding the bidirectionally predicted frames and turning off the robust Flexible MB Ordering (FMO). FMO offers a small advantage in low-motion sequences, although the computational complexity increases by many folds [44].

With the intention of boosting confidence in our results, the 45-frame video sequence test is repeated 250 times with 26 system iterations, and the average value is used in the results. The technique of Walsh coding is utilized as a spreading code. Walsh codes are specifically employed to enhance BER performance with relatively lower computational

complexity [45]. The Signal-to-Noise Ratio (SNR) of the system improves greatly as the length of the Walsh code is increased, but the flip side is a reduction in data rates [46]. For higher rates, the transmitter is allocated with a large number of codes [45]. Walsh functions form the basis of such codes. Each bit is spread by a separate Walsh function, and all of the functions are fully uncorrelated. Hence, all of the Walsh functions and codes are orthogonal and are generated using the Hadamard matrix, a square matrix containing one row of all zeros and remaining with an equal number of zeros and ones [47]. Such codes are flexible enough to be concatenated with other codes, like for synchronized multiuser systems because of its orthogonal features [46, 48].

4. Linked Terminologies and EXIT Chart Analysis

The iterative decoding schemes are mainly deployed to achieve an enhanced BER performance. EXIT charts are useful to predict the convergence pattern of an iteratively decoded system [49, 50]. Proposed by ten Brink, EXIT charts are based on the exchange of mutual information between the constituent Soft-Input Soft-Output (SISO) decoders [42]. EXIT analysis is convenient in the sense that it expeditiously predicts the SNR value where an infinitesimal BER occurs, without performing the tiresome bit-by-bit decoding [50, 51]. The EXIT chart relies on two major assumptions for accuracy; firstly, the *a priori* logarithmic-likelihood ratio (LLR) information should be uncorrelated, and secondly, its probability density function (PDF) must be Gaussian distributed. These two requirements are generally fulfilled by employing higher interleaver lengths [52]. The following basic relations govern the EXIT curves for a generalized wireless system, as discussed in [50, 53]:

$$0 \le I_A \le 1, \tag{3}$$

$$0 \le I_E \le 1, \tag{4}$$

$$I_E = T\left(I_A, \frac{E_b}{N_o}\right), \tag{5}$$

$$T(0) \le I_E \le T(1). \tag{6}$$

In Equations (3) and (4), I_A refers to the *a priori* information, whereas I_E denotes extrinsic information. I_A is the intrinsic information about the bit, known even before the decoding process begins. I_E is generated when we subtract I_A from the first output of the constituent decoder. Interleaving renders I_E to serve as I_A for the other decoders. The *a posteriori* information is the output of any decoder accepting the channel's input and I_A [53]. The symbol T in Equation (5) represents the transfer function, converting I_A to I_E at the specified E_b/N_0 value. The inverse of this transfer function exists on the range specified in Equation (6). It is notable that different values of E_b/N_0 result in distinct EXIT curves. The greater value (maximum upto 1) of I_A signifies that the greater number of bits becomes known, due to which the value of I_E also increases. As a general rule, the closer the value of I_A/I_E to 1, the better the decoding [53].

Interleavers are particularly useful for introducing time diversity and delay in any communication system. It renders the data highly uncorrelated such that it can be constructively exploited in the EXIT visualization [53]. It was demonstrated in [54] that a system with a higher number of interleaving bits will significantly reduce the number of iterations required to reach the point of convergence. This is because of the fact that the higher value of interleaving bits renders the distribution more close to Gaussian, resulting in a performance close to Shannon's limit. The logarithmic-likelihood ratios (LLRs) or L values symbolize the logarithm of the ratio of probability of any bit. The concept of L values is adopted in many systems involving iterative decoding and was initially studied by Robertson [54], as presented in

$$L(d_k) \coloneqq \ln\left(\frac{P(u_k = +1)}{P(u_k = -1)}\right), \tag{7}$$

where $L(d_k)$ represents the L value of bit d_k and $P(u_k)$ is the probability of bit u_k for its two legitimate values of +1 and -1.

For a specific E_b/N_0 value, an infinitesimally lower BER is achieved only if the EXIT curves meet the (1,1) point of perfect convergence [53]. If the curves converge at the point of perfect convergence, there must be an open tunnel in between the (0,0) and (1,1) points, commonly known as the *convergence tunnel*. The EXIT curves for the advocated system of DSTS-SP SECCC at several E_b/N_0 values are depicted in Figure 4. The two EXIT curves corresponding to the two hypothetical decoders are essentially a mirror of each other along the 45-degree diagonal line. This corroborates our claim that the SECCC decoder consists of a single SISO MAP decoder. Monte-Carlo simulation-based stair-shaped trajectories at various E_b/N_0 values are also shown in Figure 4, iterating between the EXIT curves. We can visualize the *convergence tunnels* and hence the convergence patterns for different values of E_b/N_0. The EXIT curves of the DSTS-SP SECCC do not succeed to reach the point of perfect convergence at the E_b/N_0 value of -2 dB, as the curves intersect prior to the (1,1) point, providing no open tunnel. However, the curves provide the required tunnel for convergence at E_b/N_0 of -1 dB or greater values as seen in Figure 4. The corresponding trajectories confirm this claim by iterating between the component EXIT curves to attain the highest value of extrinsic information. Furthermore, it is worth mentioning that the DSTS-SP SECCC offers a fruitful approach by converging at considerably lower E_b/N_0 than the solely SECCC scheme of [8].

5. Simulation Results

Figure 5 depicts the BER performance of the proposed system in comparison with the identical rate DSTS-SP RSC and SECCC scheme as in Table 2. As expected, the

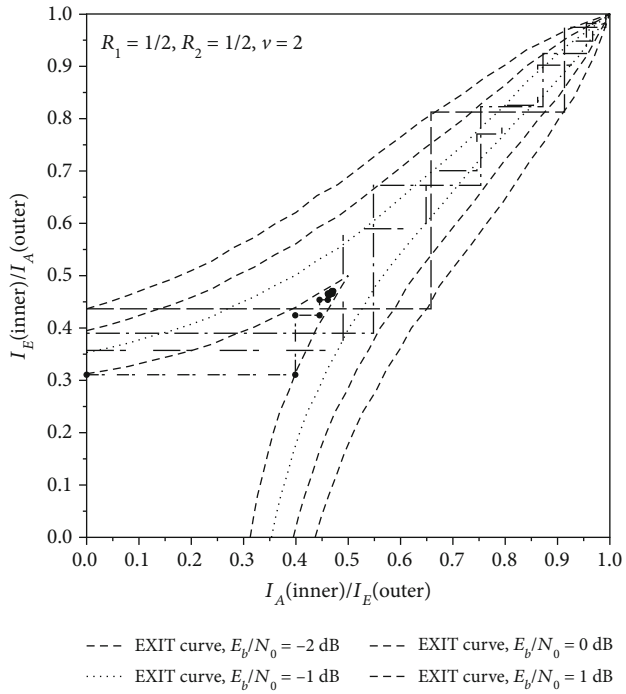


Figure 4: EXIT curves and corresponding decoding trajectories at E_b/N_0 values of (-2, -1, 0, 1) dB for the DSTS-SP SECCC system with R_1 and R_2 of 1/2 each.

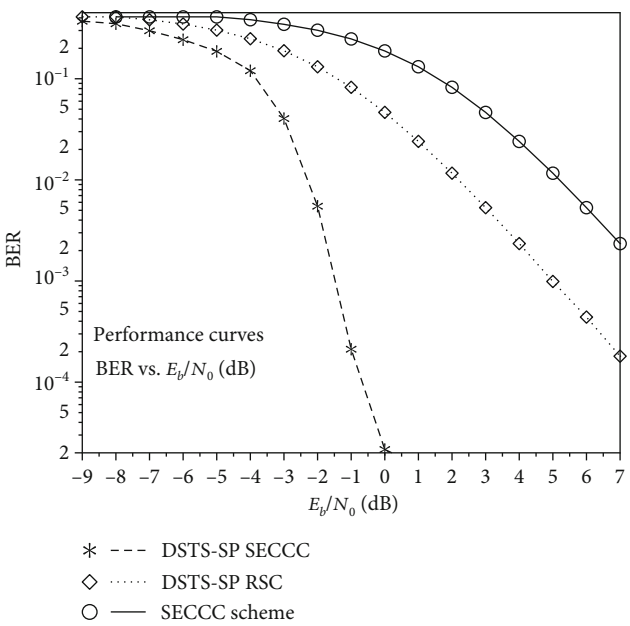


Figure 5: BER performance comparison of the schemes specified in Table 2.

incorporation of DSTS-SP results in significantly lower BER, achieving higher bandwidth efficiency. For the proposed DSTS-SP SECCC system, the 10^{-4} value of BER is achieved at a lower E_b/N_0 value in comparison with the other similar rate schemes. More precisely, there is an E_b/N_0 gain of 8 dB

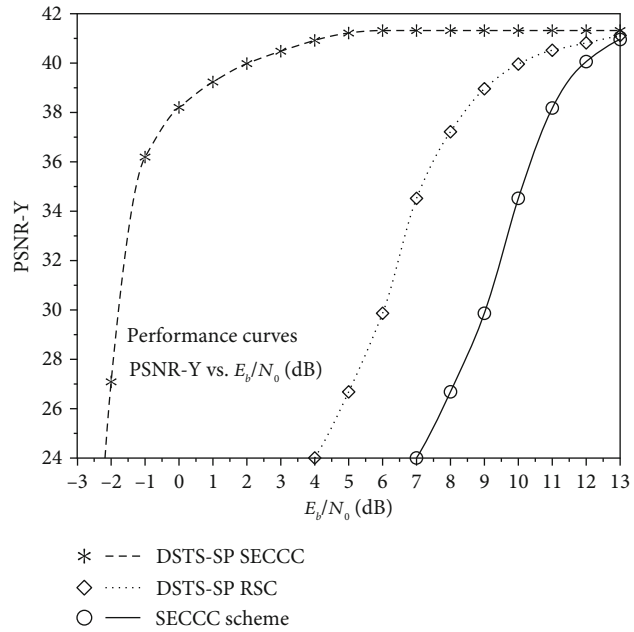


Figure 6: PSNR performance comparison of the schemes specified in Table 2.

at the BER degradation point of 10^{-4} when considering the DSTS-SP SECCC system relative to the DSTS-SP RSC. Moreover, the proposed DSTS-SP SECCC outperforms the SECCC scheme dispensing with the DSTS-SP by 11 dB at the BER degradation point of 10^{-4}. Thus, the BER performance metric advocates the proposed system.

Figure 6 provides the Peak Signal-to-Noise Ratio (PSNR) performance of the error protection schemes given in Table 2. PSNR, an objective video quality metric, provides an insight into the strength of the signal by computing the ratio between the original signal strength and noise of the related channel. An accurate estimate of the perceptual video or image quality is provided with the PSNR when the content, codec, and transmission system remain fixed [55]. It becomes obvious from Figure 6 that the PSNR performance of the proposed DSTS-SP SECCC system is better across the entire region of E_b/N_0. Explicitly, an E_b/N_0 gain of 8 dB and 10 dB is recorded at the PSNR degradation point of 1 dB, considering the proposed system over the identical rate DSTS-SP RSC and SECCC scheme, respectively.

Finally, the subjective video quality performance indicator of the proposed system is compared with the DSTS-SP RSC and can be visualized via Figure 7. The frames of Figure 7 were averaged 30 times prior to its presented form after the transmission of both luminance and chrominance parts of *Akiyo* sequence. It is plainly visible that the incorporation of DSTS-SP to SECCC has a substantial impact on the perceived video quality with reference to the DSTS-SP RSC. At 4 dB, there is a major difference in the performance of both schemes as the perceived quality of the advocated system is far better than that of the DSTS-SP RSC (annoying perceptual video distortions).

FIGURE 7: Subjective video quality performance of the *Akiyo* sequence frame using the proposed system (a) and DSTS-SP RSC scheme (b) summarized in Table 2 at E_b/N_0 values of (1, 2, 3, 4) dB (left to right).

6. Conclusion

This article presents the SECCC iterative channel decoding system for the H.264 video standard over the Rayleigh channel using DSTS-SP-based smart antennas. The presented setup is designed with a motive to operate the system close to channel capacity for wireless video communication. The DSTS-SP approach profusely improves the performance of SECCC in terms of BER and PSNR. The presented system productively exploits the diversity gain resulting from the two transmit antennas, without imploring any computational complexity at the single receiving antenna. Likewise, EXIT curves confirm the usefulness of the advocated system, as it shows fruitful convergence behaviour for deployment in efficient and flexible transceivers. Furthermore, the subjective video quality of the proposed system is found to be significantly better than that of the DSTS-SP RSC. To be more precise, the developed system exhibits an E_b/N_0 gain of 8 dB over the identical rate DSTS-SP RSC at the PSNR degradation point of 1 dB. Similarly, there exists an E_b/N_0 gain of 10 dB for the proposed system at 1 dB degradation point when compared with the SECCC scheme avoiding DSTS-SP. We aim to extend the approach of DSTS to be generally invoked with other modulation techniques and supported for transmission over dispersive channels. Finally, another promising topic of research would be to propose efficient interleaver designs helpful in overcoming or mitigating the detrimental effects of correlation due to channel and differential encoding.

Data Availability

The authors approve that data used to support the finding of this study are included in the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The financial support of the National Centre in Big Data and Cloud Computing, University of Engineering and Technology, Peshawar (NCBC-UETP), under the auspices of the Higher Education Commission, Pakistan, is gratefully acknowledged. The authors would like to thank and acknowledge Qatar National Library, Qatar, for funding the publication charges of this article.

References

[1] F. Cogen, E. Aydin, N. Kabaoglu, E. Basar, and H. Ilhan, "Generalized code index modulation and spatial modulation for high rate and energy-efficient MIMO systems on Rayleigh block-Fading channel," *IEEE Systems Journal*, pp. 1–8, 2020.

[2] Cisco Corphttps://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html.

[3] A. Banerjee, K. Vaesen, A. Visweswaran et al., "Millimeter-wave transceivers for wireless communication, radar, and sensing," in *2019 IEEE Custom Integrated Circuits Conference (CICC)*, Austin, TX, USA, USA, 14-17 April 2019.

[4] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[5] Yuan Zhang, Wen Gao, Yan Lu, Qingming Huang, and Debin Zhao, "Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks," *IEEE Transactions on Multimedia*, vol. 9, no. 3, pp. 445–454, 2007.

[6] H. Kalva, "The H. 264 video coding standard," *IEEE multimedia*, vol. 13, no. 4, pp. 86–90, 2006.

[7] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, 2000.

[8] M. F. Nasruminallah, U. Butt, S. X. Ng, and L. Hanzo, "H.264 wireless video telephony using iteratively-detected binary self-concatenated coding," in *2010 IEEE 71st Vehicular Technology Conference*, pp. 1–5, Taipei, Taiwan, 2010.

[9] X. Gao, L. Zhuo, S. Wang, and L. Shen, "A H. 264 based joint source channel coding scheme over wireless channels," in *2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 683–686, Harbin, China, 2008.

[10] Nasruminallah and L. Hanzo, "EXIT-chart optimized short block codes for iterative joint source and channel decoding in H.264 video telephony," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4306–4315, 2009.

[11] G. J. Sullivan, J. M. Boyce, Y. Chen, J.-R. Ohm, A. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)," *IEEE Journal of selected topics in Signal Processing*, vol. 7, no. 6, pp. 1001–1016, 2013.

[12] N. Nasaruddin, B. Yuhanda, E. Elizar, and S. Syahrial, "Design and performance analysis of channel coding scheme based on multiplication by alphabet-9," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 9, no. 1, pp. 7–13, 2017.

[13] J. C. Moreira and P. G. Farrell, *Essentials of Error-Control Coding*, John Wiley & Sons, 2006.

[14] M. Franklin, R. Gelles, R. Ostrovsky, and L. J. Schulman, "Optimal coding for streaming authentication and interactive communication," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 133–145, 2015.

[15] Nasaruddin and T. Tsujioka, "A novel design of reconfigurable wavelength-time optical codes to enhance security in optical CDMA networks," *IEICE Transactions on Communications*, vol. E91-B, no. 8, pp. 2516–2524, 2008.

[16] T. Zhang and Q. Ding, "Design of (15, 11) Hamming code encoding and decoding system based on FPGA," in *2011 First International Conference on Instrumentation, Measurement, Computer, Communication and Control*, pp. 704–707, Beijing, China, 2011.

[17] R. Ma and S. Cheng, "The universality of generalized hamming code for multiple sources," *IEEE Transactions on Communications*, vol. 59, no. 10, pp. 2641–2647, 2011.

[18] R. Kurnia, "Hamming coding for multi-relay cooperative quantize and forward networks," in *2016 IEEE Region 10 Symposium (TENSYMP)*, pp. 321–325, Bali, Indonesia, 2016.

[19] A. Bravo-Santos, "Polar codes for the Rayleigh fading channel," *IEEE Communications Letters*, vol. 17, no. 12, pp. 2352–2355, 2013.

[20] Y. Away, "Performance of trajectory plot for serial concatenation of BCH and convolutional codes," in *2013 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pp. 26–30, Yogyakarta, Indonesia, 2013.

[21] L. Hanzo, L.-L. Yang, E. L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronisation, Networking and Standards*, John Wiley & Sons, Chichester, UK, 2003.

[22] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding: For Transmission over Fading Channels*, John Wiley & Sons, Ltd, 2002.

[23] L. Hanzo, S. X. Ng, W. T. Webb, and T. Keller, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*, IEEE Press-John Wiley, 2004.

[24] G. Forney, *Concatenated Codes*, MIT Press, Cambridge, MA, 1966.

[25] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo-codes. 1," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, vol. 2, pp. 1064–1070, Geneva, Switzerland, Switzerland, 1993.

[26] H. V. Nguyen, C. Xu, S. X. Ng, and L. Hanzo, "Near-capacity wireless system design principles," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1806–1833, 2015.

[27] P. Elias, "Coding for noisy channels," *IRE convention record*, vol. 3, pp. 37–46, 1955.

[28] J. M. Wozencraft, "Sequential decoding for reliable communication," *IRE national convention record Conv. Rec*, vol. 5, part 2, pp. 11–25, 1957.

[29] R. Fano, "A heuristic discussion of probabilistic decoding," *IEEE Transactions on Information Theory*, vol. 9, no. 2, pp. 64–74, 1963.

[30] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[31] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[32] B. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 48–60, 2001.

[33] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169–1174, 2000.

[34] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2626–2631, 2001.

[35] M. el-Hajjar, O. Alamri, Soon Xin Ng, and L. Hanzo, "Turbo detection of precoded sphere packing modulation using four transmit antennas for differential space-time spreading," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 943–952, 2008.

[36] S. Sugiura, S. Chen, and L. Hanzo, "Cooperative differential space–time spreading for the asynchronous relay aided CDMA uplink using interference rejection spreading code," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 117–120, 2010.

[37] N. S. Othman, M. El-Hajjar, O. Alamri, S. X. Ng, and L. Hanzo, "Iterative AMR-WB source and channel decoding using differential space–time spreading-assisted sphere-packing modulation," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 484–490, 2009.

[38] M. El-Hajjar, R. G. Maunder, O. Alamri, S. X. Ng, and L. Hanzo, "Iteratively detected irregular variable length coding and sphere-packing modulation-aided differential space-time spreading," in *2007 IEEE 66th Vehicular Technology Conference*, pp. 1238–1242, Baltimore, MD, USA, 2007.

[39] J. Li, A. Bose, and Y. Q. Zhao, "Rayleigh flat fading channels' capacity," in *3rd Annual Communication Networks and*

Services Research Conference (CNSR'05), pp. 214–217, Halifax, NS, Canada, Canada, 2005.

[40] M. Divya, "Bit error rate performance of bpsk modulation and ofdm-bpsk with rayleigh multipath channel," *International Journal of Engineering and Advanced Technology*, vol. 2, no. 4, pp. 623–626, 2013.

[41] Soon Xin Ng, M. F. U. Butt, and L. Hanzo, "On the union bounds of self-concatenated convolutional codes," *IEEE Signal Processing Letters*, vol. 16, no. 9, pp. 754–757, 2009.

[42] J. G. Proakis, *Digital Communications*, McGraw Hill Higher Education, 4th edition, 2000.

[43] M. F. U. Butt, *Self-concatenated coding for wireless communication systems [PhD thesis]*, University of Southampton, 2010.

[44] T. Stockhammer, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657–673, 2003.

[45] A. C. McCormick, P. M. Grant, and J. S. Thompson, "A comparison of convolutional and Walsh coding in OFDM wireless LAN systems," in *11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings (Cat. No.00TH8525)*, vol. 1, pp. 166–169, London, UK, United Kingdom, 2000.

[46] C. K. Ho, J. H. Cheong, J. Lee et al., "High bandwidth efficiency and low power consumption Walsh code implementation methods for body channel communication," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 9, pp. 1867–1878, 2014.

[47] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Communications Magazine*, vol. 36, no. 9, pp. 48–54, 1998.

[48] S. Samanta, G. K. Maity, and S. Mukhopadhyay, "All-optical Walsh-Hadamard code generation using MZI," in *2019 Devices for Integrated Circuit (DevIC)*, Kalyani, India, India, 2019.

[49] S. Ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEU International Journal of Electronics and Communications*, vol. 54, no. 6, pp. 389–398, 2000.

[50] S. Ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE transactions on communications*, vol. 49, no. 10, pp. 1727–1737, 2001.

[51] M. F. U. Butt, R. A. Riaz, S. X. Ng, and L. Hanzo, "Near-capacity iteratively decoded binary self-concatenated code design using EXIT charts," in *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pp. 1–5, New Orleans, LO, USA, 2008.

[52] L. Hanzo, M. El-Hajjar, and O. Alamri, "Near-capacity wireless transceivers and cooperative communications in the MIMO era: evolution of standards, waveform design, and future perspectives," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1343–1385, 2011.

[53] N. Minallah, M. F. U. Butt, I. U. Khan et al., "Analysis of near-capacity iterative decoding schemes for wireless communication using EXIT charts," *IEEE Access*, vol. 8, pp. 124424–124436, 2020.

[54] M. El-Hajjar and L. Hanzo, "EXIT charts for system design and analysis," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 127–153, 2013.

[55] J. Korhonen and J. You, "Peak signal-to-noise ratio revisited: Is simple beautiful?," in *2012 Fourth International Workshop on Quality of Multimedia Experience*, pp. 37-38, Yarra Valley, VIC, Australia, 2012.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8873912, 16 pages
https://doi.org/10.1155/2020/8873912

Research Article

On the Performance of Wireless Video Communication Using Iterative Joint Source Channel Decoding and Transmitter Diversity Gain Technique

Amaad Khalil,[1] Nasru minallah,[1] Muhammad Asfandyar Awan,[2] Hameed Ullah Khan,[1] Atif Sardar Khan,[1] and Atiq ur Rehman[2]

[1]*Department of Computer Systems Engineering, University of Engineering and Technology Peshawar, Peshawar, Pakistan*
[2]*Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar*

Correspondence should be addressed to Muhammad Asfandyar Awan; mawan@hbku.edu.qa

Received 27 July 2020; Revised 12 November 2020; Accepted 5 December 2020; Published 23 December 2020

Academic Editor: Farman Ullah

In this research work, we have presented an iterative joint source channel decoding- (IJSCD-) based wireless video communication system. The anticipated transmission system is using the sphere packing (SP) modulation assisted differential space-time spreading (DSTS) multiple input-multiple output (MIMO) scheme. SP modulation-aided DSTS transmission mechanism results in achieving high diversity gain by keeping the maximum possible Euclidean distance between the modulated symbols. Furthermore, the proposed DSTS scheme results in a low-complexity MIMO scheme, due to nonemployment of any channel estimation mechanism. Various combinations of source bit coding- (SBC-) aided IJSCD error protection scheme has been used, while considering their identical overall bit rate budget. Artificial redundancy is incorporated in the source-coded stream for the proposed SBC scheme. The motive of adding artificial redundancy is to increase the iterative decoding performance. The performance of diverse SBC schemes is investigated for identical overall code rate. SBC schemes are employed with different combinations of inner recursive systematic convolutional (RSC) codes and outer SBC codes. Furthermore, the convergence behaviour of the employed error protection schemes is investigated using extrinsic information transfer (EXIT) charts. The results of experiments show that our proposed *Rate* – 2/3 SBC-assisted error protection scheme with high redundancy incorporation and convergence capability gives better performance. The proposed *Rate* – 2/3 SBC gives about 1.5 dB E_b/N_0 gain at the PSNR degradation point of 1 dB as compared to *Rate* – 6/7 SBC-assisted error protection scheme, while sustaining the overall bit rate budget. Furthermore, it is also concluded that the proposed *Rate* – 2/3 SBC-assisted scheme results in E_b/N_0 gain of 24 dB at the PSNR degradation point of 1 dB with reference to *Rate* – 1 SBC benchmarker scheme.

1. Introduction

Generally, multimedia communication systems require high data rate, which also results in high demand for transmission power and available bandwidth. Therefore, to transmit wireless multimedia information over limited available bandwidth, high compression efficiency is required. The H.264/AVC codec is a predominant wireless multimedia compression standard because of high compression capabilities required for heterogeneous communication networks and applications [1]. Predictive coding technique and variable-length coding (VLC) increase the H.264/AVC codec compression efficiency required for transmission system, but it also makes the transmitted bitstream more prone to the error [2]. Even a single error in the received bitstream reduces the decoding ability to recover the correct codeword. The predictive coding technique also results in propagating the channel error to its next neighbour video frame. In a wireless system, because of limited bandwidth and varying behaviour of the channel, it makes the video transmission

a difficult task. Layered video coding using unequal equal protection (UEP) technique is used in the H.264 codec for robust video transmission [3]. H.264/AVC with the cross-layered architecture gives better error resilience capability, when used with medium access control (MAC) as discussed in [4]. A transmission system with reversible variable-length codes (RVLC) using irregular convolutional codes (IRCC) that helps in compressing and protecting video codec and uses maximum a posteriori (MAP) algorithm for decoding is discussed in [5]. In [6], the authors discussed convolutional codes with different modulation schemes. Maximum slope (MS) convolutional code is used along with hard and soft decision Viterbi algorithm for decoding. The codeword was mapped to quadrature amplitude modulation (QAM) symbols and quadrature phase shift keying (QPSK) modulation using the additive white Gaussian noise (AWGN) channel. The simulation results of the paper conclude that binary convolutional codes give better results when they were used as inner code in the broadcast channel. Similarly, in [7], the authors have discussed the decrease in energy cost for the communication systems over the wireless channel through an orthogonal coding scheme. Transmission is carried out over the AWGN channel using the differential phase shift keying (DPSK) modulation techniques. The results show a substantial improvement in BER by the use of orthogonal coding and efficient use of transmission signal energy. Furthermore, in [8], the authors have discussed the improvement in concatenated codes. The transmission system employs in its inner code the convolutional code while the outer code comprises of the Hamming code. Block interleaver is used to disperse burst error. The simulation results show better results in BER, when the Hamming code is used as an outer code. The use of space-time coding (STC) to enhance the robustness of data transmission over the wireless channels is investigated in [9]. STC has different coding matrix for multiple input-multiple output (MIMO) transmission, but STBC4 algorithm due to maximum clock transmission steps gives a better peak signal-to-noise ratio (PSNR) and bit error rate (BER). In [10], the authors have presented an H.264/AVC-coded video transmission system using iterative source and channel decoding (ISCD). A novel source bit coding (SBC) and recursive systematic convolutional (RSC) code-assisted IJSCD approach is proposed. The data-partitioned-coded bitstream of H.264/AVC is transmitted with the help of SP-assisted DSTS [10]. The employed SBC scheme improves the performance of our proposed transmission system in the presence of ISCD.

The research paper is organized as fallows. In section 2, we have presented the related works, and section 3 gives details about the H.264/AVC data partitioning. In section 4, we have provided information related to the transmission mechanism of the proposed experimental setup. Section 5 provides the system overview. Furthermore, iterative source and channel decoding are explained in section 6. Details about the EXIT charts are analysed in Section 7, and system performance and results of the paper are presented in Section 8. The conclusion of the paper is presented in section 9.

2. Related Works

Abdullah et al. in [11] debate that H.264/AVC wireless video transmission has problems such as need of higher data networks and error proneness. The study gets its motivation from the use of ultrawide bands for usage of audio-visual signals. The simulations of various scenarios explain the importance of hierarchical and adaptive modulation schemes in various combinations for better video reconstruction. The results from the simulations show a 15 dB increase in PSNR and an increase of 20 dB when added with the various wireless channel adaptive modulation techniques. Nasruminallah et al. in [12] compare three different bandwidth efficient and flexible transceivers for video transmission system using iterative decoding and the simulated Rayleigh channel. The considered three schemes include self-concatenated convolutional, convergent serial concatenated coding, and nonconvergent serial concatenated schemes. Extrinsic information transfer (EXIT) charts show that the SECCC scheme exceeds in performance as compared to the CSCOC and NCSCOC schemes. The BER and PSNR curves also demonstrate that the SECCC scheme performs better for video transmission using iterative decoding. Kadhim et al. in [13] discuss real-time high-quality video transmission with reliability and delay constraints. The typical error protection techniques for example forward error correction and also the automatic repeat request result in the degradation of the video. This paper introduces a partial reliability-based real-time streaming (PERES) technique which is a solution to the application layer that executes partial reliable transfer. The proposed technique consists of acknowledgement and negative acknowledgment system for video transmission and scheduling algorithms with network adaptive algorithms and reliability adaption. Jiyan et al. in [14] propose the design of the H.264 video transmission medium for stationary or mobile user, using the JM tool packet employing the optimization, error protection, and adaptation techniques along the way. The system uses both standard-definition television (SDTV) and high-definition television (HDTV) to input format videos. A complete simulation model with encoder, channel, and decoder is developed. The BER and PSNR values are analysed with varying schemes as GOP, QP, reference frames, and subpixel motion estimation, and the results are shown as graphs. Hadi et al. in [15] present the joint photographic expert group (JPEG2000) image transmission using unequal error protection (UEP) in the presence of polar codes. The proposed transmission scheme achieved better results by using the polarization property of channel codes without significant modification in the overall system. They proposed a joint source channel decoding by using the belief propagation algorithm. The proposed scheme takes the error-resilience tool advantage of the JPEG2000 decoder, which reduces the complexity of system. The experimental results manifest that our designed system has better results as compared to the conventional equal error protection for polar codes. Mhamdi et al. in [16] propose the JPEG2000 image transmission for ISCD using concatenated codes. In this scheme, flexibly UEP is deployed to split the data into several layers so that important source information gets more

protection as compared to less important information. This technique provides better protection along with better decoding performance. The good performance of the designed system is evaluated in the term of a PSNR gain of 10 dB and better subjective quality. The author also presented an adaptive rate allocation scheme which gives better result as compared to static strategy. Hosany suggests in [17] the generalized framework for UEP to evaluate the error performance for rate-compatible puncture convolutional (RCPC) codes and the concatenated Reed-Solomon codes. The transmission system uses 8 PSK modulation schemes in the existence of the Rayleigh fading noise. The designed system uses the MATLAB Simulink and provides better performance with 5 dB difference but increases the overall computational complexity of the system. Chaoui et al. present in [18] the image transmission using joint source channel decoding scheme with arithmetic coding (AC) and resilience technique. The AC technique is very useful in detecting any error occurred in wireless transmission. In the proposed scheme, the JSCD combines the error-detection information feedback of AC decoder with error-free information feedback of the AC decoder. In case of erroneous segment, bit reliabilities are calculated in performing bit back tracking. Bitstream of AC decoder is input to the iterative MPA algorithm, and the result shows 4 to 8 dB better performance as compared to separate source channel model. Balsa proposes [19] the analog JSCD system designed for still images transmission. The proposed system results are compared with digital images such as JPEG and JPEG without entropy. The designed systems show better performance from its alternatives on the basis of the structure similarity (SSIM) index and time required for image transmission. This system does not need to transmit the metadata information, and at the receiver end, analog data is always processed. The proposed analog scheme confirms computational capabilities, low power consumption, and a negligible delay.

3. H.264/AVC

Multimedia transmissions require high compression efficiency owing to limited bandwidth and battery power constraint of wireless systems. Every multimedia application has specific stipulations in terms of compression efficiency, video quality, computational complexity, error resilience, and delay [20]. The H.264/AVC coding scheme is a best solution for such broad-ranged multimedia applications. H.264/AVC is originated as a results of combined efforts of the ITU-T video-coding expert group (VCEG) and International Organization for Standardization (ISO) moving picture experts group (MPEG).The first draft of H.264/AVC was presented in 1999 and after changes in design new draft of this standard was finalized in 2003, which is used for all multimedia application ranging from HD video storage to mobile services. The main goal of introducing this standard is to design a low bit rate and network-friendly video codec that could support a large number of multimedia applications. H.264/AVC delivers better results in terms of robustness in transmission, coding efficiency, and rate distortion efficiency as compared to the predecessor video codecs. H.264/AVC is an efficient video codec design that provides the best performance in real-time communication applications like video conferencing and nonreal-time communication applications like digital television broadcast and video streaming [2].

3.1. H.264/AVC Data Partitioning (DP). Every slice of a macroblock is further subdivided into three partitions based on the importance of data transmission. Data partitioning (DP) is one of the H.264/AVC error resilience techniques in which instead of transmitting the entire video bitstream as a single block video slice, the coded bitstream is partitioned into three slices [2]. The coded information of a macroblock (MB) may be encoded into different video streams called partitions. Each partition has a different sensitivity level. The H.264/AVC video codec supports three different partitions that are types A, B, and C which are discussed below.

(i) *Type A* partitions contain the header information, motion vectors, MB types, and quantization parameters. This partition contains the most sensitive and vulnerable information coded video. If the partition A is corrupted, then B and C are not useful, and the entire partition is counted as a corrupted slice. In such cases, the decoder uses an error concealment technique by using a previously decoded frame of the corresponding video segment [21]

(ii) *Type B* partition carries MB coefficients and MB coded block patterns (CBP) bits of intraframe and represents the chunk of nonzero transform coded coefficients within the block. Bitstream is recovered from errors in the intraframe encoding image regions for certain MBs by switching off interframe prediction. In intraframe coding, the encoding rate is few fractions of MBs, so that is why in this partition, each slice encodes the fewest number of bits [21]

(iii) *Type C* partition holds the interframe motion-compensated error residual (MCER), interframe CBP bits, and uses motion-compensated prediction for encoding MBs bits. In H.264/AVC, the intraframe prediction mode is used for intraframe CBP and intraframe MCER bits for encoding MBs [21].

In the H.264/AVC video codec, partition A is the most vital and essential chunk of video bitstream. In the absence of partition A, it is not possible to decode partitions B and C. Intraframe macroblock information is added in the presence of partition B, with partition A to reconstruct the slice. Similarly, in the presence of partition C with partition A, the reconstructed MCER slice is attached to the motion-compensated slice [21].

4. Transmission Mechanism

The proposed transmission mechanism comprising sphere packing (SP) modulation and differential space-time spreading (DSTS) channel diversity gain technique is presented as follows.

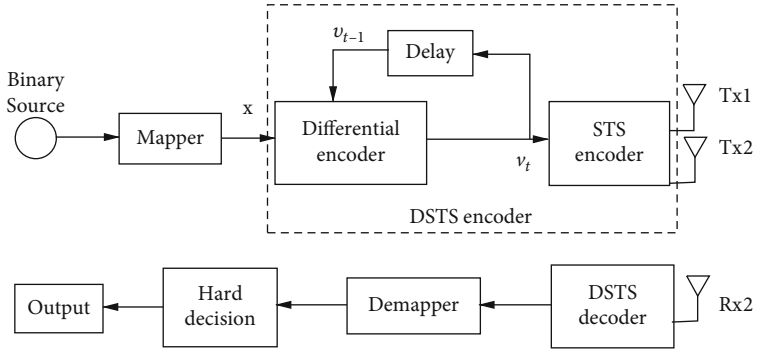


Figure 1: DSTS Encoder.

4.1. Sphere Packing (SP). Sphere packing (SP) modulation is used for modulated symbols to keep the maximum possible Euclidean distance between the modulated symbols. Space-time block code- (STBC-) based orthogonal design of size (2 x 2) for two transmitted antennas are represented as follows.

$$G_2(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \\ x_2' & x_1' \end{bmatrix}, \tag{1}$$

where x_1' represents the complex conjugate of x_1 while column and rows of the above equation represents the spatial dimensions and temporal dimension for two consecutive time slots of two antennas .This scheme consists of two complex modulated symbols (x_1, x_2) that are examined by SP modulation-based orthogonal design for transmission in $T = 2$ time slots from two antennas. The signal is transmitted with L precise space-time signal in consecutive $T = 2$ time slots from the two antennas $(x_{1,l}, x_{2,l}), l = 0, 1, 2 \cdots .L - 1$, where the SP-modulated symbol is represented by L. The aim of jointly designed $x_1 \wedge x_2$ in SP modulation is to enhance the error resilience feature of the system by producing the best minimum Euclidean distance to the remaining $L - 1$ permissible transmitted space-time signals [22].

4.2. Differential Space-Time Spreading (DSTS). The space-time coding (STC) scheme is used to exploit the autonomous fading of the signal of two antennas and create an effectual diversity technique to mitigate the shortcomings of wireless channel. The aim of the STC scheme is to attain a significant power gain and diversity as compared to the single input-single output (SISO) scheme. Space-time block codes (STBC) are a type of STC, proposed by Alamouti [23]. STBC works on a block of data and provides better diversity gain. The STBC technique requires channel estimation and uses coherent detection. Due to the channel estimation technique, the ch4annel experiences an increase in the complexity and cost of the receiver. During transmission, high transmission power is required due to the overhead of fast fading, which increases the number of training symbols. In comparison to this scheme, differential space-time spreading (DSTS) is constituted, which does not require any channel estimation technique. DSTS is a specific scheme for the low-complexity MIMO system by using a noncoherent detection method. The DSTS system gives low complexity, with a trade-off around 3 dB performance loss, as compared to the complex coherent receivers. DSTS consists of two main components that are differential encoder and space-time spreading encoder. In DSTS encoder, the mapped symbols are differentially encoded first and subsequently using STS; they are spread as shown in Figure 1 [23, 24].

At time $t = 0$, the arbitrary dummy reference symbols v_0^1 and v_0^2 are passed to the STS encoder from where these are transmitted via two antennas to the receiver side. Equations (2) and (3) show that the symbols v_t^1 and v_t^2 are differentially encoded as follows [25].

$$v_t^1 = \frac{\left(x_1 \times v_{t-1}^1 + x_2 \times v_{t-1}^{2*}\right.}{\sqrt{\left(\left|v_{t-1}^1\right|^2 + \left|v_{t-1}^2\right|^2\right)}}, \tag{2}$$

$$v_t^2 = \frac{\left(x_1 \times v_{t-1}^2 - x_2 \times v_{t-1}^{1*}\right.}{\sqrt{\left(\left|v_{t-1}^1\right|^2 + \left|v_{t-1}^2\right|^2\right)}}. \tag{3}$$

The differentially encoded symbols are passed to the STS encoder, where symbols are spread assisted by spreading codes c_1 and c_2 and forwarded to antenna for transmission as shown in Figure 2. The spreading code ensures that after using the code concatenation rules, both spreading codes c_1 and c_2 are orthogonal as represented in Equation (4) and (5).

$$c_1^T = [c\, c], \tag{4}$$

$$c_2^T = [c - c]. \tag{5}$$

The differentially encoded symbols split into two substreams, and the two successive symbols are subsequently spread to both antennas for transmission as shown in Figure 2 and represented in Equations (6) and (7).

$$y_t^1 = c_1 \times v_t^1 + c_2 \times v_t^{2*}, \tag{6}$$

$$y_t^2 = c_1 \times v_t^2 - c_2 \times v_t^{1*}. \tag{7}$$

The received signal at a single-receiver antenna is to be denoted by r_t as shown in Equation (8). The nondispersive complex-valued channel impulse response for first and

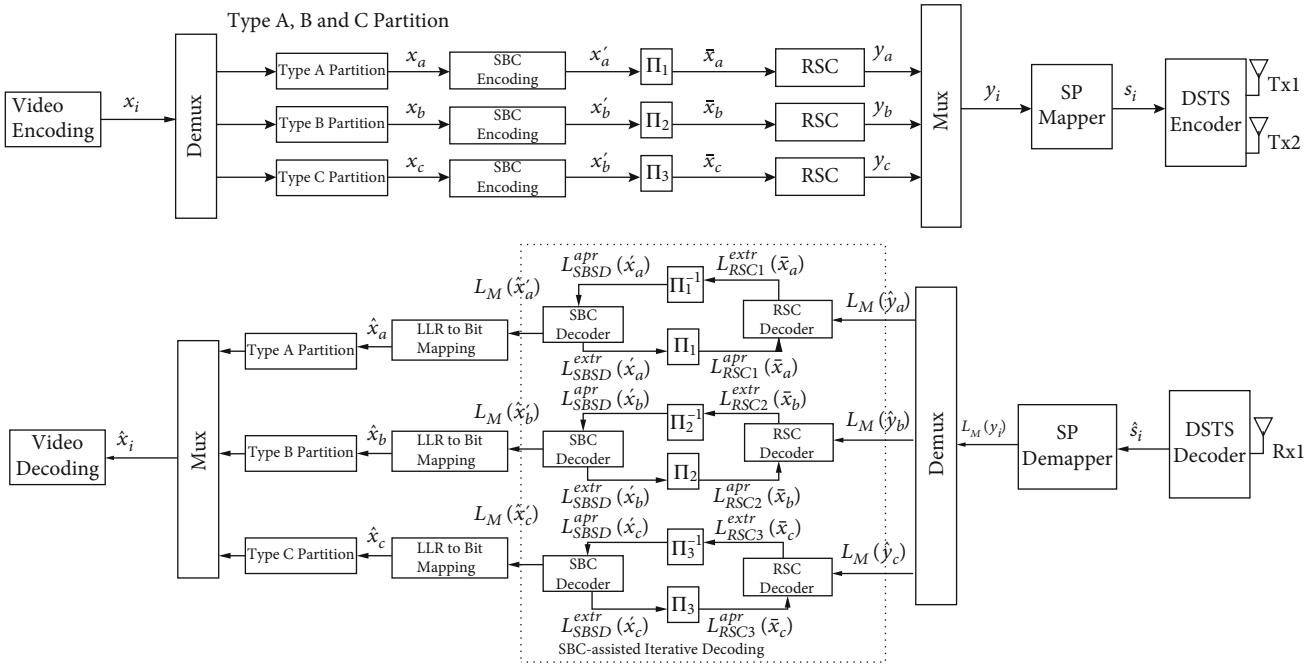


Figure 3: Proposed system design diagram.

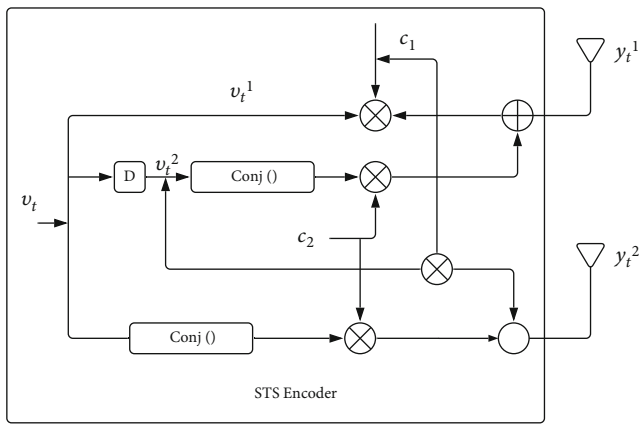


Figure 2: STS encoder.

second antennas is represented by h_1 and h_2. The AWGN channel with a variance of σ_n^2 is denoted by n_t.

$$r_t = h_1 \times y_t^1 + h_2 \times y_t^2 + n_t. \tag{8}$$

In Equations (9) and (10), codes c_1 and c_2 are correlated with received signal r_t, and two data symbols denoted by d_t^1 and d_t^2 are received. H represents the Hermitian matrix.

$$d_t^1 = r_t \times c_1^H = h_1 \times v_t^1 + h_2 \times v_t^2 + c_1^H \times n_t, \tag{9}$$

$$d_t^2 = r_t \times c_2^H = h_1 \times v_t^{2*} - h_2 \times v_t^{1*} + c_2^H \times n_t. \tag{10}$$

Differential decoding is achieved by using received data symbols of successive time slots as shown in Equations (11) and (12). The Gaussian random variables having zero mean complex value are denoted by N_1 and N_2 having a variance of σ_N^2.

$$d_t^1 \times d_{t-1}^{1*} + d_t^{2*} \times d_{t-1}^2 = \left(|h_1|^2 + |h_2|^2\right) \times \sqrt{\left|v_{t-1}^1\right|^2 + \left|v_{t-1}^2\right|^2} \times x_1 + N_1, \tag{11}$$

$$d_t^1 \times d_{t-1}^{2*} - d_t^{2*} \times d_{t-1}^1 = \left(|h_1|^2 + |h_2|^2\right) \times \sqrt{\left|v_{t-1}^1\right|^2 + \left|v_{t-1}^2\right|^2} \times x_2 + N_2. \tag{12}$$

The above equation shows that signal fading (h_1 and h_2) independently works in each transmitter. The proposed technique assures to obtain a diversity gain by the use of a low-complexity algorithm. The space-time spreading operation requires no extra spreading code for transmitting symbols from two antennas in the same time slot.

5. System Overview

In our experimental setup, 300 frames of the H.264-encoded "Akiyo" video sequence are considered for simulation. The diagram of our designed video transmission scheme is presented in Figure 3. The H.264/AVC codec is employed for encoding the video pattern at the transmitter side as shown in Figure 3. The input video sequence has been fragmented by the demultiplexer into three bitstreams, namely Stream A, Stream B, and Stream C. Each stream output contains partition A, B, and C bitstreams in a sequential concatenated manner of all slices of each frame. The output bitstream x_a, x_b, and x_c from demultiplexer are mapped by

Table 1: Different SBC schemes with corresponding symbols and $d_{(\mathrm{H,min})}$.

SBC type	Symbols in decimal	d
Rate – 1 SBC	{0,1}	1
Rate – 2/3 SBC	{0,3,5,6}	2
Rate – 3/4 SBC	{0,3,5,6,10,12,15}	2
Rate – 4/5 SBC	{0,3,5,6,10,12,15,17,18,20,23,24,27,29,30}	2
Rate – 5/6 SBC	{0,3,5,6,10,12,15,17,18,20,23,24,27,29,30,33,34,36,39,40,43,45,46,48,51,53,54,57,58,60,63}	2
Rate – 6/7 SBC	{0, 3, 5, 6, 10, 12, 15, 17, 18, 20, 23, 24, 27, 29, 30, 33, 34, 36, 39, 40, 43, 45, 46, 48, 51, 53, 54, 57, 58, 60 63, 65, 66, 68, 71, 72, 75, 77, 78, 80, 83, 85, 86, 89, 90, 92, 95, 96, 99, 101, 102, 105, 106, 108, 111, 113, 114, 116, 119, 120, 123, 125, 126}	2

using a source bit coding (SBC) scheme into bit strings. Here, $B = b_a + b_b + b_c$, and $a = 1, 2, ..b_a$, $b = 1, 2, ..b_b$, $c = 1, 2, ..b_c$. The bit interleaver Π is used after SBC encoder to interleave the mapped bitstreams and results into $\acute{x}_a$, $\acute{x}_b$, and $\acute{x}_c$. The interleaver within each partition does not affect and extend the video sequence, but it improves the performance of the iterative decoder. Then, the bit strings are encoded with different code rates by the RSC codes, while output streams after channel encoding are represented by y_a, y_b, and y_c. The bitstreams after encoding through RSC error protection codes are multiplexed and concatenated into a single bit steam y_i. The SP mapper is used to transmit the H.624/AVC bitstream with the DSTS encoder using two transmitter antennas. The SP mapper maps the bitstreams to the SP symbol streams represented by s_j as shown in Figure 3. The DSTS provides diversity gain to achieve the coding advantage with low complexity. This process does not need any information of channel estimation, which results in decreasing the BER and improves the subjective video quality. At the receiver end, the DSTS decoder decodes the received signal from the receiving antenna, and the soft information from the DSTS module is forwarded to the SP demapper. Then, the demultiplexer is used to pass the information to the RSC decoder towards its corresponding partition. Each RSC decoder exchanges the extrinsic information with its SBC decoder in the presence of deinterleaver. The deinterleaver helps the SBC decoder module to utilize the residual redundancy. The SBC decoding uses a zero-order Markov model for generating extrinsic information as shown in Equation (13).

$$P\left[\widehat{y}_{(n,k)} \mid y_{(n,k)}\right] = \prod_{i=1}^{n} P\left[\widehat{y}(i)_{(n,k)} \mid y_{(n,k)}\right]. \tag{13}$$

Received n^{th} bit of the k^{th} symbol is represented by $\widehat{y}_{(n,k)}$, and $P[\widehat{y}^{[\mathrm{ext}]}_{(n,k)} \mid y^{[\mathrm{ext}]}_{(n,k)}]$ expresses the extrinsic channel output information as represented in Equation (14).

$$P\left[\widehat{y}^{[\mathrm{ext}]}_{(n,k)} \mid y^{[\mathrm{ext}]}_{(n,k)}\right] = \prod_{i=0, i\neq\lambda}^{n} P\left[\widehat{y}(i)_{(n,k)} \mid y_{(n,k)}\right]. \tag{14}$$

The channel output information and a priori information of the k^{th} symbol give the values of resultant extrinsic *LLR* as represented in Equation (15).

$$LLR\left[y(\lambda)_{(n,k)}\right] = \log\left[\frac{\sum_{y^{[\mathrm{ext}]}_{(n,k)}} P\left(y^{[\mathrm{ext}]}_{(n,k)} \mid y(\lambda)_{(n,k)} = +1\right).P\left[\widehat{y}(i)_{(n,k)} \mid y_{(n,k)}\right]}{\sum_{y^{[\mathrm{ext}]}_{(n,k)}} P\left(y^{[\mathrm{ext}]}_{(n,k)} \mid y(\lambda)_{(n,k)} = -1\right).P\left[\widehat{y}(i)_{(n,k)} \mid y_{(n,k)}\right]}\right]. \tag{15}$$

6. Iterative Joint Source and Channel Decoding

The main goal of iterative joint source and channel decoding (IJSD) is to aid inner and outer decoders in iterative manner to find the maximum possible extrinsic information. SBC uses the residual and artificial redundancy from the encoded bit pattern of video for extraction of extrinsic information. *Rate* – 1 SBC is not capable to achieve better performance gain due to limited redundancy of encoded bitstream. In the H.264/AVC video, to achieve better performance gain in the presence of IJSCD, we add redundant source-coded bits of video, and the method is referred to as the source bit coding (SBC). The SBC scheme is a new approach created on extracting the property of extrinsic information transfer (EXIT) charts. Low BER can be attained by using the iterative decoding method in which there is an EXIT curve in the form of an open tunnel between the inner and outer decoder. To achieve convergence when there exists open tunnel between the inner and outer EXIT curves, they intersect at the upper-right corner of EXIT chart where $(I_A, I_B) = (1, 1)$. Kliewer in [26] discusses the satisfying condition of perfect iterative convergence which is the minimum Hamming distance $d_H = 2$ between the codeword. This encourages the development of an innovative SBC technique where all codewords of SBC have code rate < 1. This can be searched in finding the code table in which necessary condition is $d_H = 2$. This SBC mapping table guarantees that the outer EXIT curve of the SBC outer code will reach with perfect convergence to point $(I_A, I_B) = (1, 1)$. SBC achieves low BER with perfect convergence curve, and its theoretical justification is discussed above. Here, SBC performance analysis is demonstrated with an example in which optimized SBC mapping with *Rate* – 2/3, 3/4, 4/5, 6/7 (as presented in

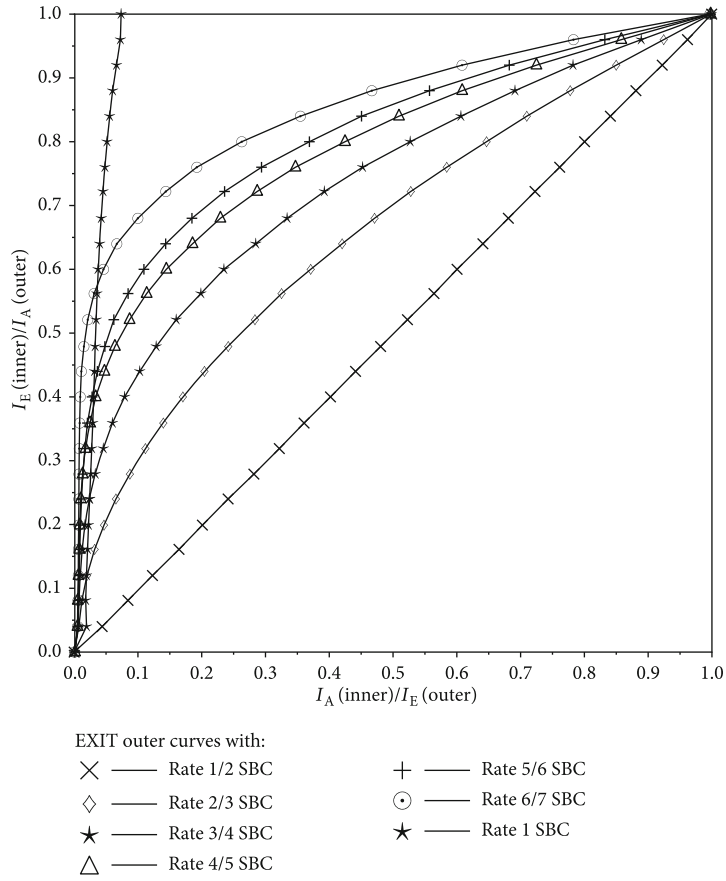


Figure 4: EXIT outer characteristics of different rate SBC coding schemes.

Table 2: Code rate of the different proposed error protection schemes.

S. No.	Outer code (code rate)	Inner code (code rate)	Overall system (code rate)
1	SBC *Rate* – 1	RSC *Rate* – 1/2	*Rate* – 1/2
2	SBC *Rate* – 2/3	RSC *Rate* – 3/4	*Rate* – 1/2
3	SBC *Rate* – 3/4	RSC *Rate* – 2/3	*Rate* – 1/2
4	SBC *Rate* – 4/5	RSC *Rate* – 5/8	*Rate* – 1/2
5	SBC *Rate* – 5/6	RSC *Rate* – 3/5	*Rate* – 1/2
6	SBC *Rate* – 6/7	RSC *Rate* – 7/12	*Rate* – 1/2

Table 1), which is discussed with reference to their EXIT outer curves (as presented in Figure 4). Firstly, it can be observed from the bit mapping presented in Table 1 that all the considered SBC codes of Table 2 ensure the minimum Hamming distance, i.e., $d_H = 2$. As a result, the presented optimized mapping of m to n bit symbols are capable to reach point $(I_A, I_B) = (1, 1)$ of the perfect convergence of the EXIT charts.

7. EXIT Chart Analysis

The inner EXIT characteristic curves of SBC scheme with *R ate* – 1, 2/3, 3/4, 4/5, 6/7 of Table 1 are presented in Figure 4. Figure 4 shows that the EXIT curve for the SBC scheme having code rate < 1 meets at the top-right corner $(I_A, I_B) = (1, 1)$ of the perfect convergence of the EXIT chart. Contrary to this, the *Rate* – 1 SBC scheme falls short of reaching the perfect convergence point. It is important to note that both rate < 1 and *Rate* – 1 SBC scheme maintain an identical bit rate budget for all the employed combinations of outer SBC and inner RSC codes of Table 2. This convergence property of the SBC scheme with rate < 1 is due to the incorporation of artificial residual redundancy in the SBC coding process. Therefore, logically it is clear that rate < 1 SBC is potentially capable to take the maximum advantage of the iterative decoding mechanism by exchanging the beneficial mutual information to achieve lower BER. On the other hand, EXIT curves for SBC scheme with *Rate* – 1 fail to reach and meet at the top-right corner and are not capable to gain any advantage of

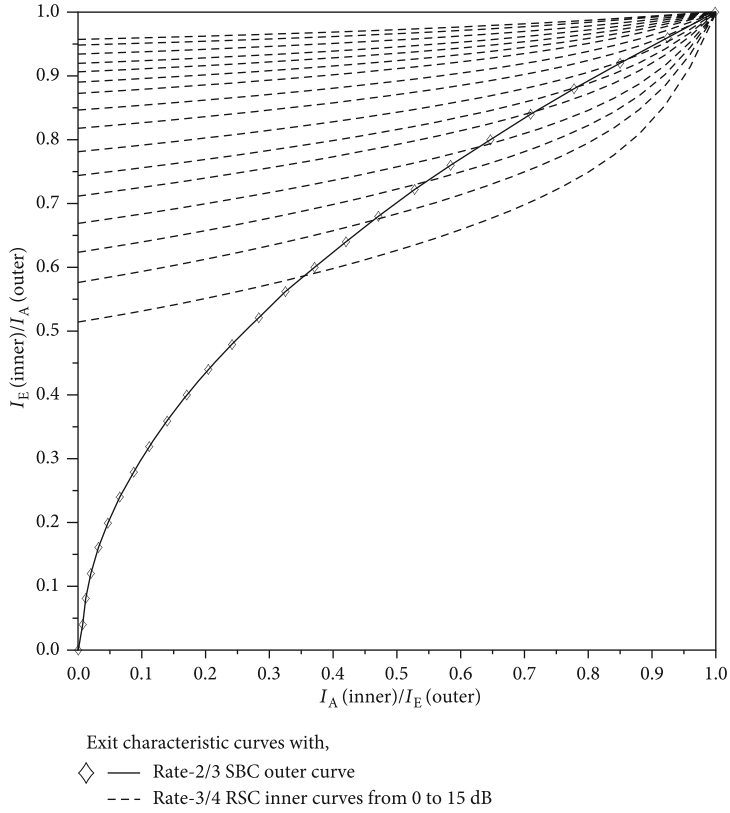


Figure 5: EXIT characteristics curves with *Rate* – 2/3 SBC outer code and *Rate* – 3/4 RSC inner code.

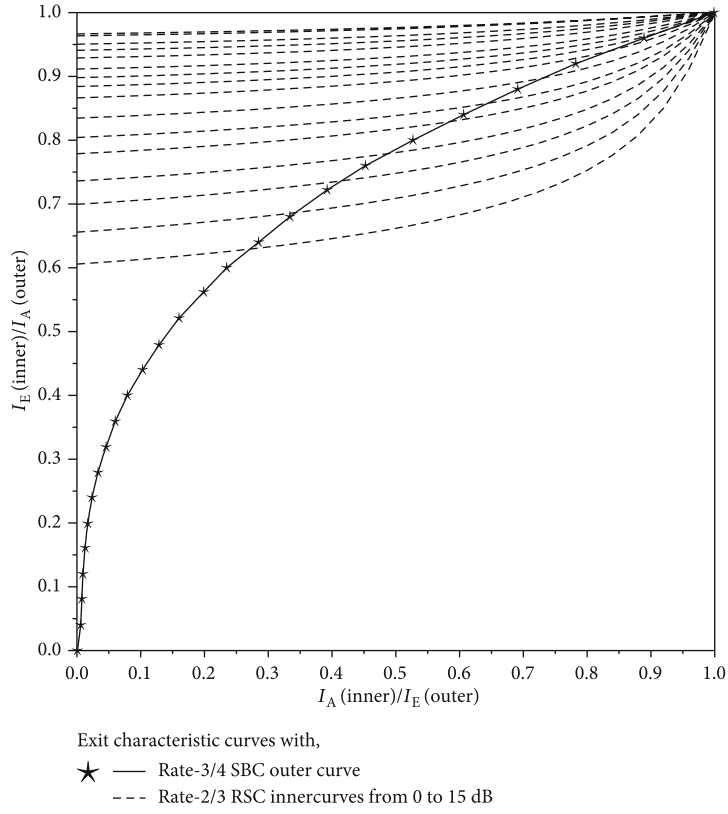


Figure 6: EXIT characteristics curves with *Rate* – 3/4 SBC outer code and *Rate* – 2/3 RSC inner code.

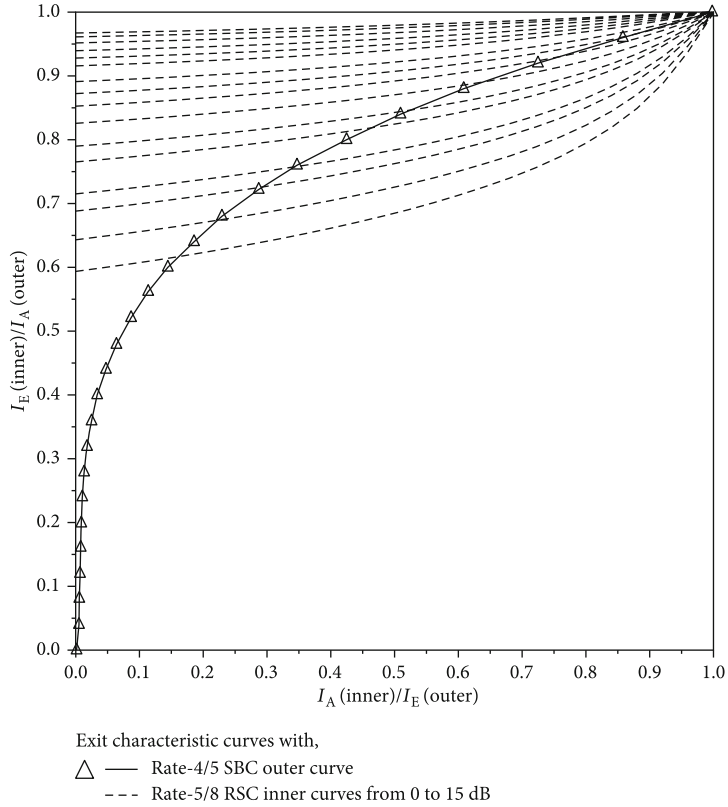


Figure 7: EXIT characteristics curves with *Rate* – 4/5 SBC outer code and *Rate* – 5/8 RSC inner code.

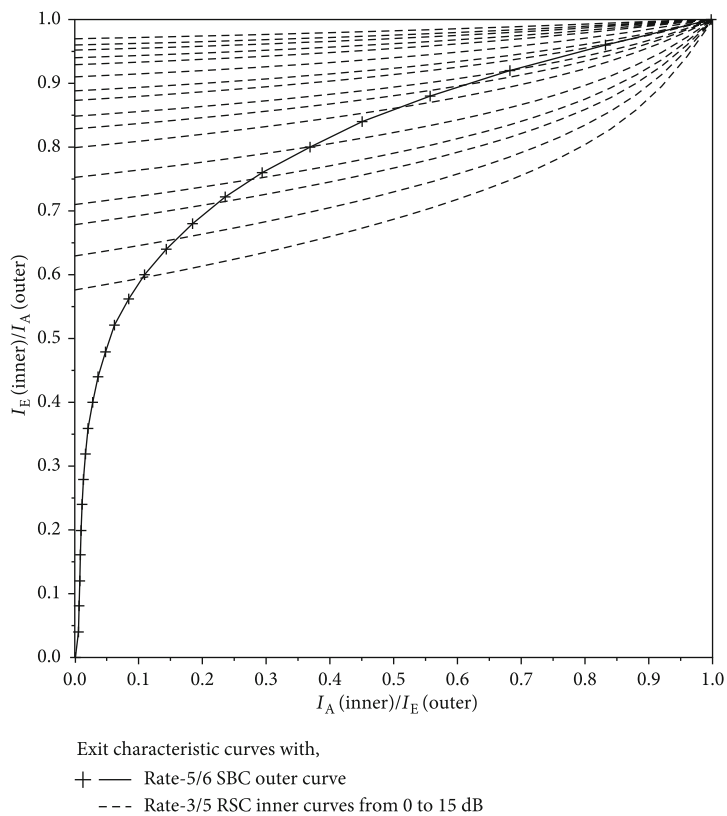


Figure 8: EXIT characteristics curves with *Rate* – 5/6 SBC outer code and *Rate* – 3/5 RSC inner code.

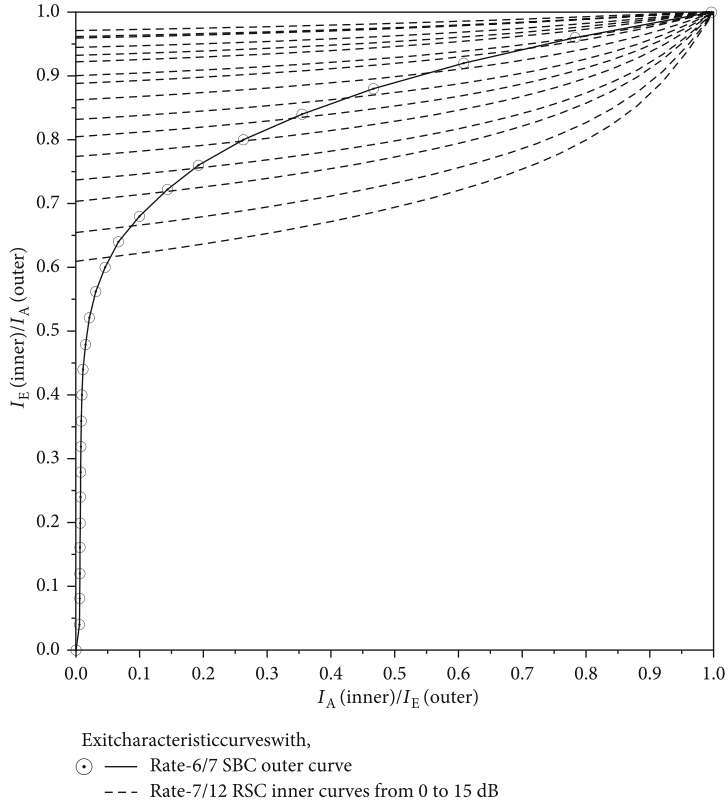


Figure 9: EXIT characteristics curves with *Rate* – 6/7 SBC outer code and *Rate* – 7/12 RSC inner code.

the iterative decoding procedure. With reference to the EXIT outer curves of Figure 4, generated for the different SBC schemes of Table 1, their EXIT characteristic curves along with the corresponding inner RSC curves are presented in Figures 5–9. The presented EXIT characteristic curve shows that the open EXIT tunnel approaches closer to point $(I_A, I_B) = (1, 1)$ of the perfect convergence while employing lower rate SBC as compared to the relatively high rate SBC scheme for the same E_b/N_0 value. More specifically, considering an E_b/N_0 value of 4 dB, the open EXIT tunnel for *Rate* – 2/3 SBC with *Rate* – 3/4 RSC reaches to point $(I_A, I_B) = (0.72,0.85)$ as presented in Figure 5. Similarly, the EXIT tunnels for *Rate* – 3/4, 4/5, 5/6, 6/7 SBC with corresponding RSC code *Rate* – 2/3, 5/8, 3/5, 7/12 of Table 2 reach to points $(I_A, I_B) = (0.6,0.85), (0.47,0.8)$, and $(0.38,0.8)$, respectively. Hence, it can be concluded that the open EXIT tunnel feature of the SBC as the outer decoder and RSC as the inner decoder is more promising while considering a lower rate SBC of Table 2.

Table 3: Systems parameters.

Systems parameters	Value
Source coding	H.264/AVC
Frame rate (fps)	15
Bit rate (kbps)	64
No. of MB's/slice	11
No. of slices/frame	9
Intraframe MB update/frame	3
Channel coding	RSC
Overall code rate	½
MIMO scheme	DSTS
Modulation scheme	SP ($L = 16$)
Number of transmitters	2
Number of receivers	1
Spreading code	Walsh code
Spreading factor	8
Number of users	4
Channel	Correlated Rayleigh fading
Normalized Doppler frequency	0.01

8. System Performance and Results

This part of the paper deals with the explanation of the performance outcome for the suggested schema. The "Akiyo" video pattern [1] contains a quarter common intermediate format (QCIF) of 45 frames, and each frame is 176x144 pixels. The video uses the H.264/AVC JM 19 video codec for encryption, and it is encoded with 64 kbps bit rate for our test sequence at 15 frames per second. Every single QCIF frame is divided into nine segments, and every segment consists of a row of 11 MBs within each QCIF frame. The resultant video sequence contains an intracoded "I" frame, and then 44 predicted frames "P" are placed such that the IPPP

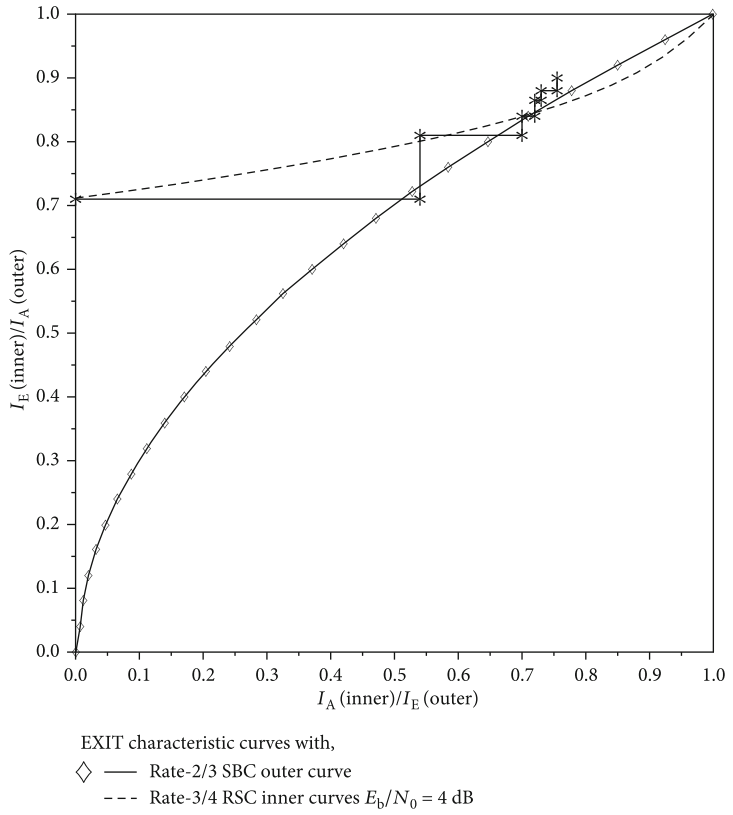


Figure 10: EXIT chart and simulated decoding trajectory of *Rate* – 2/3 SBC scheme of Table 2.

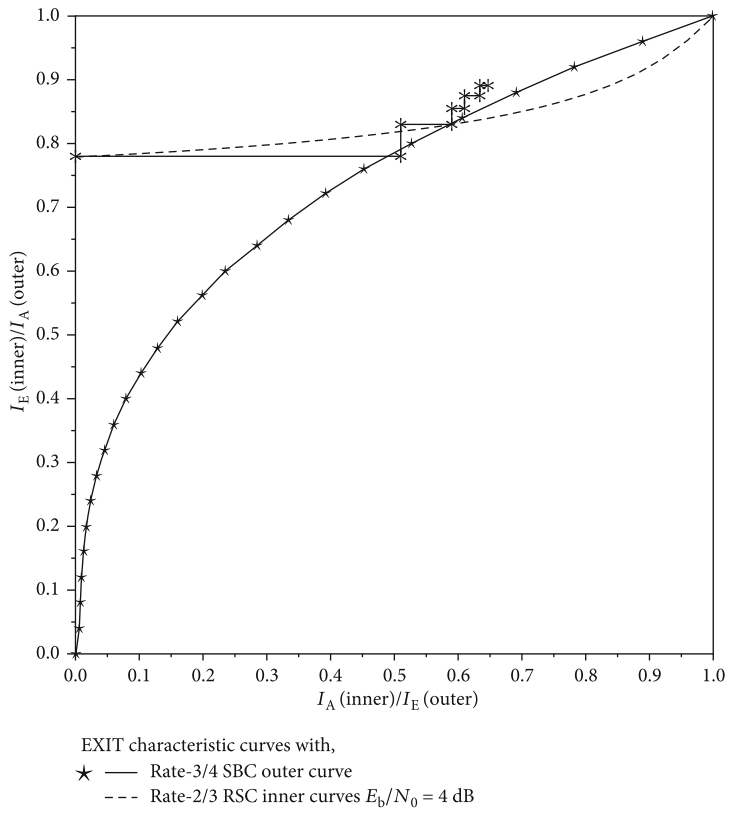


Figure 11: EXIT chart and simulated decoding trajectory of *Rate* – 3/4 SBC scheme of Table 2.

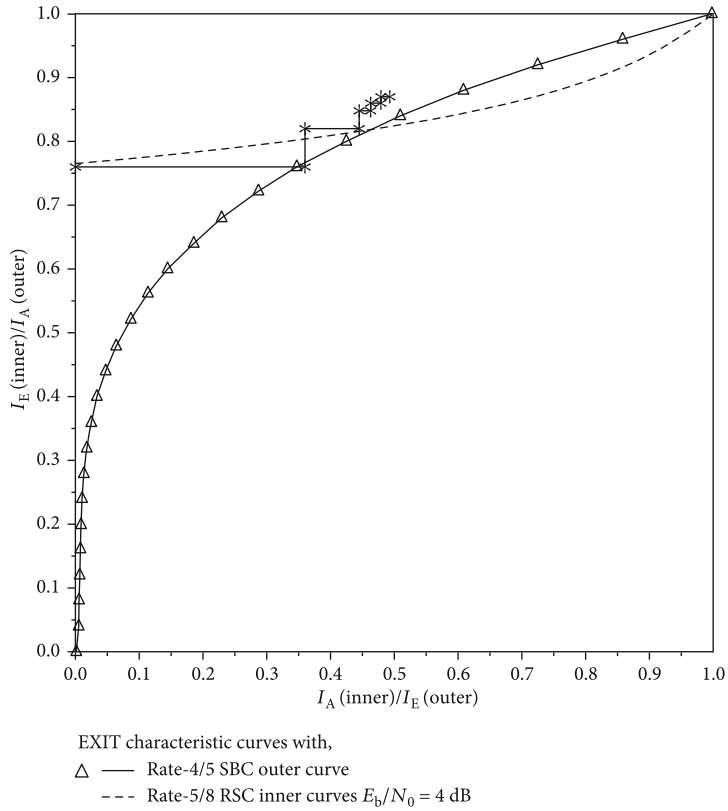


Figure 12: EXIT chart and simulated decoding trajectory of *Rate* – 4/5 SBC scheme of Table 2.

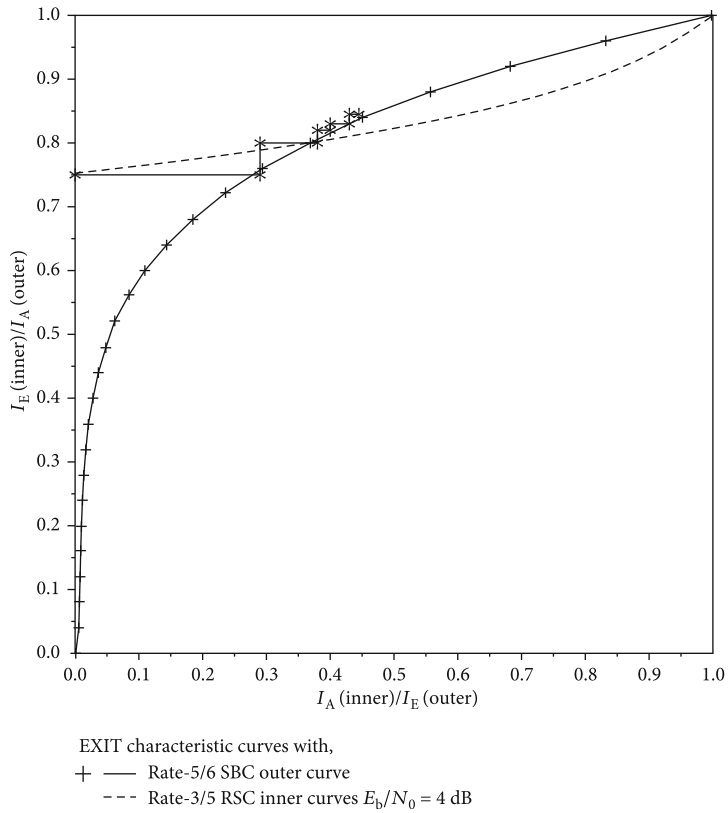


Figure 13: EXIT chart and simulated decoding trajectory of *Rate* – 5/6 SBC scheme of Table 2.

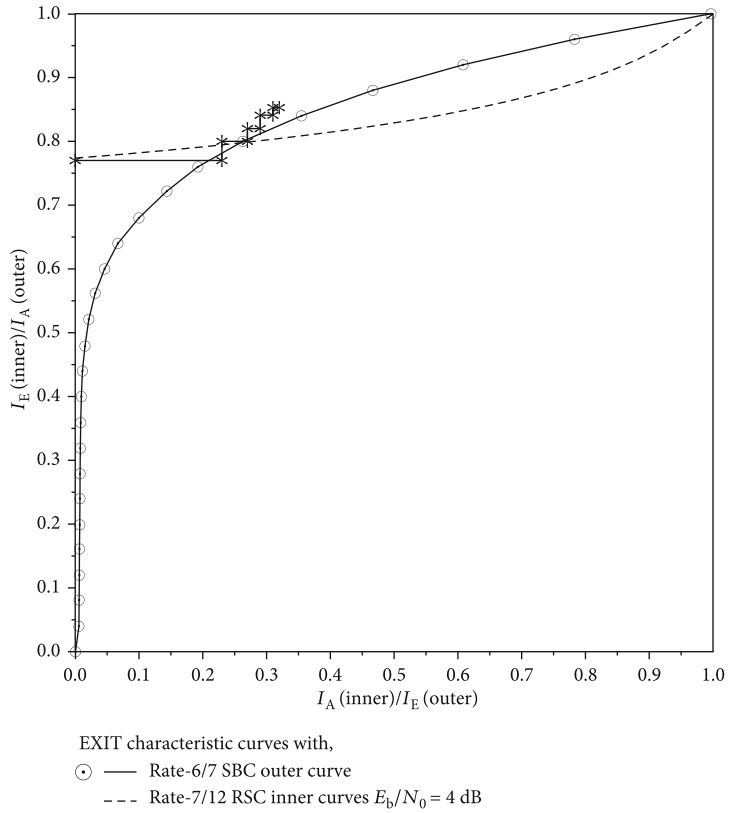


Figure 14: EXIT chart and simulated decoding trajectory of *Rate* – 6/7 SBC scheme of Table 2.

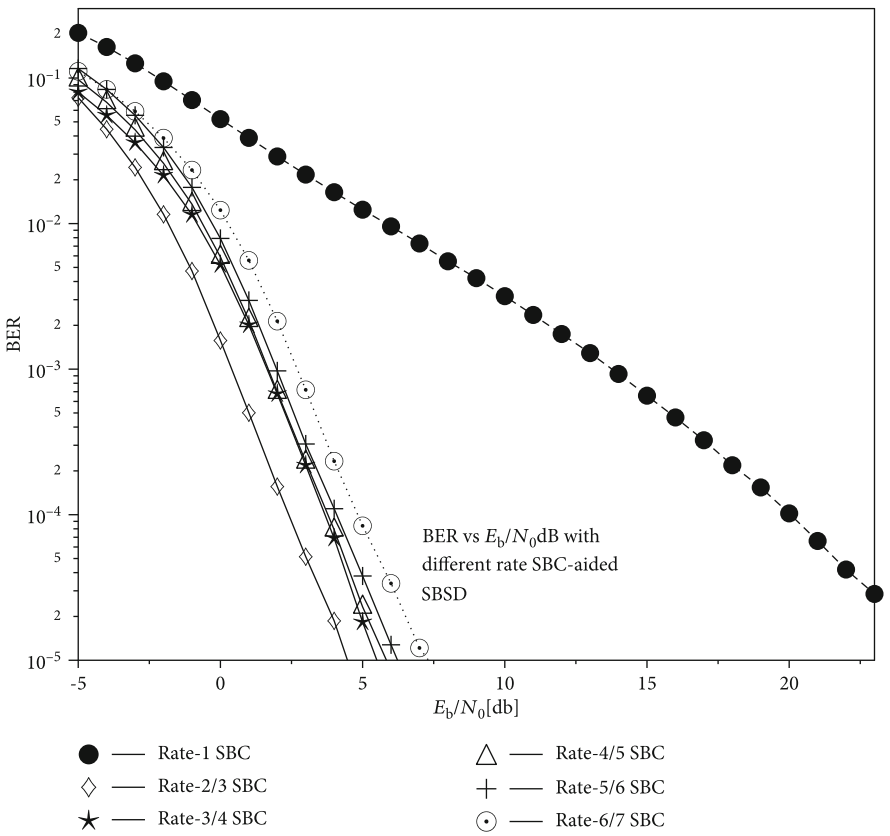


Figure 15: BER performance curves of the coding scheme presented in Table 2.

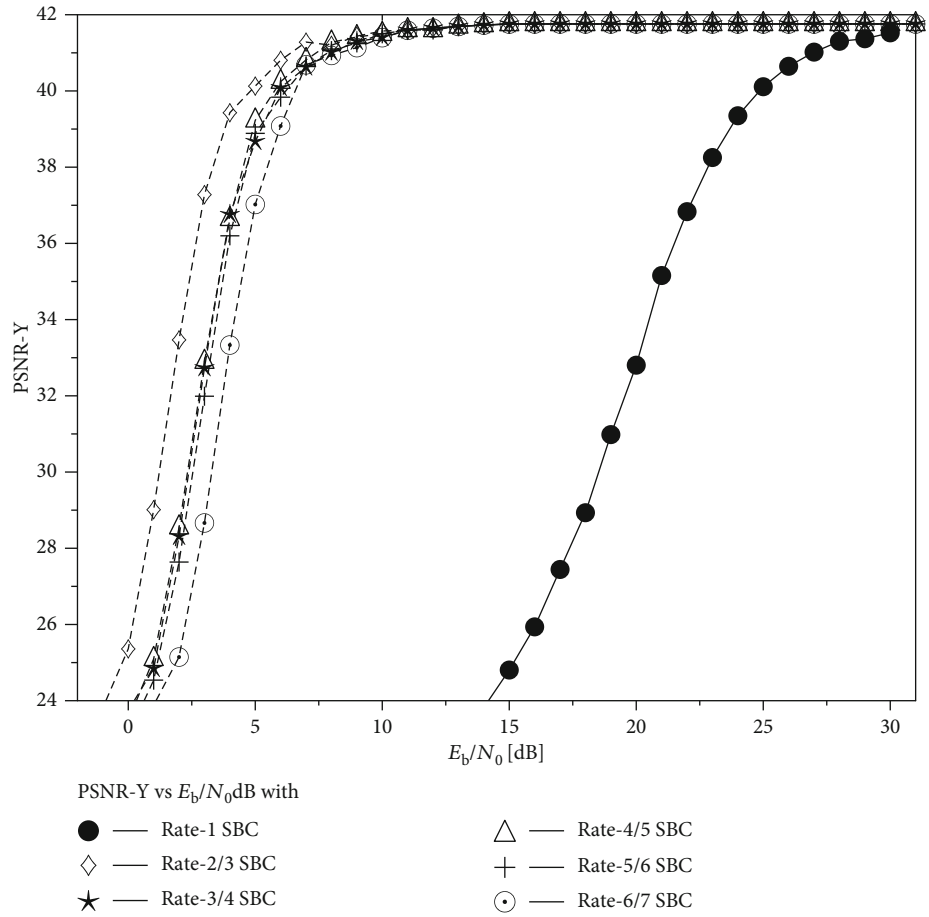


Figure 16: PSNR performance curves of the coding scheme presented in Table 2.

PP… frame sequence is considered in which the "I" frame is repeated after 45 frames within a 3-second duration at 15 frames per second. The intracoded frame has additional benefits in controlling error propagation, so that is why our considered video sequence has a special pattern of "I" and "P" frames. Details about the system parameters of this proposed experimental scheme are presented in Table 3. Flexible macroblock ordering (FMO) and various reference frames utilization, employed for the interframe motion compensation, with additional computational complexity do not have processing performance in a low bit rate video telephony video sequence. Therefore, they were not considered for our H.264/AVC-coded video stream. Source bitstream contains limited residual redundancy. The Monte Carlo simulations were carried out using 45 frames of the "Akiyo" video sequence; experiments were repeated for 260 times, and the average results are considered. SBC with $Rate - 1$ mapping has limited residual redundancy in the coded stream, and therefore, the number of iterations is limited to $I = 3$. For SBC with rate < 1, as presented in Table 1, the mapping obeys the necessary and sufficient condition to reach the upper-right corner of the EXIT chart, and hence, the number of iterations is fixed to $I = 5$. The performance of the various error protection schemes with diverse SBC coding rate was evaluated with the overall same video rate and code rate. From the perspective of H.264/AVC coding, it is pertinent to know that when the frames of low-motion video clips are corrupted due to loss of partition A, the corresponding partitions B and C are not usable, and hence, they are also dropped and the previously decoded frame is used for concealment. A mechanism of motion-compensated prediction is utilized to conceal the lost segment of the future frames. However, a scenario where partition A is received correctly, with loss of partition B of the corresponding video segment, will result in loss of intraframe-coded MB information contained in partition B and hence will result in loss of quality of the corresponding video sequence. The decoding trajectories for the $Rate - 2/3, 3/4, 4/5, 6/7$ SBC schemes of Table 2 are recorded at $E_b/N_0 = 4$ dB as shown in Figures 10–14. Performance analysis of the designed systems using the SBC mapping $Rate - 2/3, 3/4, 4/5, 6/7$ and $Rate - 1$ on the basis of achievable BER and PSNR is shown in Figures 15 and 16, respectively. The SBC $Rate - 2/3$ scheme, with highest redundancy incorporation capability, results in the best BER performance as compared to the other coding schemes of Table 2. Furthermore, it is also observed that the $Rate - 1$ SBC scheme along with $Rate - 1/2$ RSC as inner coding scheme results in worst BER performance, due to its nonconvergence capability in the iterative decoding process. Moreover, it is4 also observed that owing to best BER performance of the

Rate – 2/3 SBC coding scheme, it results in its best PSNR performance, relative to the counterpart coding schemes of Table 2, as shown in Figure 16. More specifically, the *Rate* – 2/3 SBC scheme results in E_b/N_0 gain of 1.5 dB at the PSNR degradation point of 1 dB as compared to the *Rate* – 6/7 SBC scheme having an equivalent overall bit rate. Furthermore, it is also observed from Figure 15 that the proposed *Rate* – 2/3 SBC scheme results in E_b/N_0 gain of 24 dB, with reference to the benchmarker *Rate* – 1 SBC coding scheme, at the PSNR degradation point of 1 dB. Furthermore, it is important to note that both the *Rate* – 2/3 and *Rate* – 1 SBC coding schemes are having an identical overall code rate.

9. Conclusion

In this research work, data-partitioned H.264/AVC video bitstream is transmitted using the iterative joint source and channel decoding (IJSCD) scheme. The performance of different diverse-rated SBC outer-coding schemes was investigated in combination with RSC inner codes, while keeping the overall bit rate budget constant. The source- and channel-coded video stream is SP modulated and transmitted using the DSTS-assisted transceiver. It was demonstrated that the designed IJSCD scheme using the *Rate* – 2/3 SBC scheme gives better BER performance due to incorporation of high level of redundancy in the source bitstream. The convergence behaviour of the presented IJSCD error protection schemes is investigated with the aid of the EXIT charts. The experimental result shows that our *Rate* – 2/3 SBC-assisted error protection scheme with high redundancy incorporation capability gives better results with about 1.5 dB E_b/N_0 gain at the PSNR degradation point of 1 dB as compared to *Rate* – 6/7 SBC-assisted error protection scheme while maintaining the overall bit rate budget constant. Furthermore, it is also concluded that the proposed *Rate* – 2/3 SBC-assisted scheme results in E_b/N_0 gain of 24 dB at E_b/N_0 degradation point of 1 dB with reference to the *Rate* – 1 SBC benchmarker scheme.

Data Availability

The authors approve that data used to support the finding of this study are included in the article.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research work is funded by the National Center of Big Data and Cloud Computer (NCBC), University of Engineering and Technology, Peshawar, under the auspices of Higher Education Commission, Pakistan.

References

[1] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo Transceivers*, Wiley-IEEE Press, 2007.

[2] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657–673, 2003.

[3] M. M. Ghandi, B. Barmada, E. V. Jones, and M. Ghanbari, "H.264 layered coded video over wireless networks: channel coding and modulation constraints," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, Article ID 085870, 2006.

[4] A. Ksentini, M. Naimi, and A. Gueroui, "Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 107–114, 2006.

[5] A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo, "An iterative detection aided irregular convolutional coded wavelet videophone scheme using reversible variable-length codes and map equalization," in *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*, pp. 2404–2408, Dublin, Ireland, April 2007.

[6] V. V. Zyablov and V. G. Potapov, "Error correcting coding schemes for a broadcast channels," in *XVI International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY)*, pp. 32–35, Moscow, Russia, Russia, October 2019.

[7] A. V. Rabin, "Encoding and decoding schemes in communication systems using orthogonal coding for noise immunity's increase," in *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, pp. 1–4, IEEE, Saint-Petersburg, Russia, 2019.

[8] S. Jihwan and H. Lee, "Burst error correction for convolutional code concatenated with Hamming code with a block interleaver," in *International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pp. 531–533, Fukuoka, Japan, February 2020.

[9] M. Ivanov, A. Timoshenko, N. B. Molenkamp, and M. Sokolov, "Implementation of space-time coding model for communication systems MIMO 4x4," in *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 1700–1702, St. Petersburg and Moscow, Russia, January 2020.

[10] M. EL-Hajjar, O. Alamri, S. X. Ng, and L. Hanzo, "Turbo detection of precoded sphere packing modulation using four transmit antennas for differential space-time spreading," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 943–952, 2008.

[11] A. B. Abdullah, A. Zibri, A. Dziri, and F. Tlili, "H.264/AVC video transmission over UWB AV PHY IEEE 802.15.3c using UEP and adaptive modulation techniques," in *International Conference on Advanced Communication Technologies and Networking (CommNet)*, Rabat, Morocco, April 2019.

[12] N. Minallah, M. F. U. Butt, I. U. Khan et al., "Analysis of near-capacity iterative decoding schemes for wireless communication using EXIT charts," *IEEE Access*, vol. 8, pp. 124424–124436, 2020.

[13] H. F. Kadhim, A. H. A. Mahmood, and N. K. Nasser, "H.264 video transmission with high quality and low bitrate over wireless network," *International Advanced Research Journal in Science, Engineering and Technology*, vol. 5, no. 5, 2018.

[14] W. Jiyan, T. Rui, and W. Ming, "Streaming high-definition real-time video to mobile devices with partially reliable

transfer," *IEEE Transactions on Mobile Computing*, vol. 18, no. 2, pp. 458–472, 2019.

[15] A. Hadi, E. Alsusa, and A. Al-Dweik, "Information unequal error protection using polar codes," *IET Communications*, vol. 12, no. 8, pp. 956–961, 2018.

[16] M. Mhamdi, A. Zribi, C. Perrine, and Y. Pousset, "Efficient multiple concatenated codes with turbo-like decoding for UEP wireless transmission of scalable JPEG 2000 images," *IEEE Access*, vol. 7, pp. 6327–6336, 2019.

[17] M. A. Hosany, "Performance evaluation of an unequal concatenated error protection system for the HEVC standard over wireless channels," *Arabian Journal for Science and Engineering*, vol. 45, no. 8, pp. 6489–6500, 2020.

[18] S. Chaoui, O. Ouda, and C. Hamrouni, "A joint source channel decoding for image transmission," *Advances in Science, Technology and Engineering Systems Journal*, vol. 4, no. 6, pp. 183–191, 2019.

[19] J. Balsa, T. Domínguez-Bolaño, O. Fresnedo, J. A. García-Naya, and L. Castedo, "Transmission of still images using low-complexity analog joint source-channel coding," *Sensors*, vol. 19, no. 13, p. 2932, 2019.

[20] S. Wenger, "H.264/AVC Over IP," *IEEE transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 645–656, 2003.

[21] T. Stockhammer and M. Bystrom, "H.264/AVC data partitioning for mobile video communication," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, pp. 545–548, Singapore, October 2004.

[22] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups*, Springer-Verlag, New York, USA, 1999.

[23] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[24] B. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 48–60, 2001.

[25] M. El-Hajjar, O. Alamri, and L. Hanzo, "Differential space-time spreading using iteratively detected sphere packing modulation and two transmit antennas," in *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, pp. 1664–1668, Las Vegas, NV, USA, April 2006.

[26] R. G. Maunder, J. Kliewer, S. X. Ng, J. Wang, L. Yang, and L. Hanzo, "Joint iterative decoding of trellis-based VQ and TCM," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1327–1336, 2007.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8897026, 21 pages
https://doi.org/10.1155/2020/8897026

Research Article

Dimensionality Reduction for Internet of Things Using the Cuckoo Search Algorithm: Reduced Implications of Mesh Sensor Technologies

Azeema Yaseen,[1] **Mohsin Nazir**,[2,3] **Aneeqa Sabah**,[3] **Shahzadi Tayyaba**,[4] **Zuhaib Ashfaq Khan**,[5] **Muhammad Waseem Ashraf**,[6] **and Muhammad Ovais Ahmad**[7]

[1]*Maynooth University, Ireland*
[2]*Asian Institute of Technology, Thailand*
[3]*Lahore College for Women University, Pakistan*
[4]*The University of Lahore, Pakistan*
[5]*Comsats University Islamabad, Attock Campus, Pakistan*
[6]*Government College University Lahore, Pakistan*
[7]*Dept. of Mathematics and Computer Science, Karlstad University, Sweden*

Correspondence should be addressed to Muhammad Ovais Ahmad; ovais.ahmad@kau.se

Received 24 April 2020; Revised 21 September 2020; Accepted 15 October 2020; Published 15 December 2020

Academic Editor: Sungchang Lee

The internet of things is used as a demonstrative keyword for evolution of the internet and physical realms, by means of pervasive distributed commodities with embedded identification, sensing, and actuation abilities. Imminent intellectual technologies are subsidizing internet of things for information transmission within physical and autonomous digital entities to provide amended services, leading towards a new communication era. Substantial amounts of heterogeneous hardware devices, e.g., radio frequency identification (RFID) tags, sensors, and various network protocols are exploited to support object identification and network communication. Data generated by these digital objects is termed as "Big Data" and incorporates high dimensional space with noisy, irrelevant, and redundant features. Direct execution of mining techniques onto such kind of high dimensionality attribute space can increase cost and complexity. Data analytic mechanisms are embedded into internet of things to permit intelligent decision-making capabilities. These notions have raised new challenges regarding internet of things from a data and algorithm perspective. The proposed study identifies the problem in the internet of things network and proposes a novel cuckoo search-based outdoor data management. The technique of the feature extraction is used for the extraction of expedient information from raw and high-dimensional data. After the implementation for the cuckoo search-based feature extraction, few test benchmarks are introduced to evaluate the performance of mutated cuckoo search algorithms. The consequential low-dimensional data optimizes classification accuracy along with reduced complexity and cost.

1. Introduction

The next generation of internet and computers will decline the conventional approach of the internet, reaching to the end-users by promoting the model of interconnecting "smart" objects. It will not replace the internet but will be an addition to (internet) as an infrastructure by integration of physical objects with processing and transmission technologies delivering immense range of services and applications in a more reliable, fast, and accessible way. Such revolution leads towards ubiquitous computing in which every object is embedded with microprocessors and communicate proficiently [1]. It will make physical objects "smart" and let them integrate with worldwide cyber physical frameworks. This trend will pave a way for new openings and innovations in information and communication technologies (ICT), offering new services and applications by connecting physical and virtual objects. It is emerging as a trend in which most

of the surrounding objects will be on network in various forms. This is a shift from conventional internet approaches to the internet for connecting physical objects that interact with each other and humans [2].

The scope of this research work is limited to the application of the cuckoo search algorithm for dimensionality reduction issues in data mining for internet of things application. This work is more related to feature extraction as compared to the feature selection. The further scope of this research work is based on transformation of existing features by using cuckoo search algorithms for data mining. Among indoor and outdoor data service, our focus would be on outdoor data services. The outdoor data could be of many types, e.g., text, images, hypertext, audio, and video, but we are only considering textual or numerical regression of our data for the proposed model [3].

The internet of things enables us to connect with anyone, anytime, and at any place. Technological realms are constructing societies, where everything will be connected. Things have capability to be identified uniquely, operating in smart environments with the help of intellectual interfaces to connect and communicate within the physical world. Interconnected smart objects are heterogeneous and equipped with smart devices (sensors and actuators). Data sensed/captured by these objects is huge in amount and can be termed as "big data" [4]. Knowledge discovery and data mining techniques (classification, clustering, and pattern analysis) are proposed for the internet of things (IoT) by researchers to provide a suitable environment and quality services to people. The extensive volume of data (big data) produced by smart commodities with high dimensions, noisy, irrelevant, and redundant attributes generate a huge search space. If mining techniques are applied on such rough and fuzzy data, it can reduce performance, increase cost, and computation of mining algorithms. Therefore, preprocessing techniques are required to map the original datasets onto new reduced attribute subset, which can represent the original space with high accuracy [5].

In the current era, scope of real time networks and applications is not limited to social and enterprise activities. They are emerging as an extensive discipline to provide advanced and competitive environments for diverse activities including health, home, and business processes. To maintain network robustness and accessibility of proficient services, data analytic techniques are crucial. Data purification to reduce computational complexity of preprocessing and mining models is mandatory. Existing techniques are complex, thus involve large computations [6]. These facts introduce research gap, and formulation of the proposed research work is based on the following objectives:

(i) Outlining untaken challenges concerned to deal with big data analysis for the internet of things and with performance analysis and limitations of existing techniques

(ii) Signifying importance of preliminary processing paradigms for self-organized networks to reduce complexity and enhance performance of mining techniques

(iii) Presenting a framework to reduce curse of dimensionality for the internet of things, based on a nonlinear metaheuristic approach, i.e., cuckoo search algorithm

(iv) Investigating a proposed framework performance

The conducted research work identifies challenges for the outdoor data management in the internet of things and the purpose of feature extraction to extract expedient information from raw and high-dimensional data by the technique of the cuckoo search algorithm [7]. The consequential low-dimensional data optimizes classification accuracy along with reduced complexity and cost [8]. The presented work is fundamentally focused to propose a suitable preprocessing framework for internet of things to outdoor data services [9]. Moreover, the proposed technique is only related to single objective optimization, and it can be expanded for multiobjective nonlinear optimization [10]. Implementation within this document only covers the basic concept of the proposed technique and explicitly for the internet of things' outdoor data services [11]. Dataset used in this research is based on the results of virtual hardware devices and may contain some ambitious attributes as well [12].

1.1. Dimensionality Reduction. In an IoT network, data collected by various sensing entities (e.g., sensors and RFIDs) includes redundant, irrelevant, and noisy information about a dataset. The improved performance and effectiveness of IoT services are achieved by applying frequent mining techniques. Data is collected in unwavering speed which increases the complexity of mining algorithm classifiers due to high data dimensions [13]. Data dimension depicts the several aspects or features to describe an input dataset.

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

Here, D is the dimension, and f_i represents the n features describing the dataset L.

As already mentioned, the accuracy of mining techniques can be affected by high dimensional datasets. So, some preprocessing strategies are necessitated for the transformation of high-dimensional data into lower dimensions (see Figure 1). Data with low dimensionality will optimize classification accuracy along with reduced complexity and cost. Mining the data for IoT is dissimilar to conventional data approaches [14].

In IoT, mining problems occur due to traditional data mining algorithms and hence, they need to be revised for handling scalability and big data issues [15]. Classification problems symbolize many challenges and issues in mining and machine learning research areas [16]. From the IoT point of view, the aim is to classify each object falling into big data according to the characteristics illustrated by its features [17]. Appropriate data representations are needed to describe the physical world data [18], and it is problematic to distinguish the advantageous features. Noisy, redundant, irrelevant, and distorted data can reduce the performance, diminish the classification accuracy, increase cost, and

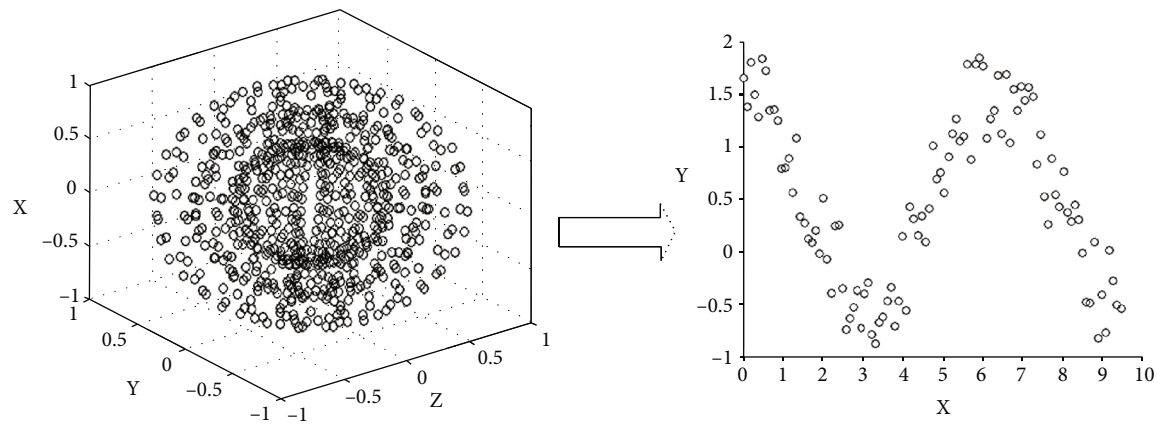


Figure 1: High-dimensional data to low dimensionality showing maximum dimensions of data in (a) 3D (XYZ) plane while graph 2D (XY) on (b) shows data represented by fewer dimensions.

computation of mining algorithms [19]. Eliminating these factors, data mining techniques can benefit and work more efficiently [20]. According to the work proposed in [21], reducing data dimensions (dimensionality reduction) can handle this problem by the selection of relevant and meaningful features only.

If L is a dataset with dimension D, such that $L = \{f_{i=1}^{n}\} \in \mathbb{R}^{D}$, then dimensionality reduction is implied, such that there will be a dataset M having fewer dimensions d, D, and $M = \{f_{i=1}^{n}\} \in \mathbb{R}^{d}$, where $d \ll D$ in such a way that M represents the dataset L with fewer feature subset or dimensions d.

The reduced set of data can expand the performance and speed of mining algorithms, which leads towards optimal classification results and better network performance. Dimensionality reduction is a complicated challenge due to a large search space as the size of data increases exponentially based on its attributes. An attribute may become relevant or redundant in various scenarios according to properties specified by its dimensions. So, optimal search techniques are indispensable for exhaustive search, which is impossible when a search space is indefinite. An inclusive range of searching techniques has been proposed (e.g., sequential forward/backward selection) to select profound attributes to reduce data dimensions which can represent the data in the most appropriate form for classification and other mining strategies. Despite these approaches, attribute reduction practices are undermined from data and algorithm perspectives and going through challenges for dynamic local optimum, exceeding cost and computational complexity [22].

1.2. Techniques for Dimensionality Reduction. Reducing the number of dimensions from the extensive search space is a challenging and demanding call for current networks and computing technologies. Figure 2 describes an abstract process for knowledge discovery and dimensionality reduction. Feature selection (FS) and feature extraction (FE) are two approaches for dimensionality reduction and will be discussed in following sections [23].

Dimensionality reduction (branch of statistics and machine learning) is emerging as a new realm in the solution domain for big data issues to map an original feature space onto a new space This process minimizes the number of random variables (attributes) under consideration and transforms data from high dimensions to low dimensions. Two approaches for mapping can either be choosing a subset of the original feature space (feature selection) or by forming a new space using a transformation function (feature extraction). Feature selection (or feature subset selection) from available datasets is considered more efficient to represent original data comparatively [24].

1.2.1. Feature Selection. Preprocessing techniques are indispensable for data mining to reduce complexity, processing, storage, and cost of classifiers. In IoT, input datasets on data collection layers are represented by high dimension variables and raise the processing complexity of mining algorithms. Feature selection outlines the problem of selecting a feature subset from available candidate features representing originally measured datasets [25]. Feature selection and extraction are also used extensively in image processing and computer vision field of research [26].

A more technical work on the local manifold representation with usage of affinity matrix in the field of dimensionality reduction in hyperspectral imagery [27]. In the field of hyperspectral dimensionality reduction for remote senting data, machine learning models are also extensively used for labeling the graphs along with learning features [28].

Let input to a mining or training classifier is a set of n datasets. Each dataset L can be a set of N features describing the original set of features. The instance L is a tuple including D dimensions as given below.

$$L = \{F_1(L), F_2(L), F_3(L), \cdots \cdots, F_i(L)\}, L \in \mathscr{R}^{D}$$
$$|L| = N, \rightarrow (N = D) \quad (2)$$

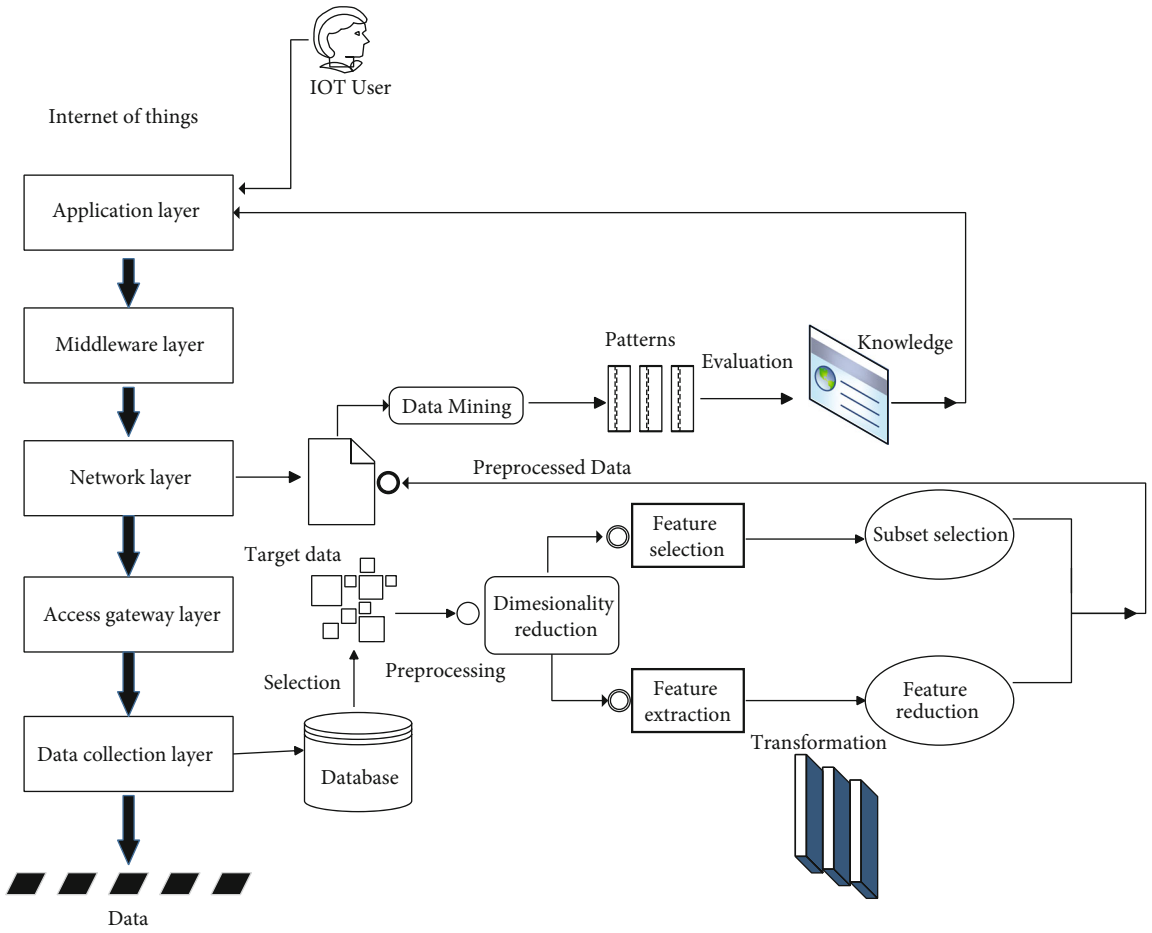


Figure 2: Knowledge discovery and dimensionality reduction for internet of things representing techniques of feature selection or feature extraction to reduce the number of original dimensions for outdoor data in internet of things introduced on the network layer.

where F_i is the domain of the i^{th} feature, and cardinality of L is N, having dimensions D. Let n represents the selected dimensions d in M, $M \subseteq L$.

$$M = \{f_i(M), f_{i+1}(M)\} \rightarrow M \subseteq L \subseteq \mathscr{R}^D, \tag{3}$$

$$|M| = n \rightarrow (n = d). \tag{4}$$

Equations (3) and (4) define a new subset of L which is smaller than the original set and belong to the same search space that is $M \subseteq L \subseteq \mathscr{R}^D$, where number of features d in M is equal to n. The process of feature selection is shown in Figure 3 which is based on two phases. The first part of the process selects subset from original spaces, and the second section evaluates the newly generated subset. Let L be the representation vector and $P(M)$ be the selection criteria for optimal subset M. Formally, feature selection maps the high-dimensional space to low dimension by finding a subset $M \subseteq L$ where $|\mathrm{M}| = d$.

$$\mathrm{P(M)} = \max_{K \subseteq L, d \ll D} P(K). \tag{5}$$

According to equation (5), the higher value of $P(M)$ signifies an improved feature subset. The selected subset performs as a best input to the classifiers and expands the accuracy rate. It is given by the feature selection criterion that only a subset is selected that forms a large set of variables to represent datasets and does not incorporate any transformation and mapping to extract new information from existing high-dimensional datasets.

1.2.2. Feature Extraction. Feature extraction, an imperative data preprocessing technique, adds value to the mining techniques as a performance enhancer for IoT networks to transform existing high-dimensional datasets, i.e., uploading high accuracy data representation model for the original feature space [29].

$$\begin{aligned} &X = \{x_1, x_2, x_3, x_4, \cdots \cdots \cdots, x_n\}, \\ &X_i \in \mathscr{R}^d \Rightarrow Y_i \in \mathscr{R}^p | (p \ll d). \end{aligned} \tag{6}$$

The existence of inappropriate and redundant facts in original datasets demand a prerequisite process (feature extraction). The feature extraction process undermines two research issues: search technique and evaluation measures. The search space includes complete and feature subset,

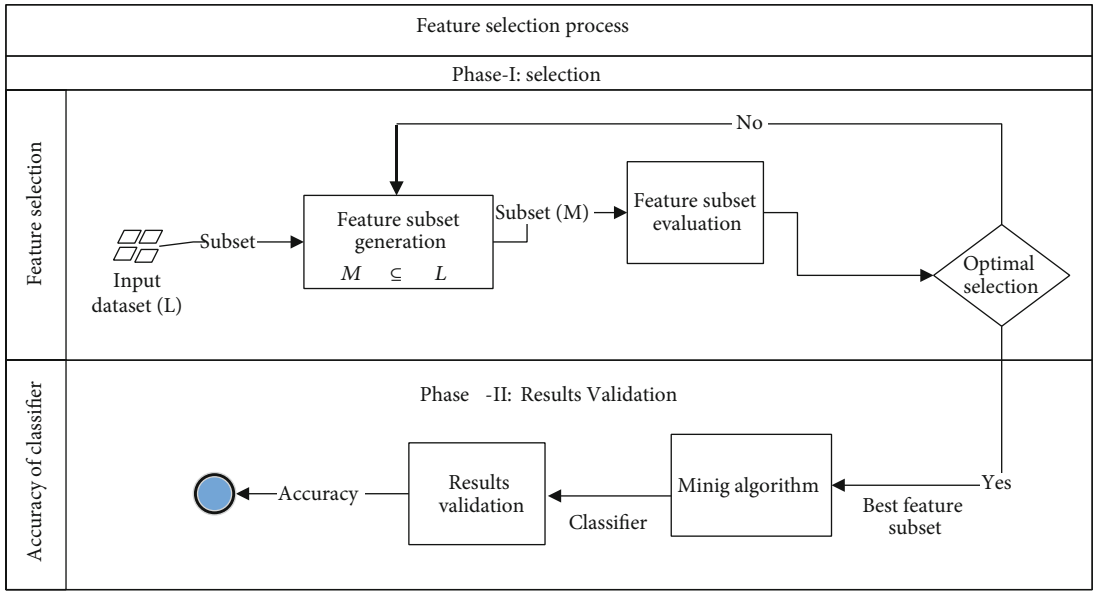


Figure 3: Feature selection-based dimensionality reduction: selection is made in phase I while phase II ensures the accuracy for chosen classifier after the feature subset selection.

and feature extraction transforms existing features to find optimal solution set [30].

$$\sum_{f}^{D} = \mathrm{O}\binom{D}{f} = (1+1)^{D} = 2^{D}. \tag{7}$$

In equation (7), D represents dimension, and f denotes the recent feature subset size. Various evaluation techniques have emerged for the optimal subset selection. Searching for an optimal feature subset is termed as a nondeterministic polynomial- (NP-) hard problem. Traditional searching algorithms are not efficient to handle the high-dimensional search space. Evolutionary computing (EC) algorithms are prominent for their optimal global search competency [31]. Dimensionality reduction is a fact-based problem that determines two basic reasons including reduction of the feature space and to enlarge the accuracy of mining strategies which are demonstrated in Figure 4.

1.3. Heuristic and Metaheuristic Search Methods. Internet of things requires abstract data representation with relatively less number of features, which is a fundamental to data analysis and decision-making tasks. Optimization tools are necessarily a way to find an optimal solution driven by dynamic optimal parameters [32] in a live network like IoT. The dynamically attuned algorithms are applicable when optimization is multiobjective, e.g., maximizing a search optima parallel to another network optimization goal. In such a dynamic network environment, where the data streams are fluent, repititive, and continuous, optimization becomes a dynamic functional requirement and hence, the passive- and problem-specific algorithms do not offer sustainable search optimization for uninterrupted service delivery. The heuristic techniques are problem-dependent and

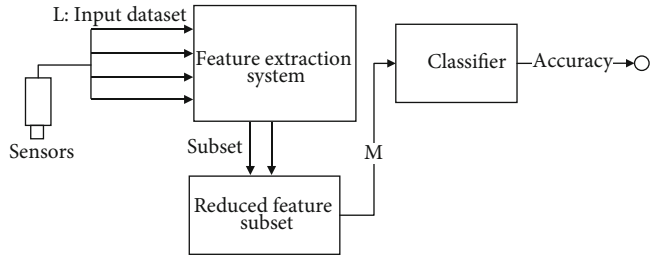


Figure 4: Feature extraction-based dimensionality reduction: data collected by sensors and input L to the feature extraction system for reducing dimensions to get a reduced feature subset M which is sent to a classifier for improved accuracy.

implicitly deliver an approximate solution for a particular situation without an exact accuracy level. The combinatory and low-rank representations based on the heuristic search algorithm in [33] provide an evidence in favour of heuristic techniques but this is only applicable when the objective function is not dependent or followed by another global minima or maxima function.

The studies presented in [34, 35] propose optimization algorithms to get an optimal solution which also ensures the quality and efficiency of the solution with some proving statements.

Metaheuristics are problem-independent techniques that can be applied to a broad range of problems. A heuristic is, for example, choosing a random element for pivoting in Quicksort. A metaheuristic knows nothing about the problem it will be applied, but it can treat functions as black boxes. As a general distribution, the algorithms are defined by two representative categories: (i) deterministic and (ii) stochastic algorithms. In contrast to stochastic algorithms, deterministic techniques are linear, where initial variables

control and determine the output with no random variables, hence, does not need to be adapted for random optimization problems. In stochastic algorithms, random output could be the end outcome of same or random input parameters depending on the triggered operations. The Stochastic algorithm can further be classified by two types of algorithms: evolutionary and metaheuristic algorithms. As given in [36], metaheuristics are naturally and biologically inspired algorithms, offering their applications in various global optimization and real-time problems. Some of these are ant colony optimization [37], particle swarm optimization (PSO), and state transition algorithms (STA) [38, 39]. In spite of the extensive optimum nature, these algorithms have some degree of randomness; it means they reduce the global search ability and easily fall into local optima. To address this, a mutated cuckoo search algorithm is proposed which establishes a solution space as a global function.

1.4. Cuckoo Search Optimization. Cuckoo search (CS) is a metaheuristic optimization technique, proposed in 2009 [40], inspired by some successful characteristics (e.g., breeding) of cuckoo's biological behavior. The growth of this algorithm depends on two terminologies: randomization (random walk) and stochastic search. The algorithm begins to explore the local search space (R_n) for local optima. Moreover, the algorithm is not bounded by local optima; instead, it expands as the problem becomes global and offers a global optimal solution [41]. Though the originally proposed algorithm is tuned to have relatively less number of parameters and dedicatedly targets local optimization, but according to the revised and enhanced CS algorithms in [32], it can be tailored as a dynamic and global optimization algorithm to amend its performance boundary by adjusting the step size and parametric values.

1.4.1. Cuckoo's Living Behavior. Cuckoo is an obligate and brood parasitic organism which depends on other host birds for their reproduction and to grow its offspring [42]. The cuckoo search (CS) is instigated by the influence of cuckoo's genetic activities, e.g., foraging (search for food) [43]. Cuckoo lays its egg in the nest of host birds where eggs hatch and offspring seeks for host attention to get food [44]. Moreover, it also imitates some exterior attributes of host eggs. It is based on two approaches: exploration and exploitation. CS makes use of levy flights to generate a new solution. Cuckoo may throw the eggs of a host bird to raise the hatching probability of its own egg [45].

1.4.2. Algorithm Constraints

(a) Each cuckoo lays one egg at a time and pitches it in a randomly chosen nest

(b) The nest with the best eggs will grow as a next generation

(c) The number of nests is fixed. There is a probability $P_a(0, 1)$ that alien eggs can be identified by the host bird. If it happens, then the host bird either discards eggs or leaves the nest and builds a new one

1.4.3. Algorithm Formulation. Cuckoo selects a nest and dumps its egg into it that is owned by some host bird. The selection of nests depends on the random walk. The randomization is depicted by the foraging and flight behavior of a cuckoo. Each egg laid by a cuckoo represents a new solution V^{t+1}.

$$v^{t+1} = A_{\mathrm{CS}}\left(v^t, P(t)\right), \tag{8}$$

where A_{CS} is a nonlinear cuckoo search algorithm that maps existing IoT d-dimensional vector v^t with parameters $P(t)$ to a relatively new vector v^{t+1}. According to several observations, it is deduced that the natural flying pattern of a cuckoo and the characteristics of a levy distribution are quite similar. These random walks are not isotropic, i.e., vary in directions and magnitude.

1.4.4. Random walk. The formulation of the CS algorithm is an equilibrium consolidation of local and global random walk. Hence, it does not only optimize outdoor IoT data but can also converge to local optima when required. A random walk is a sequence of successive random or stochastic processes.

$$X_i\,(\mathrm{t}+1) = X_i\,(\mathrm{t}) + \text{step size} \oplus \mathscr{L}(\beta), \tag{9}$$

where $\mathscr{L}$ indicates levy flights and $\oplus$ denotes sequenced multiplication for each new step, which is then added to the previous candidate solutions. In each new iteration, a solution is generated through levy flight, and steps for search are taken from the levy distribution bound to heavy-tailed distribution. As compared to the normal distribution, the heavy-tailed distribution is not exponentially bound, and most of the values during generation meet the criteria of the fitness value (objective function) [46]. A random walk is shown in Error! Reference source not found. For 10,000 steps taken at a time t to choose a better position than the previous one, levy flight is preferable when the search space is exponentially unbounded and continuously expands in any dimension and size. Cuckoo search is highly recommendable because of its levy flights to handle network data for global optimization [47].

In information and communication technology, requirements for fast and self-organizing algorithms are indispensable when data is huge in amount and various data analytic activities that are initialized to improve network performance. Metaheuristic algorithms are one of the global optimization techniques that are designed to sort out current global optimization problems. Among several metaheuristics algorithms such as harmony search and bat algorithm, cuckoo search is a newly developed algorithm that can best fit the future smart IoT networks and their continuous outdoor data to provide valuable services with improved machine learning techniques [48]. Figure 5 shows a random walk graph with 10,000 steps starting from 0, where the x-axis represents the time lapse t, while the y-axis shows the position of movement.

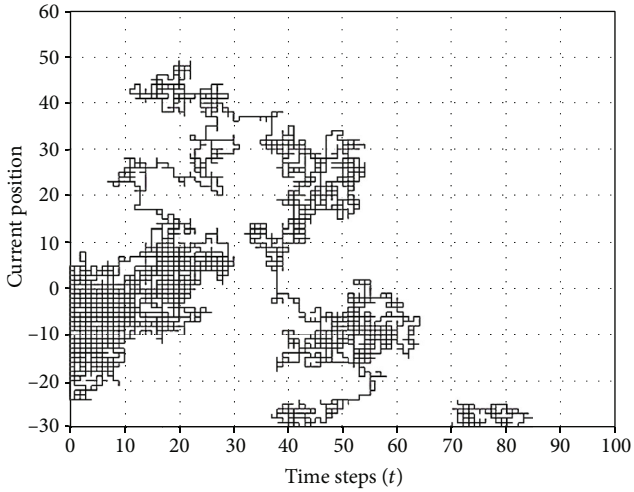


FIGURE 5: Random walk: a graph showing a random walk of 10,000 steps starting from 0, where y-axis represents the current position and x-axis represents time t.

1.4.5. Parameter Tuning. The Cuckoo Search Strategy (CSS), in its originated form, is optimal for local search optimization but its algorithmic constraint can be tuned to broaden its applications for global search optimization problems. In order to achieve the best performance, a number of parameters in the cuckoo search algorithm need to be tuned, namely, the nest size, the elitism probability (probabilistic selection of fittest candidates), and the repetition [49]. Similar to many nature-inspired algorithms, the algorithm starts with random parameters. On each step of the iteration, the parameters are tuned with varying step size. The selection of the step size is important to convergence or divergence of the algorithm. Based on different applications, the step size can be increased or decreased for speedy convergence or performance requirements.

1.4.6. Efficiency. Cuckoo search is a metaheuristic algorithm which is nature-inspired and is now among the most widely used algorithms for optimization. It has many advantages over conventional algorithms due to the inherent randomness in its approach. Metaheuristic algorithms are very diverse, including genetic algorithms, simulated annealing, differential evolution, ant and bee algorithms, bat algorithm, particle swarm optimization, harmony search, firefly algorithm, and cuckoo search [50]. These algorithms are nature-inspired and work without any central computing paradigm. Most of the parameters are tuned with neighbouring nodes interacting with each other. The interconnection among peers makes them as nonexponentially complex.

1.4.7. Limitations. The performance of the cuckoo search algorithm is compromised if the problem is discrete and multiobjective though it performs well for continuous optimization problems. Therefore, the algorithm has limited scope when processing some real-time problems; it demands further study to overcome its limitations. Other than the continuous problems, there has been much development in terms of the step size, parameter adjustment, intercoupling with other algorithms, and other factors used to improve the performance-related markers. Meanwhile, this algorithm also has problems in adaptability and getting the best possible search results, and its algorithmic ability to solve complex problems is inadequate for real-world applications. Future research should be focused towards studying and exploration of new methods and strategies to improve high coupling functions between variables [51].

1.5. Motivation. In wireless sensor networks and IoT-based system, data is generated in enormous volume. This enormous volume is cumbersome to analyze for any fruitful analysis of data. We have to clean this data in any of the phases before the analysis of results is generated from data. Sometimes, this huge volume of data is cleaned during the nonoperational time of data processing, and existing data is updated with removed redundant data. This is also possible that data is cleaned before any analytical processing during runtime [52]. Both approaches have its own pros and cons. With reduced dimensions on scientific and mathematical basis, the data is safe with less volume. The reduced amount of data based on less dimension is always a hot topic of research with diverse implementation of IoT-based systems [53].

The tricky difference between selection and reduction is compromises on selecting the features required during run time execution in feature selection and dropping the undesired and unimportant features during the data cleansing phase. In many real-time applications, searching the huge amount of data with the high-dimensional search space is not practically feasible. The existence of unimportant features causes interferences due to redundancy, irrelevancy, and triviality of the data search space. Many of the evolutionary algorithms lack this inherent attribute for accurate selection of features for deletion and reduction of the attributes. The results in the paper show that cuckoo search outperforms many of the existing and applied algorithms for dimensionality reduction phenomenon [54].

2. Proposed Technique and Implementation

In any IoT network, data from the physical world is highly nonlinear as its environment changes dynamically depending on local or global activities. So, to manage data, preprocessing to mining and decision-making purification is needed for accurate classification and reduced cost. Objects can leave or join the network from time to time, which will need to restart the mining algorithms to deal with immediate and abrupt changes. To reduce the huge search space, selection of reduced attribute subset, and minimizing the cost of mining algorithms, feature extraction (FE) can be an essential strategy to enhance the performance of classifiers and other mining techniques. Moreover, novel algorithms are needed to overcome shortcomings of traditional approaches [55].

2.1. IoT Vectors and Dimensions. The exhaustive search space $\mathbb{R}^d$ consisting of enormous dimensions D is reduced through

feature extraction before any further tasks relevant to data analytics involving data mining techniques. The resultant moderated vectors will increase prediction accuracy with least complexity and cost.

2.1.1. Problem Description. Data gathered by sensors is stored in a database as a combination of rows and columns. Each row represents a distinct vector v_i or dataset and each column $\mathscr{C}$ with its corresponding dimensions d_j. A schema for IoT databases including dimensions and vectors is given in Table 1.

Suppose all distinctive objects in IoT are represented in vector set V, where each vector v_i stores data of a single object with several dimensions (attributes) $d^j \in D$, then the number of vectors in V is equal to the number of tuples in the database, so that $1 \le V \le n$. A set of n vectors can be demonstrated as a column matrix given below:

$$V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \\ v_n \end{bmatrix}, \forall^{v_1}\ V \in \mathbb{R}^n, \tag{10}$$

where each vector v_i can be represented along with its dimensions as a single row matrix.

$$v_i = [d_1\ d_2\ d_3 \cdots d_{m-1}\ d_m], \forall \in \mathbb{R}^D. \tag{11}$$

If we replace each vector with its dimensions, then we get a $m \times n$ matrix.

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

With respect to vector v_i and its dimensions, $[d_1\ d_2\ d_3 \cdots d_{m-1}\ d_m]$ transforms the existing features that belong to the original dimension space $\mathbb{R}^D$ into reduced and transformed vector v_i' and space $\mathbb{R}^d$ such that

$$\begin{aligned} &\text{If } v_i = [d_1\ d_2\ d_3 \cdots d_{m-1}\ d_m], \forall^{v_1} \in \mathbb{R}^D. \\ &\text{Then } T(v_i) = T[d_1\ d_2\ d_3 \cdots d_{m-1}\ d_m], \\ &\qquad => T(v_i) = v_i' = \left[d_1', d_2', d_3' \cdots \cdots \cdots, d_m'\right]. \end{aligned} \tag{13}$$

Before we formulate the algorithm for feature extraction, a brief description for used terms is given in Table 2.

Table 1: Two dimensional data for internet of things: representation for datasets $v_i \in V$ and their corresponding dimensions $d^j \in D$.

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

2.2. Feature Extraction. The component V is an original vectors' set of dimension n, and $\boldsymbol{v_i}$ is a distinct untransformed vector of dimension m. We can illustrate v_i' as a transformed vector of dimension m' such that $v_i' \in \mathbb{R}^d \ll \mathbb{R}^D$. The task of feature extraction is divided into the following steps:

(i) Feature construction (FC) or feature transformation (FT)

(ii) Features' subset selection (searching technique)

(iii) Result efficiency

Depending on these tasks the algorithm for feature extraction can be formulated to maintain a general procedure for meaningful and efficient feature extraction. An algorithm for extraction of high-quality features from the original space is given below.

2.3. Feature Construction. The feature extraction is a technique for transformation of a vast range of features or dimensions into a reduced set of features for various data analytics tasks. Transformation of original features needs some parameters and technique to construct new features from the existing one. These techniques can vary depending on the type of attributes to be constructed. In the following sections, algorithms are presented for nominal and numeric attributes. It will reduce dimensionality of data to enhance search optimization for any machine learning task, e.g., classification and data mining to target the suitable data for further processing in IoT networks. A vector with comprehensive and well-constructed attributes can benefit to achieve high prediction accuracy. Construction of nominal and numeric attributes required different operators and operations to identify hidden information which can be beneficial to data analytics for decision-making [56].

2.3.1. Algorithm for Numeric Attributes. The IoT database includes various dimensions, and each dimension can be represented as a numeric value or combination of characters and strings. In both cases, a set of operators are changed to construct features accordingly. Algorithm given below specifies the steps to construct new features for numeric

Table 2: List of variables and their representation for Internet of Things.

Term	Representation
Dataset/vector in IoT	$V = \{v_i\}$
Number of vectors	$n\{v_i, i = 1, 2, 3, \cdots.., n\}$
Features/dimensions of vector	$D = \{d_j\}$
Number of dimensions	$m\{d_j, j = 1, 2, 3, \cdots.., m\}$
Feature space	$\mathbb{R}^D$
Vector space	$\mathbb{R}^n$
Reduced feature space	$\mathbb{R}^d$

attributes based on arithmetic operators (+,/, -). Figure 6 gives the tree representation for feature construction with various mathematical operations performed for dimensions d^i. Here, the selection of the operators depends on the problem and desired outcome [57].

Algorithm starts with initial input vectors $v_i^{\sim}$, where each distinctive vector is a collection of characters or strings. First loop selects dimensions of an untransformed vector $\boldsymbol{v_i^{\sim}}$ and copies each feature d^j to a new vectorv_i. After construction of selected attributes, original attributes are discarded to avoid duplication of similar attributes. Internal two loops choose attributes from a new vector v_i and select operators from a list of arithmetic operators. Last loop selects each vector one by one and adds all constructed features from $j = 1, 2, 3, \cdots, m$ for each vector v_i, $i = 1, 2, 3, \cdots, n$.

2.3.2. Algorithm for Nominal Attributes. Other than arithmetic operations, concatenation of strings or characters is used to generate new features, if the type of attributes is not numeric. Algorithm for construction of nominal attributes is described below.

Nominal attributes are concatenated to construct new features by making various pairs among all fields in a dataset $v_i^{\sim}$. First loop copies values from the original vector to a new vector v_i.

Last two loops select each dimension from a new array $\boldsymbol{v_i}$ and select each vector sequentially to construct features for all vectors v_i, $i = 1, 2, 3, \cdots., n$. Figure 7 gives the Heaviside function $\theta(p_a - \sigma)$ a representing dimension scaling factor at time t.

The local random walk for local optimum solution is isotropic and can be represented as follows:

$$v_i^{t+1} = v_i^t + \psi s \bigotimes \theta(p_a - \sigma) \bigotimes \left(v_i^j(t) - v_k^j(t), \right. \quad (14)$$

where $v_i^j(t)$ and $v_k^j(t)$ represent two distinctive vectors $(i \wedge k)$ with the j^{th} d dimension at time (t). ψ is a scaling factor for size transformation to control the search space in IoT.

$$\theta(p_a - \sigma) = \begin{cases} 0, & p_a > \sigma \\ 0.5, & p_a - \sigma, \\ 1, & p_a < \sigma \end{cases}$$

$$v_i^j(t+1) = v_i^j(t) + \psi \bigotimes \mathscr{L}(s, \lambda),$$

$$\mathscr{L} \sim u = s^{-\lambda}(1 < \mathscr{L} \leq 3), u \sim (0, 1),$$

$$\psi \bigotimes \mathscr{L} = \{\psi \cap \mathscr{L} : \forall \psi > 0, \forall 1 < \mathscr{L} \leq 3, \psi \cap \mathscr{L} \neq 0\}. \quad (15)$$

Here, λ is taken from a uniform distribution, and each step in levy flight is taken from a heavy-tailed distribution. The levy distribution based on the heavy-tailed distribution increases the probability for selection of each dimension d^j in vector v_i.

2.4. Mutated Cuckoo Search=Based Feature Extraction (CSFE) Algorithm. The improved version of CS is used to extract old and new dimensions along with reduction of overall existing dimensions. Algorithm starts with the step of constructing new features based on original input vectors and selects enhanced attributes for each vector that is given in Table 3 [58].

The maximum optimization is achieved when in each iteration dimensions d^j at extensive distance are chosen to identify how they are compatible to each other in one vector v_i. This task is based on the dimension section (DS) performed according to the rule inspired by cuckoos' strategy of laying eggs in habitat. The DS for each new transformed vector can be calculated as

$DS = \psi$ Number of current dimensions in v_i Total Number of dimensions $\times (d^m - d^i)$.

Algorithm input parameters are as follows:

Dataset environment: $V = \{v_i, i = 1, 2, 3, \cdots, n\}$

Number of datasets in V: n

Discarding probability: p

Scaling factor: ψ

Number of dimensions: $D = \{d^j, j = 1, 2, 3, \cdots, m\}$

Number of iterations: T

Output: globally optimized V'

Auxiliary parameters are as follows: fitness vector v_i with dimensions m', global fitness $\mathscr{G}$, and local fitness $\mathfrak{L}$.

The algorithm for the cuckoo search-based feature extraction is provided in the following section where the objective function is chosen according to the cited problem of outdoor IoT data. The algorithm will generate a feature subset in each iteration and continue this procedure until an optimized cost or performance is achieved.

2.4.1. End Procedure_CSFE. The mutated cuckoo search-based feature extraction (CSFE) algorithm includes three procedures demonstrated as procedure_1 for construction of numeric attributes, procedure_2 for nominal attributes construction, and last is procedure_3 for apply global selection strategy onto constructed features to find more appropriate features which describe each vector $\boldsymbol{v_i}$ to improve prediction exactness. Procedure_3 starts with two fitness

Input initial vectors set $V\sim$ *with original features* $D\sim$
For $(\forall v_i \in V\sim(t-1), i = 1, 2, 3, \cdots, n)$ ***do***
Construct new features F_n *for each* v_i ***(transformation)***
Update each vector v_i *and add it to new vectors set* V_c
$$V_c(t) = C\,(\,V(t-1)) = F_c - v_i$$
Select subset of constructed features from V *through* ***Cuckoo Search***
$$F(V_c(t)) = (V(t)) \rightarrow {}^{CS}V'(t)$$
Generate optimized vectors set V' *with reduced dimensions* D'
End Procedure_1

Algorithm 1: Initiate procedure_1.

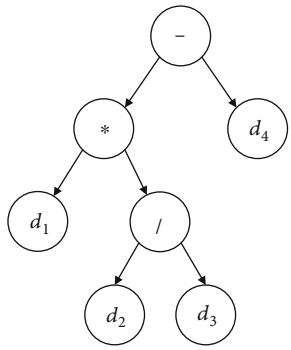


Figure 6: Representation for feature construction: various mathematical operations are performed for dimensions d^i.

IF *(attributes ! = numeric attributes)* ***Then***
Go to procedure_2 for Nominal Attributes
Else *Start Procedure _1 for Numeric Attributes*
Input vector $\boldsymbol{v_i^{\sim}}$*, a set of nominal attributes*
$d^j \in v_i^{\sim}$*, single feature in each vector*
$V\sim = \{t_1, t_2, t_3, \cdots, t_n\}$ *, sequence of n tuples*
$V_c(t) = \varnothing$*, set of newly constructed features*
A *: set of arithmetic operators*
For *each* $(d^j \in v_i^{\sim})$***do***
$v_i = v_i^{\sim} - d^j$ *(Prevent duplication of a feature)*
For$(\forall d^j \in v_i)$
For$(\forall a \in A)$
For$(\forall v_i \in V)$
$F_c = \sum_{i=1}^{m}(\,d^j,\ d^{j+1})$
$V_c(t) = F_c \cup V_c(t)$
End Procedure_2

Algorithm 2: Initiate procedure_2.

functions $\mathcal{F}_V$ for global optimization and $\mathfrak{L}_{v_i}$ to evaluate each vector for local optimization, initially sets the existing n vectors $(v_i, i = 1, 2, 3, \cdots, n)$ as an input to procedure_3. First loop calculates fitness $\mathfrak{L}(v_i)$ for original n vectors sequentially before selection through CS. The next loops select each dimension from (v_i) at time T until it is not equal to the length of tuples in original vectors space and select dimensions randomly through levy flight and generate a new vector v_i' including optimal dimensions m'. In the next step, fitness for new vector v_i' is calculated. If the fitness of the reduced vector is maximum than the existing one, then the replacement is conducted and abandons the worst vector, assembles all best fitted vectors v_i' to V', and finds fitness to compare with the existing one and replace if necessary. Last, **IF** statement finds efficiency for newly built vector space to decide whether it should be discarded or placed for further processing.

Input to the algorithm is the original feature set, and construction of features is performed according to the identified type of input features (numeric or nominal). After the construction, updated space is relocated to cuckoo search for selection of appropriate features through random walk and levy flight. The output is a new and enhanced feature subset for each vector $\boldsymbol{v}$.

2.5. Dataset Generation. The internet of things has evolved as a preeminent and exquisite source to provide valuable services to consumers, business analysts, and industries in their daily professional and personal lives where things can connect themselves to the internet and serve without any delay. Despite this rapid evolution, familiarity and adeptness to the IoT network are quite gradual. Only few highly recognizable industries are providing valued services to their consumers. In literature, most of the work regarding IoT only demonstrates the fundamental concepts and architectural aspects. In the current era when technological advancement is more beyond than the internet, the real-time networks are facing challenges to accommodate continuous and abrupt amount of collected data for various mining and machine learning tasks to accomplish the goal of smart and intelligent networks with self-continuation ability without human intervention.

The internet of things maintains data collected by the entities that are part of it. These entities are computing devices (scanner, thermostat) that can communicate over the internet to share their information and services. These smart objects are distinct, and the EPC of each object can be stored as a primary key in the database to maintain its record. The organizations with advancement of IoT are not willing to share their private data publicly for security and confidentiality. That is why in literature and over the internet dataset relevant to IoT networks are quite unavailable. To apply proposed solutions (CSFE) for dimensionality

IF *(attributes ! = Nominal attributes)* **Then**
Go to procedure_1 for Nominal Attributes
Else *Start Procedure_2 for Numeric Attributes*
Input vector $v_i^{\sim}$, *a set of nominal attributes*
$d^j \in v_i^{\sim}$, *single feature in each vector*
$V \sim = \{t_1, t_2, t_3, \cdots, t_n\}$, *sequence of n tuples*
$V_c(t) = \varnothing$, *set of newly constructed features*
con *():concatenate features in* $v_i^{\sim}$
For *each* $(d^j \in v_i^{\sim})$**do**
$v_i = v_i^{\sim} - d^j$ *(Prevent duplication of a feature)*
For$(\forall d^j \in v_i)$
For$(\forall v_i \in V)$
$F_c = con(d^j, d^{j+1})$
$V_c(t) = F_c \cup V_c(t)$
End Procedure_3

Algorithm 3: Start procedure_3.

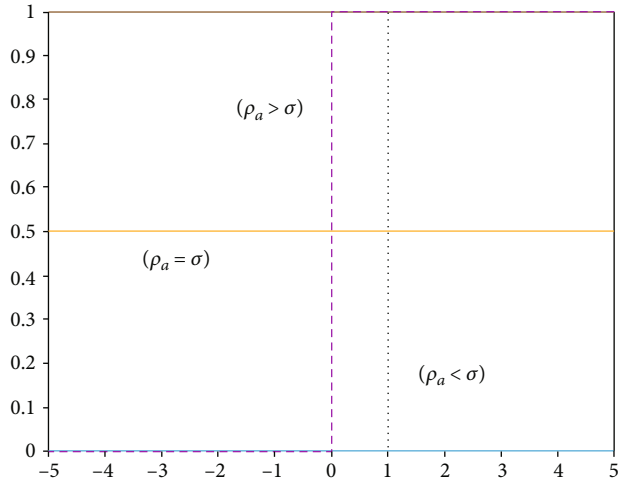


Figure 7: Heaviside function: $\theta(p_a - \sigma)$, representing dimension scaling factor at time t.

reduction onto outdoor IoT data, an appropriate and relevant dataset is required. It will facilitate to produce appropriate results through implementation of suggested technique in MATLAB for reduction of the overall extensive search space $\mathbb{R}^n$ and feature space $\mathbb{R}^D$.

IoTify is a web-based platform for simulation to develop IoT applications by using virtual hardware devices, e.g., sensors. It facilitates a virtual lab and enables the creation and building of virtual IoT devices in JavaScript. The IoTify database is generated using JavaScript object naming (JSON) with extension JSON. Table 4 demonstrates the dataset that is used to accomplish results and for analysis of proposed algorithms (CSFE). IoT-based devices can extract specific and required facts from a patient's blood and will share the generated report to the doctor when an alarming situation arises. Figure 8 outlies a flowchart for the mutated cuckoo search-based feature extraction including procedures for nominal and numerical attributes with generation of final reduced feature subset.

Table 3: List of parameters and their symbolic representation in the cuckoo search-based feature extraction algorithm for dimensionality reduction.

Parameters	Algorithmic representation
Dataset environment	V
Fitness function	$\mathcal{F}(x)$
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	$\mathcal{L}(\lambda)$
Number of iterations	T
Time instance	T
Step size	S
Step size scaling factor	$\psi > 0$
Normal distribution	$\mathcal{u}$ (0,1)

Reduction of original dimensions for outdoor IoT data is performed through the task of feature extraction. Here, the task of feature extraction as subtasks of feature construction and selection from newly constructed features is introduced. Selection from the extensive new search space is done using the cuckoo search-based optimization technique. To evaluate the results for suggested research techniques, an IoT-based dataset is used. This technique will lessen the exhaustive search space and generate a new organized search space that will improve the accuracy of machine learning tasks or mining classifiers.

2.6. Algorithm Result Analysis and Visualization. The mutated cuckoo search-based feature extraction is the proposed algorithm implemented in MATLAB, and results are visualized with graphs, plots, distributions, and statistical operation (mean, minimum, and maximum). Figure 9 represents the plots for newly generated space V' including n rows and m' reduced dimensions. Plots are relatively at distance and scattered that indicates that dimensions are chosen from the extensive search space. Subplot shows the number of iterations for random generations, and residuals are calculated for each dimension to check its weightage for selection. The dimension selection is constructed through levy flight (λ) and step size scaling factor ψ, and the steps are chosen randomly from the levy distribution.

Selection through CSFE searches through the extensive search space $\mathbb{R}^n$ and $\mathbb{R}^d$. Search is exponentially increasing as new objects enter the IoT and need an algorithm to modify itself to adjust for immediate changes, since data generated by IoT is continuous and needs a global optimization solution to enhance network efficiency.

Figure 10 shows the heavy-tailed distribution for CSFE produced for the input dataset of IoT-based Patients' CBC

IF *(Attributes ! = Nominal A ttributes)* **Then**
Go to Procedure_1 for Numeric Attributes
Else *Start Procedure_2 for Numeric Attributes*
Start Procedure_3 for Optimal Subset Selection
$\mathscr{F}_V = n/\text{execution time}(V) - \text{execution time}(V')$
$\mathfrak{L}_{v_i} = m - m'/\text{execution time}(v_i')$
Global Fit= $(\mathscr{G}) = \mathscr{F}_V$
Local Fitness = $\mathfrak{L}_{v_i}$ *= max_ fitness*
Initiate Population of n vectors in V
For *each vector* $v_i(\forall i, i = 1, 2, 3, \cdots, n)$
  Calculate fitness for current vector $\mathfrak{L}(v_i)$
  For *each dimension* $d^j(\forall j,\ j = 1, 2, 3, \cdots, m)$ *in*
    While $(T \neq n)$ *at time instance t,* **do**
      Find v_i' *through levy flight for* v_i *in which* $(v_i \neq 0)$
        $v_i^j(t+1) = v_i^j(t) + \psi \bigotimes \mathscr{L}(s, \lambda$
      Compute fitness for v_i': $\mathfrak{L}(v_i')$
      If $(\mathfrak{L}((v_i') > \mathfrak{L}(v_i))$ *Then*
        Max_ fitness = $\mathfrak{L}(v_i') \wedge$ *discard worst vector* $\mathfrak{L}(v_i)$
      Else *max_ fitness =* $\mathfrak{L}(v_i) \wedge$ *discard worst vector* $\mathfrak{L}(v_i')$
      Set best fitted vector v_i^j *as a new reduced vector to* V'
      If$(F_v' > F_V)$ *Then*
        Set V' *as a new solution with reduced dimensions*
      Else *build new vectors to get required fitness*

Algorithm 4: Start procedure_CSFE.

Table 4: Representation for number of vectors and dimensions for each vector in dataset: internet of things-based patients' CBC results.

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

results. The highest peak represents the global optimization solution for IoT data as tail is exponentially increasing and consistently provides best fitness. The best selection is estimated trough DS and coherence estimation factor π. Area under the curve is not exponentially bounded, that means that as data becomes extinct, it increases the number of fitness values more close to best optimum. In Figure 11, the size of bars shows that most of the fitness values were globally best.

The term "Internet of Things" is considered to represent innovation that relies on both the resulting network by the integration of smart objects along with developed internet technologies and a variety of supporting devices, equipment, and machines that are important to ensure this technological evolution. Applications and services are developed to take advantage of these technologies for new business trends and offering daily life conveniences. IoT is an infrastructure based on networked smart objects and integrated networks as a supplement to internet services by ensuring availability for all kinds of services anytime and anywhere to anyone. It is emerging as a trend in which most of the objects in our surroundings will be on network in various forms. This shifts from conventional internet approaches to the internet for connecting physical objects that interact with each other and humans. These kinds of technologies are producing immense amounts of data, and it becomes critical when analytics and machine learning techniques are applied to make them intelligent with self-organizing capabilities.

2.7. Performance Comparison and Evaluation. Many standards are given in literature to test the efficiency, modality, or validity of any new optimization algorithm. After implementation of the CSFE algorithm, results are analyzed through global optimization test functions.

2.7.1. Rastrigin's Function. Rastrigin's function is a nonlinear optimization function introduced by Rastrigin as a 2-dimensional and extended by Mühlenbein et al. For a D dimensional space, this function can be illustrated as

$$\text{Ras}\left(V'\right) = Ad + \sum_{j=1}^{m}\left(v^j - A\cos\left(2\pi v^j\right)\right), \tag{16}$$

where $A = 10$ and range for this function are $v^j \in [-5.12, 5.12]$. Figure 12 shows results for Rastrigin's function for the D dimensional space with individual dimensions d_i. As compared to the cuckoo search-based feature extraction, the scattered plots show that the dimension space is still extensive, and each selected dimension is similar to previous selected which is not good to represent the whole data.

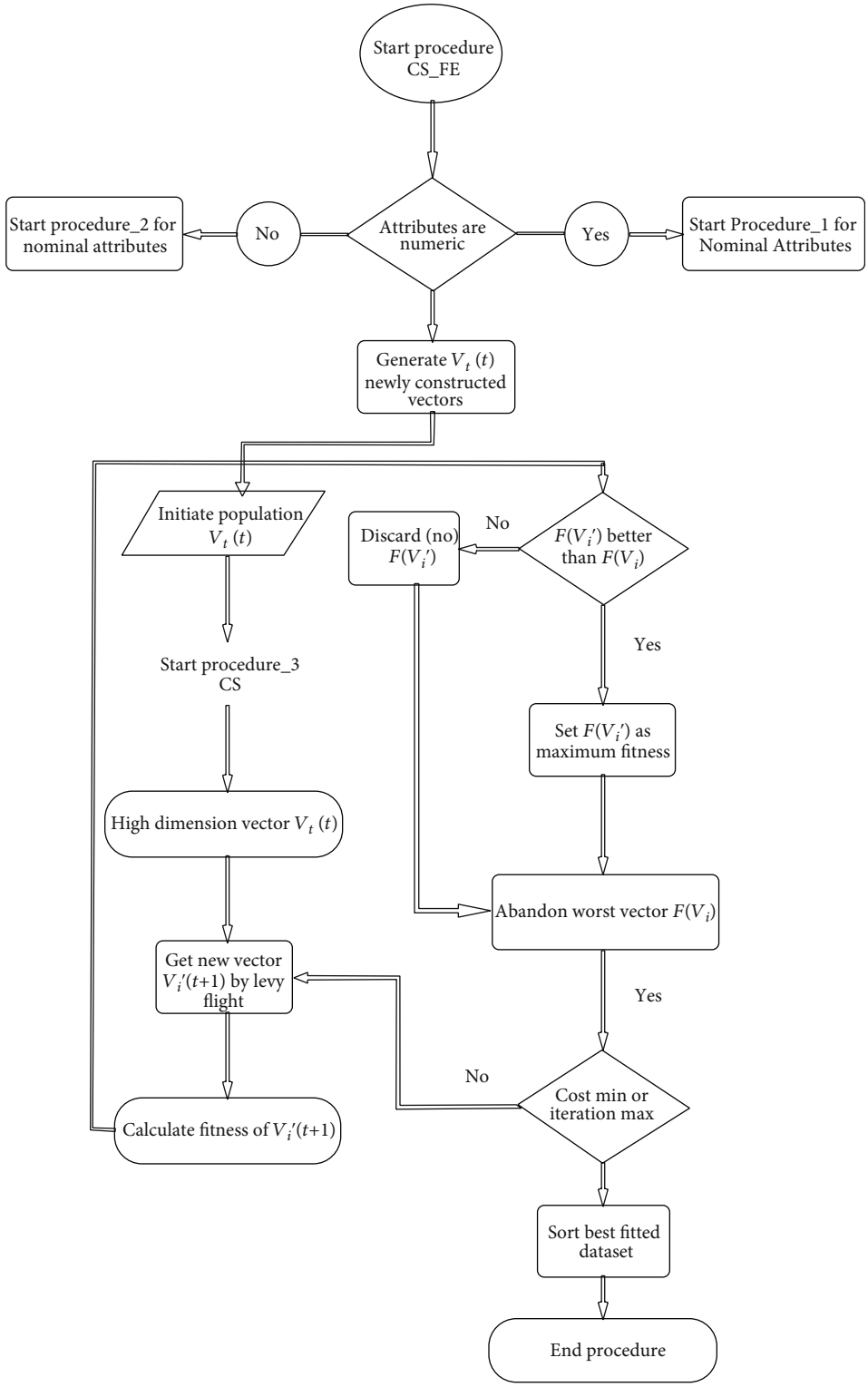


Figure 8: Flowchart for the mutated cuckoo search-based feature extraction including procedures for nominal and numerical attributes while generation of final reduced feature subset through the cuckoo search-based feature extraction.

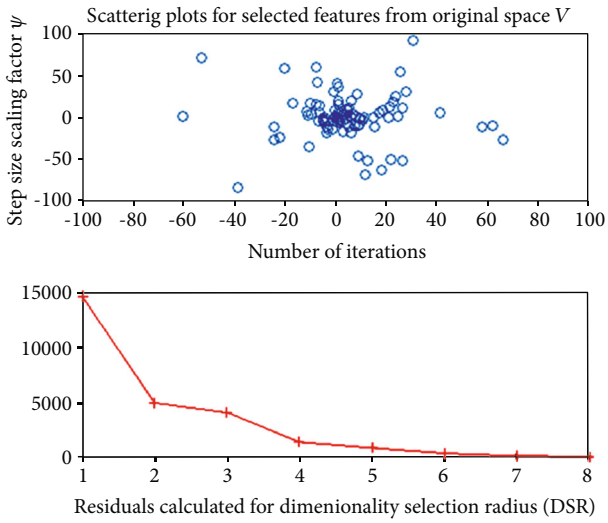


FIGURE 9: Scattering plot for the cuckoo search-based feature extraction, where each point represents the reduced and compressed amount of dimension extracted by CSFE.

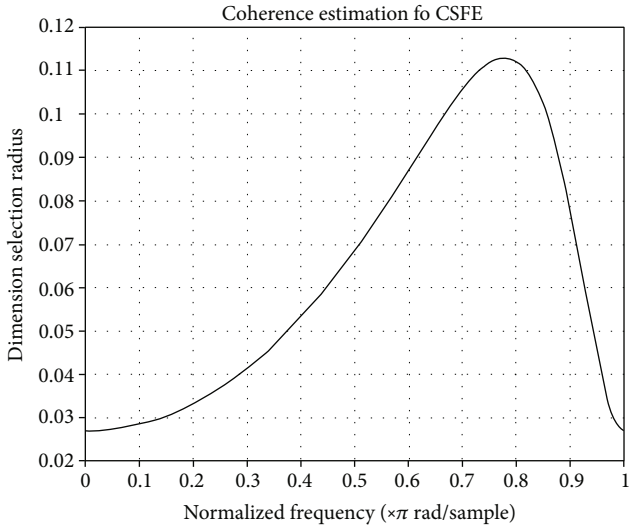


FIGURE 10: Heavy-tailed distribution: representing consistency for the cuckoo search-based feature extraction algorithm, whereas the highest peak level shows that most of the values chosen in each iteration are globally best.

2.7.2. Mccormick's Function. Mccormick's function is a benchmark to test an optimization algorithm. It is defined as given below:

$$\mathrm{Mck}\left(V'\right) = \mathrm{Mck}\left(v_i^j\right) = \sin\left(i+j\right) + \left(i-j\right)^2 - 1.5i + 2.5j + 1, \tag{17}$$

where i and j are vector number and dimension number sequentially. Searching range function is $-1.5 \le i \le 4$ and $-3 \le j \le 4$.

2.7.3. Cross-in-Tray Function. The cross-in-tray function is a continuous and multimodal test standard based on two-dimensional space initially and extended later on. The equation for this function takes the following form:

$$CIT\left(V'\right) = CIT\left(v_i^j\right) - 0.0001\left[\left|\sin\left(i\right)\sin\left(j\right)\exp\cdot\left(\left|100 - \frac{\sqrt{i^2+j^2}}{\pi}\right|\right)\right| + 1\right]^{0.1}, \tag{18}$$

where V' is a D dimensional space and domain range for the cross-in-tray function that is $(i, j)\epsilon[-10, 10]$.

2.7.4. Rosenbrock Function. The Rosenbrock function is a nonlinear benchmark to test the performance of optimization problems, introduced by Howard H. Rosenbrock in 1960. It is also termed as Rosenbrock's Valley or Rosenbrock's banana function.

The mathematical definition for Rosenbrock is mentioned below:

$$\mathrm{Ros}\left(v_i^j\right) = \left(1-i\right)^2 + 100\left(j-i\right)^2, \tag{19}$$

subjected to $(i-1)^3 - j + 1$ and $x + y - 2 < 0$.

Range for the Rosenbrock function is $i \in [-1.5,1.5]$ and $j \in [-0.5,2.5]$. Figure 13 represents the fitness performance for these functions. Local maximum and global minimum for Rosenbrock are shown in Figure 14.

2.7.5. Easom Function. Easom is a multimodal and nonscalable test function to find the global minimum for a search space. It is defined as a following mathematical equation:

$$\mathrm{Easom}\left(v_i^j\right) = -\cos\left(i\right)\cos\left(j\right)\exp\left(-\left(\left(i-\pi\right)^2 + \left(j-\pi\right)^2\right)\right). \tag{20}$$

Search domain for the Eason function is $-100 \le i, j \le 100$. These test functions are used for the comparison of CSFE with particle swarm optimization and harmony search optimization algorithms.

Figure 15 displays the fitness curve for both local and global optimization. For N generations, global fitness is achieved at the early stage that indicates that the running time for the cuckoo search-based feature extraction will be minimal. Figure 16 provides a graph for the best cost value through harmony search (HS). Minimum cost for the firefly algorithm (FFA) is shown in Figure 17. As compared to HS, PSO, and FFA, CSFE gives the minimum cost value in minimum iteration and less elapsed time.

Table 5 provides an overview for the assessment of CSFE against PSO and HS. Performance is measured according to the minimum cost and elapsed time corresponding to each algorithm for maximum generations to calculate the best fitness value (cost). Here, CSFE is compared with few global

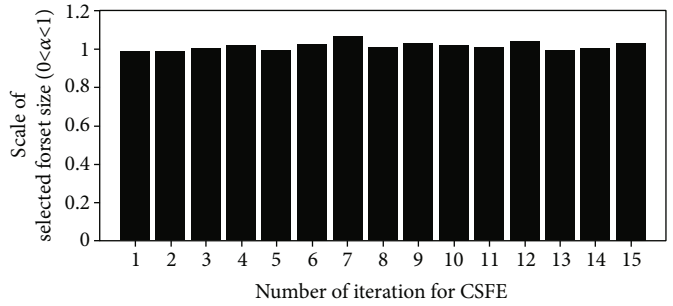


Figure 11: Bar representation for the feature selection during each iteration representing maximum value near to best maximum fitness.

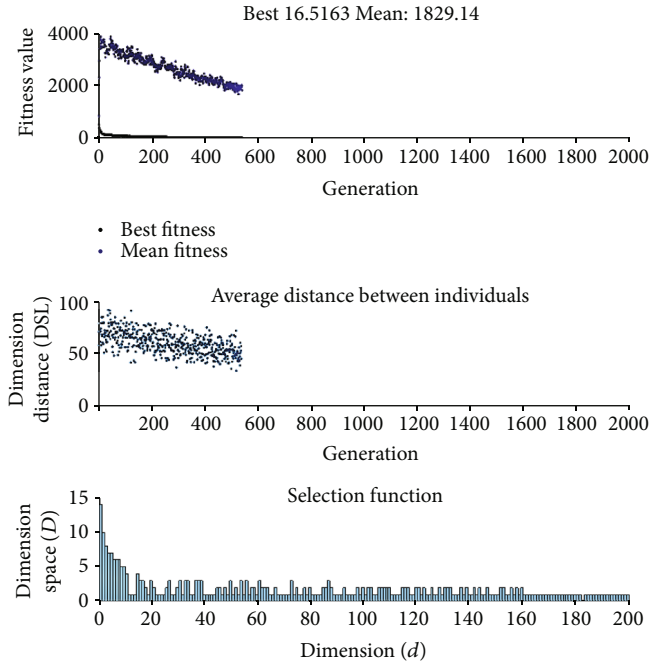


Figure 12: Fitness plot for Rastrigin's function: demonstration of best fitness and mean value for given number of dimensions and selection radius for each individual dimension with selection function for the overall D dimensional space.

optimization techniques. Particle swarm optimization introduced by Kennedy and Eberhart in 1995 provides best mutation results but it is not suitable for complex tasks as it is slow due to its complex structure and mutation.

As compared to CSFE, the peak and tail for the normal distribution are narrow and exponentially bounded from which the fitness values for PSO are generated. It means that local optimization points available through PSO are relatively rare. All global optimization techniques can provide the best possible solution for continuous data generated by the internet of things for a given problem of interest. Algorithms other than CSFE took more time to run and provide minimum cost in more number of iterations.

Another optimization technique is the firefly algorithm introduced by Xin-She Yang in 2008 inspired by the flashing behavior of fireflies. The random numbers for FA are drawn from the uniform distribution $[0, 1]$ with constant probability. Because of constant probability, this optimization technique is not appropriate for continuous and multiobjective optimization problems. In Table 6, PSO and CSFE are compared after 1000 runs for abovementioned test functions. Figure 18 shows the normal distribution $u(0, 1)$ for particle swarm optimization, where maximum height of tail shows that only best values are limited to this small area.

After the evolution of PSO and its limitations, harmony search (HS) was developed by Geem et al. in 2001 based on the concepts of music composition. In Table 7, mutated CSFE is compared with HS. After implementing CSFE and harmony search in MATLAB, the algorithms are compared according to their results generated by almost 1000 iterations.

After comparison, it is analyzed that CSFE has given more accuracy and global fitness for few mentioned test functions. Performance for both algorithms is measured corresponding to each test function as a pair of mean and standard deviation while the accuracy rate is given as

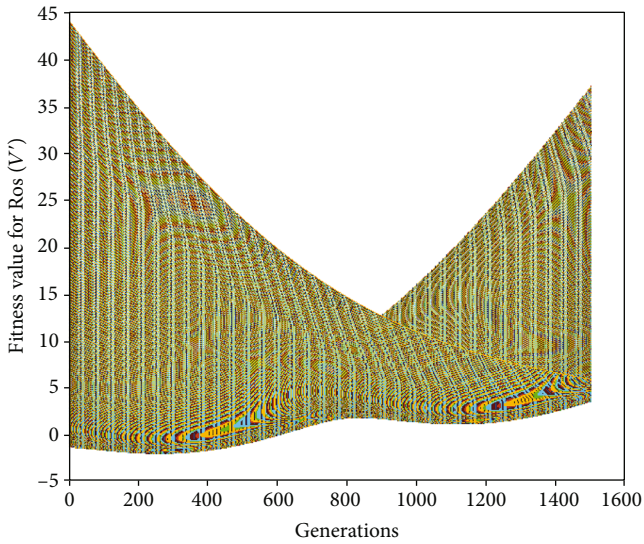


Figure 13: Plot for the Rosenbrock function: as the number of generations increases along the x-axis, the fitness value onto the y-axis is taking the form of local maxima for the given problem.

percentage. CSFE has given more success rate as compared to PSO and HS due to randomization and exploration. CSFE can converge to a global maximum state when required.

In literature, many test benchmarks are introduced to evaluate the performance of any new optimization technique. In this section, few test functions have been used to assess the functionality of the proposed cuckoo search-based feature extraction technique. CSFE is evaluated individually, and performance comparison is established using few global optimization techniques. At last, the accuracy rate delivered by CSFE is more stable and consistent than PSO and HS for global optimization.

2.8. Contribution. The major contribution of this paper is analysis of different aspects regarding dimensionality reduction and discussion of evolutionary approach with a special focus on cuckoo search algorithms. We have given the detailed discussion on existing dimensionality reduction techniques outlining feature selection and extraction. A comparison of heuristic and metaheuristic search methods are described in this paper. Cuckoo search is an optimization technique which is widely used in resource allocations in operations research, and here, we have used it for finding the attributes in data which may be dropped without affecting the meaning and information coherently in the database. The cuckoo living behavior is analyzed with reference to our own problem formulation. This algorithm also possesses some inherent features which limits its working for dimensionality reduction scenario in IoT and discussed in this paper. An algorithm corresponding to the problem of dimensionality reduction in internet of things scenario is formulated to further investigate the performance measures of the cuckoo search optimization algorithm.

We have transformed our problem into IoT vectors having distinct dimensions and explained the feature selection phenomenon. An algorithm is used for feature construction with numeric and nominal attributes. We have specially introduced mutated cuckoo search-based feature extraction algorithms to work with the generated dataset. We have analyzed the result of application of the cuckoo search algorithm on dimensionality reduction and compared its performance. This cuckoo search optimization algorithm proved to be very effective in feature selection and dimensionality reduction techniques and can be used in similar kinds of future applications.

3. Recommendation and Future Work

In information and communication technology, a number of innovative trends have emerged to facilitate humans, businesses, and industries with improved and efficient services. These next generation technologies manifest new challenges and complexities. Homogenize objects, wireless, and sensor networks, addressing schemes, and visualization build a multiplex structure of IoT. Data storage and analytics is one of the most important elements that formulate the network and emphasize dealing with unpredictable amounts of raw data collected by smart objects. Requirement for cost, time, and energy is directly proportional to an incredibly increasing amount of data. In the coming era, scientists are introducing "Green Computing Devices and Networks" with reduced cost, least time, and minimum energy resources.

3.1. Energy Proficient Algorithm for Green Internet of Things (GIoTs). The fundamental aim for IoT is to empower the smart world without greenhouse influences. To interact with real-world objects, these kinds of networks are equipped with numerous sensors, protocols, and communication technologies with high amounts of energy, sufficient cost, and complexity. Efficient algorithms are required for IoT services and applications to reduce the existing greenhouse effects or to build a new with minimum energy consumption. The proposed technique can be used to get maximum accuracy for any machine learning task with minimum cost without complex computation.

Cuckoo search-based techniques are suitable to build future GIoT with maximum accuracy and lower complexity. Cuckoo search is a global optimization technique and provides global maximum solutions for real-time systems (IoT) who generate continuous and high amounts of data with massive dimensionalities. The reason for this recommendation is that CS does not implicate a lot of mathematical computation which will definitely decrease the overall complexity of the system.

3.2. Future Internet of Things for Patient's Monitoring. In medical scenarios, patients are monitored manually, e.g., patient's history, current disease, and their daily health report. Individual files with distinct patient numbers are maintained including some health parameters such as heart rate, blood pressure temperature, and blood samples. These records are assessed by the concerned doctor for further treatment. Instead of all these manual procedures, a smart health monitoring IoT device can perform these actions smartly without extensive human intervention.

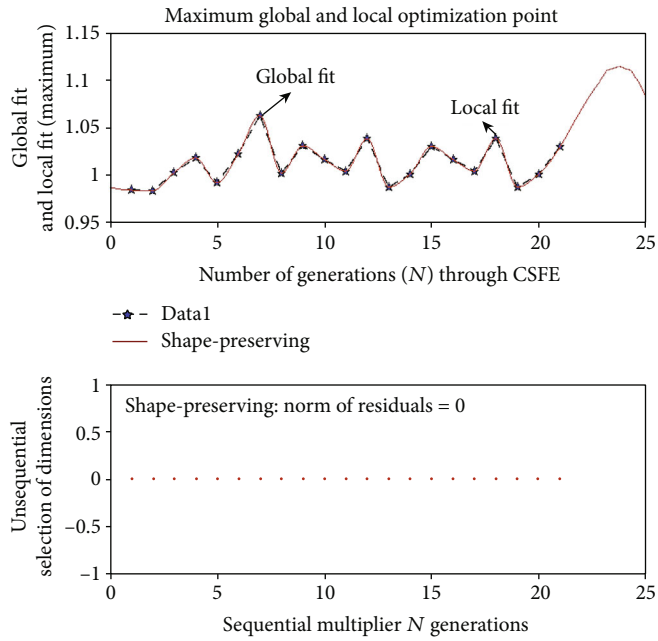


FIGURE 14: Curve fitting plot for the cuckoo search-based feature extraction: indication of global and local maximum for N generations. Subplot provides the nonsequential flow during multiplication of each element at time t.

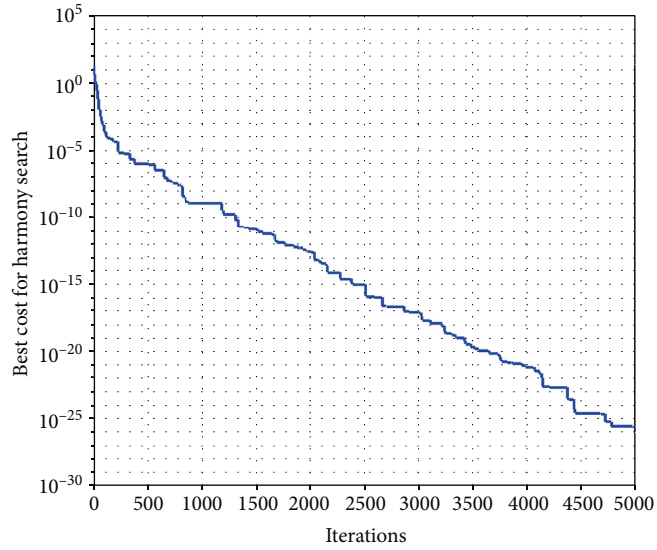


FIGURE 15: Cost estimation for harmony search shows the best cost value for harmony search after 4500 iteration, and it will increase overall algorithm running time.

The architecture in Figure 19 demonstrates an abstract layout for a complete health care system (CHCS) based on an intelligent IoT device to monitor health of patients. The patient's record managed and handled by smart IoT devices is available to concerned doctors and for users as well. A situation handled by an IoT device for patient's monitoring can be a complete blood count (CBC) report of the patient. A continuous blood report will be generated by this intelligent IoT device and whenever there is an alarming situation some action would be triggered.

In spite of tremendous efforts made regarding ICT, there is a need to execute the emerging and evolutionary trends without negative environmental effects to compensate for the increasing amount of data with a smaller amount of energy and computations. Indispensable measurements are required to minimize negative technological effects on the health and society. It is concluded that the mutated cuckoo search-based feature extraction can be an advantageous approach towards the recent internet of things and for future green internet of things as well. Moreover, it can be

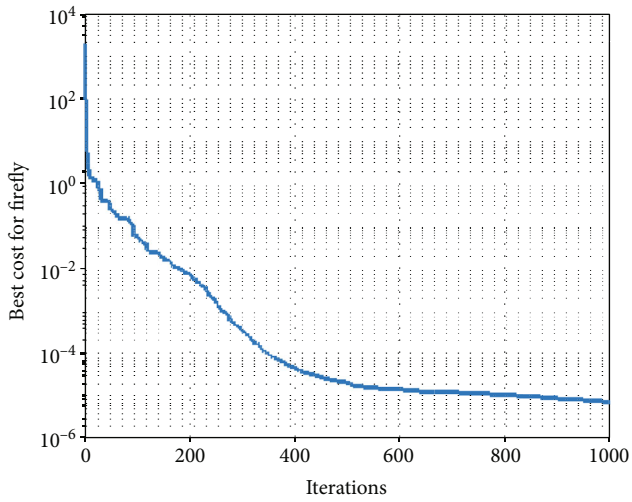


Figure 16: Cost estimation for the firefly algorithm: the best cost value is available after a large number of iterations.

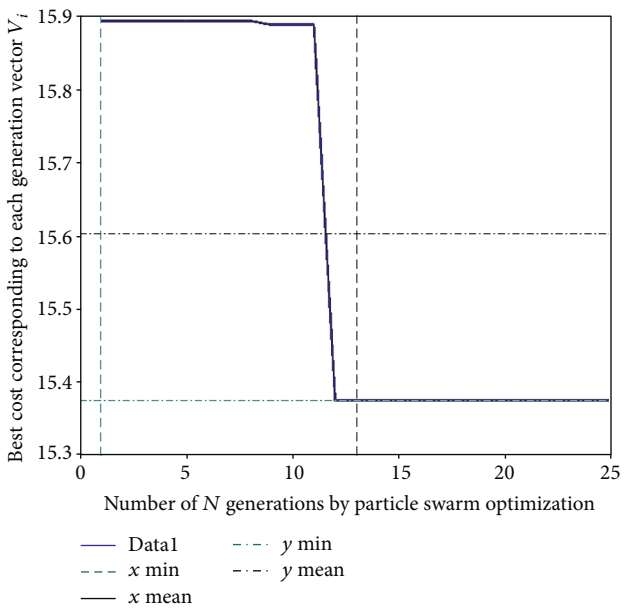


Figure 17: Minimum and maximum cost by particle swarm optimization to generate transformed datasets $\boldsymbol{v}_i^l$.

Table 5: Summarized performance table for HS, PSO, CSFE, and firefly algorithm: analysis of metaheuristic global optimization algorithms with respect to minimum cost and elapsed time.

Algorithm	Best cost value	Elapsed time
Cuckoo search-based feature extraction	4.0788e-26	11.228997 s
Particle swarm optimization	14.375e-4	42.284616 s
Harmony search	7.9050e-06	12.435746 s
Firefly algorithm	7.2753e-06	21.244669 s

Table 6: Comparison of particle swarm optimization and cuckoo search-based feature extraction.

Algorithm	Particle swarm optimization	Cuckoo search-based feature extraction
Rastrigin	4112 ± 279 (93%)	1025 ± 103 (97%)
Mccormick	7877 ± 503 (97%)	3421 ± 209 (100%)
Cross-in-tray	13901 ± 2131 (95%)	4239 ± 276 (100%)
Rosenbrock	90571 ± 3036 (100%)	42330 ± 9211 (100%)
Easom	55491 ± 4023 (97%)	78901 ± 1043 (98%)

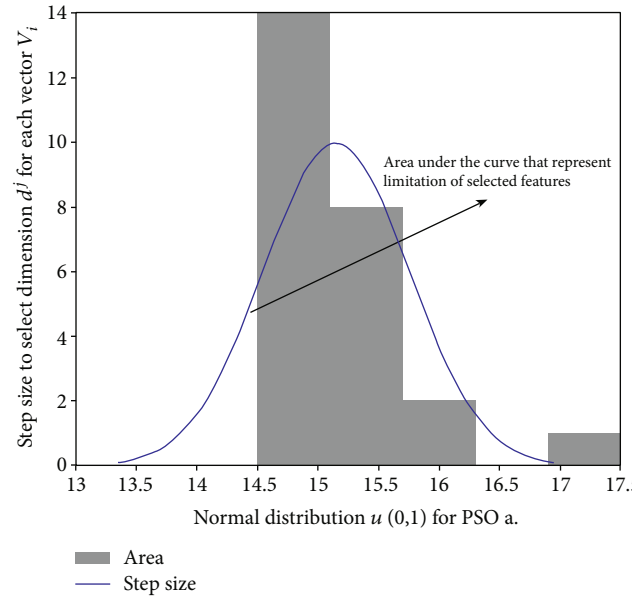


Figure 18: Normal distribution $u(0, 1)$ for particle swarm optimization, where maximum height of tail shows that only best values are limited to this small area.

Table 7: Comparison of harmony search and cuckoo search-based feature extraction.

Test functions	Harmony search	Cuckoo search-based feature extraction
Rastrigin	54412 ± 2301 (94%)	1025 ± 103 (97%)
Mccormick	60925 ± 3324 (87%)	3421 ± 209 (100%)
Cross-in-tray	70215 ± 1051 (100%)	4239 ± 276 (100%)
Rosenbrock	39571 ± 3036 (100%)	42330 ± 9211 (100%)
Easom	10033 ± 4023 (97%)	78901 ± 1043 (98%)

utilized to enhance the performance for future IoT devices for uninterrupted monitoring of patients. As a brief description, cuckoo search optimization is a metaheuristic approach and applicable to situations where the system is in the local state or will grow up towards a global phenomenon in future.

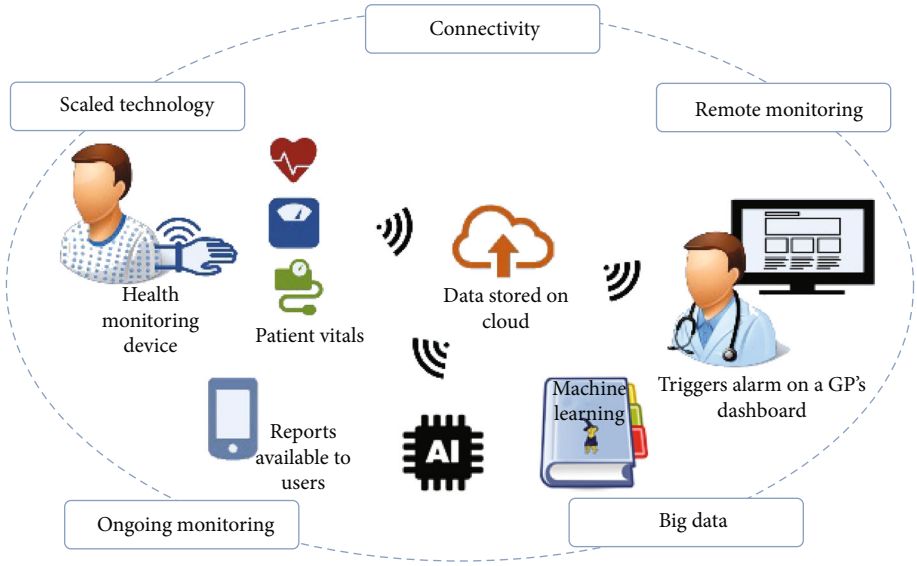


Figure 19: A future IoT for patient's monitoring: an overview for an internet of things-based complete health care system with big data analytics and machine learning techniques.

4. Conclusion

Although IoT networks have emerged and performing exacting tasks competitively but to stable their performance, enhancement for future challenges is mandatory. It becomes possible when networks are not only fast with preeminent servers but also smart enough to cope with unpredictable circumstances. Introducing efficient and global optimization algorithms can help to achieve this target. In this research, a metaheuristic global optimization algorithm is established to reduce dimensions of outdoor data for IoT. The cuckoo search-based feature extraction is a mutated algorithm that organizes itself according to the unpredictable amount of data and produces a new and an enhanced feature space. The newly generated feature space and proposed algorithm benefit for improving the accuracy for classifiers and mining algorithms. It also makes the algorithm computationally feasible, flexible, and efficient for obtaining the target of convergence. This mutated algorithm is further generalizable to IoT indoor activities as the need of near future. The scenario to train the IoT network for future challenges and increase sphere of knowledge is also discussed. Among all abovementioned facts, CSFE can perform all activities with minimum cost and less time as evaluated. This algorithm can be further improved for multiobjective optimization problems. It can provide tremendous support to build an IoT-based smart world with no negative impacts and minimum resources.

Data Availability

Dataset is generated through IoTify. This is a web-based platform for simulation to develop IoT application by using virtual hardware devices, e.g., sensors. It facilitates like a virtual lab and enables to create a virtual IoT device in JavaScript. The IoTify database is generated using JavaScript object naming (JSON) with extension.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] O. Vermesan, P. Friess, P. Guillemin et al., "Internet of things strategic research roadmap," *Internet things-global Technol. Soc. trends*, vol. 1, no. 2011, pp. 9–52, 2011.

[2] E. Borgia, "The Internet of Things vision: Key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.

[3] H. Chiroma, T. Herawan, I. Fister Jr. et al., "Bio-inspired computation: Recent development on the modifications of the cuckoo search algorithm," *Applied Soft Computing*, vol. 61, pp. 149–173, 2017.

[4] O. Vermesan, P. Friess, P. Guillemin et al., *Internet of things strategic research and innovation agenda*, River Publ. Ser. Commun., 2013.

[5] K. K. Patel and S. M. Patel, "Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges," *International Journal of Engineering in Computer Science*, vol. 6, no. 5, 2016.

[6] Y. Yoo, "Computing in Everyday Life: A Call for Research on Experiential Computing," *MIS Quarterly*, vol. 34, no. 2, pp. 213–231, 2010.

[7] S. M. Thampi, O. Marques, S. Krishnan, K.-C. Li, D. Ciuonzo, and M. H. Kolekar, "Advances in signal processing and intelligent recognition systems: 4th International Symposium SIRS 2018, Bangalore, India, September 19-22, 2018," *Revised Selected Papers*, vol. 968, 2019.

[8] W. Fan, W. Lee, S. J. Stolfo, and M. Miller, "A multiple model cost-sensitive approach for intrusion detection," *in European conference on machine learning*, pp. 142–154, Springer, Barcelona, Catalonia, Spain, 2000.

[9] N. Zhang, H. Chen, X. Chen, and J. Chen, "Semantic framework of internet of things for smart cities: Case studies," *Sensors*, vol. 16, no. 9, p. 1501, 2016.

[10] P. Maghouli, S. H. Hosseini, M. O. Buygi, and M. Shahidehpour, "A Multi-Objective Framework for Transmission Expansion Planning in Deregulated Environments," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1051–1061, 2009.

[11] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context Aware Computing for The Internet of Things: A Survey," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 1, pp. 414–454, 2014.

[12] P. Hu, S. Dhelim, H. Ning, and T. Qiu, "Survey on fog computing: architecture, key technologies, applications and open issues," *Journal of Network and Computer Applications*, vol. 98, pp. 27–42, 2017.

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

[14] D. Del Vecchio and D. Carter, "Enhancing presence awareness in instant messaging," in *in Innovations Through Information Technology: 2004 Information Resources Management Association International Conference*, vol. 1, p. 242, New Orleans, Louisiana, USA, 2004.

[15] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, "IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 75–87, 2017.

[16] J. Z. Kolter and M. J. Johnson, "REDD: a public data set for energy disaggregation research," in *in Workshop on data mining applications in sustainability (SIGKDD)*, vol. 25, pp. 59–62, San Diego, CA, 2011.

[17] C.-W. Tsai, C.-F. Lai, H.-C. Chao, and A. V. Vasilakos, "Big data analytics: a survey," *Journal of Big Data*, vol. 2, no. 1, 2015.

[18] J. Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, MIT press, 1998.

[19] R. M. Ramadan and R. F. Abdel-Kader, "Face recognition using particle swarm optimization-based selected features," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no. 2, pp. 51–65, 2009.

[20] J.-Y. Yeh, T.-H. Wu, and C.-W. Tsao, "Using data mining techniques to predict hospitalization of hemodialysis patients," *Decision Support Systems*, vol. 50, no. 2, pp. 439–448, 2011.

[21] C. A. Bhatt and M. S. Kankanhalli, "Multimedia data mining: state of the art and challenges," *Multimedia Tools and Applications*, vol. 51, no. 1, pp. 35–76, 2011.

[22] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 491–502, 2005.

[23] P. Ristoski and H. Paulheim, "Semantic Web in data mining and knowledge discovery: A comprehensive survey," *Journal of Web Semantics*, vol. 36, pp. 1–22, 2016.

[24] X. Xu, L. He, H. Lu, L. Gao, and Y. Ji, "Deep adversarial metric learning for cross-modal retrieval," *World Wide Web*, vol. 22, no. 2, pp. 657–672, 2019.

[25] S. Ramírez-Gallego, B. Krawczyk, S. García, M. Woźniak, and F. Herrera, "A survey on data preprocessing for data stream mining: current status and future directions," *Neurocomputing*, vol. 239, pp. 39–57, 2017.

[26] B. Rasti, D. Hong, R. Hang et al., "Feature extraction for hyperspectral imagery: the evolution from shallow to deep," 2020, http://arxiv.org/abs/2003.02822.

[27] D. Hong, N. Yokoya, and X. X. Zhu, "Learning a robust local manifold representation for hyperspectral dimensionality reduction," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2960–2975, 2017.

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

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

[30] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in internet of things: towards secure and privacy-preserving fusion," *Information Fusion*, vol. 51, pp. 129–144, 2019.

[31] B. Gendron and T. G. Crainic, "Parallel branch-and-branch algorithms: survey and synthesis," *Operations Research*, vol. 42, no. 6, pp. 1042–1066, 1994.

[32] L. Huang, S. Ding, S. Yu, J. Wang, and K. Lu, "Chaos-enhanced cuckoo search optimization algorithms for global optimization," *Applied Mathematical Modelling*, vol. 40, no. 5–6, pp. 3860–3875, 2016.

[33] B. He, S. Shah, C. Maung, G. Arnold, G. Wan, and H. Schweitzer, "Heuristic search algorithm for dimensionality reduction optimally combining feature selection and feature extraction," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 2280–2287, 2019.

[34] B. P. De, R. Kar, D. Mandal, and S. P. Ghoshal, "Optimal selection of components value for analog active filter design using simplex particle swarm optimization," *International Journal of Machine Learning and Cybernetics*, vol. 6, no. 4, pp. 621–636, 2015.

[35] S. M. Piryonesi and M. Tavakolan, "A mathematical programming model for solving cost-safety optimization (CSO) problems in the maintenance of structures," *KSCE Journal of Civil Engineering*, vol. 21, no. 6, pp. 2226–2234, 2017.

[36] X. Zhou, C. Yang, and W. Gui, "Nonlinear system identification and control using state transition algorithm," *Applied Mathematics and Computation*, vol. 226, pp. 169–179, 2014.

[37] A. R. Malisia and H. R. Tizhoosh, "Applying Opposition-Based Ideas to the Ant Colony System," in *in 2007 IEEE Swarm Intelligence Symposium*, pp. 182–189, Honolulu, Hawaii, USA, 2007.

[38] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *in 2007 IEEE Swarm Intelligence Symposium*, pp. 120–127, Honolulu, Hawaii, USA, 2007.

[39] X. Zhou, D. Y. Gao, and C. Yang, "A Comparative Study of State Transition Algorithm with Harmony Search and Artificial Bee Colony," in *Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013*, pp. 651–659, China, 2013.

[40] X.-y. YANG and W. U. Dan, "Atomic simulations for surface-initiated melting of Nb(111)," *Transactions of the Nonferrous Metals Society of China*, vol. 19, no. 1, pp. 210–214, 2009.

[41] X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, and M. Karamanoglu, *Swarm Intelligence and Bio-Inspired Computation*, Newnes, 2013.

[42] R. B. Payne, "The Ecology of Brood Parasitism in Birds," *Annual Review of Ecology and Systematics*, vol. 8, no. 1, pp. 1–28, 1977.

[43] H. Rathore, *Mapping Biological Systems to Network Systems*, Springer, 2016.

[44] N. Davies, *Cuckoos, Cowbirds and Other Cheats*, A&C Black, 2010.

[45] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems," *Engineering with Computers*, vol. 29, no. 1, pp. 17–35, 2013.

[46] K. Maulik and B. Zwart, "Tail asymptotics for exponential functionals of Lévy processes," *Stochastic Processes and their Applications*, vol. 116, no. 2, pp. 156–177, 2006.

[47] S. Fong, "Opportunities and challenges of integrating bio-inspired optimization and data mining algorithms," in *in Swarm Intelligence and Bio-Inspired Computation*, pp. 385–402, Elsevier, 2013.

[48] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Communication Surveys and Tutorials*, vol. 19, no. 4, pp. 2392–2431, 2017.

[49] A. B. Nasser, A. R. A. Alsewari, and K. Z. Zamli, "Tuning of cuckoo search based strategy for t-way testing," *in International conference on electrical and electronic engineering*, vol. 9, p. 10, 2015.

[50] X.-S. Yang, S. F. Chien, and T. O. Ting, "Computational Intelligence and Metaheuristic Algorithms with Applications," *The Scientific World Journal*, vol. 2014, Article ID 425853, 4 pages, 2014.

[51] G. Wang, "A comparative study of cuckoo algorithm and ant colony algorithm in optimal path problems," *MATEC Web of Conferences*, vol. 232, p. 03003, 2018.

[52] V. Mayer-Schonberger and K. Cukier, *Big Data: The Essential Guide to Work, Life and Learning in the Age of Insight*, Hachette UK, 2013.

[53] P. Raj, T. Poongodi, B. Balusamy, and M. Khari, *The Internet of Things and Big Data Analytics: Integrated Platforms and Industry Use Cases*, CRC Press, 2020.

[54] P. Tan, X. Wang, and Y. Wang, "Dimensionality reduction in evolutionary algorithms-based feature selection for motor imagery brain-computer interface," *Swarm and Evolutionary Computation*, vol. 52, p. 100597, 2020.

[55] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: a survey," *Digital Communications and Networks*, vol. 4, no. 3, pp. 161–175, 2018.

[56] S. Pittner and S. V. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83–88, 1999.

[57] M. R. Genesereth and N. J. Nilsson, *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, 2012.

[58] D. Mladenić, "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, 2005.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8832191, 18 pages
https://doi.org/10.1155/2020/8832191

Research Article

A Comparative Analysis of Different Outlier Detection Techniques in Cognitive Radio Networks with Malicious Users

Arshed Ahmed,[1] Muhammad Sajjad Khan,[2,3] Noor Gul,[1,3] Irfan Uddin,[4] Su Min Kim,[2] and Junsu Kim[2]

[1]*Department of Electronics, University of Peshawar, Pakistan*
[2]*Department of Electronic Engineering, Korea Polytechnic University, Republic of Korea*
[3]*Department of Electrical Engineering, International Islamic University, Islamabad, Pakistan*
[4]*Department of Information Technology, Superior University, Lahore, Pakistan*

Correspondence should be addressed to Junsu Kim; junsukim@kpu.ac.kr

Received 4 August 2020; Revised 12 November 2020; Accepted 22 November 2020; Published 9 December 2020

Academic Editor: Fawad Zaman

In a cognitive radio (CR), opportunistic secondary users (SUs) periodically sense the primary user's (PU's) existence in the network. Spectrum sensing of a single SU is not precise due to wireless channels and hidden terminal issues. One promising solution is cooperative spectrum sensing (CSS) that allows multiple SUs' cooperation to sense the PU's activity. In CSS, the misdetection of the PU signal by the SU causes system inefficiency that increases the interference to the system. This paper introduces a new category of a malicious user (MU), i.e., a lazy malicious user (LMU) with two operating modes such as an awakened mode and sleeping mode. In the awakened mode, the LMU reports accurately the PU activity like other normal cooperative users, while in the sleeping mode, it randomly reports abnormal sensing data similar to an always yes malicious user (AYMU) or always no malicious user (ANMU). In this paper, statistical analysis is carried out to detect the behavior of different abnormal users and mitigate their harmful effects. Results are collected for the different hard combination schemes in the presence of the LMU and opposite categories of malicious users (OMUs). Simulation results collected for the error probability, detection probability, and false alarm at different levels of the signal-to-noise ratios (SNRs) and various contributions of the LMUs and OMUs confirmed that out of the many outlier detection tests, the median test performs better in MU detection by producing minimum error probability results in the CSS. The results are further compared by keeping minimum SNR values with the mean test, quartile test, Grubbs test, and generalized extreme studentized deviate (GESD) test. Similarly, performance gain of the median test is examined further separately in the AND, OR, and voting schemes that show minimum error probability results of the proposed test as compared with all other outlier detection tests in discarding abnormal sensing reports.

1. Introduction

Radio spectrum is considered the backbone for wireless communication. The unique characteristic of the wireless sensor networks (WSNs) makes it distinguishable from the traditional networks [1]. In WSNs, a number of small sensor devices distributed spatially are allowed to cooperatively sense environmental and physical conditions. The WSN nodes have limited resources in terms of power, computational complexity, and memory [2]. Recently, the WSNs are employed in civilian applications, such as home appliance control, traffic control, checking environmental conditions, Internet of things (IoT), and robotic games [2]. The frequency spectrum assigned to the WSNs and other communication devices is not efficiently utilized that results in spectrum scarceness issues. The CR network (CRN) is a promising technology in the field of WSNs to tackle the spectrum scarcity [3].

The idea of CRN was presented for the first time by Mitola in [4]. As demand to the frequency spectrum resources is increasing with the increased number of wireless devices, therefore, static spectrum allocation (SSA) policy is considered to have limitations to meet these requirements [5]. The 300 GHz bandwidth that once seems to be sufficient

is now becoming congested [6–9]. CRN is an intelligent wireless communication technology that has the ability to sense the radio environment and act accordingly. The CRN has two main objectives: reliable communication at any time and place and efficient use of the radio spectrum [10]. As static spectrum allocation is not the solution to meet with the increasing number of wireless communication devices, therefore to overcome this challenging problem, dynamic spectrum access (DSA) has been widely proposed as one of the most promising technologies to increase spectral efficiency [11]. The CRN is considered a feasible intelligent technology for 4G wireless networks or self-organization networks. In the CRN, unlicensed users or secondary users (SUs) periodically sense the spectrum band of the PU network. The SUs utilize vacant channels in the VHF and UHF frequency bands, allocated to TV broadcasting between 54 and 862 MHz frequency range [12]. The PU spectrum availability is inspected by applying various spectrum sensing techniques [6]. The SU performs local sensing by adopting sensing techniques such as energy detector (ED), matched filter detector (MFD), and cyclostationary [7]. When statistics of the PU are not available, then the ED technique is more suitable that requires only power of the PU channel. The received energy of PU is compared with a fixed threshold value in the ED technique. In case the received energy is greater than the threshold, the presence of PU is confirmed; otherwise, the absence of the PU signal is declared [7, 8]. In the proposed work, we will follow the ED technique to sense the spectrum of the PU channel.

In CRN, individual SU is not sensitive enough to detect PU channel weak signals. The single SU sensing performance is further deteriorated by the multipath fading and shadowing effects as in [13]. In order to tackle individual user sensing issues, CSS is used to solve this problem. This allows local sensing users to forward their sensing results to the fusion center (FC), where the final decision is made about the PU status [14].

1.1. Related Work and Contribution. Information reported to the FC by the SUs through local sensing is divided into two major categories: hard decision fusion (HDF) and soft decision fusion (SDF). In the HDF, the SUs convert the sensing reports into binary decision to represent a PU signal. The HDF schemes not only reduce communication cost but also reduce the implementation complexity [15]. In the SDF scheme, the reports are in the form of energy values of the PU signal forwarded to the FC. There are many SDF schemes suggested in the literature, where soft energy information is reported to the FC [16]. Similarly, in the Bayesian model, users report probabilities to represent the confidence level of the users' local decision [17]. The FC then takes a global decision by combining all these probabilities. An SDF model proposed in [17] reports two-bit information to state the free and occupied status of the PU channel. The SDF scheme known as the likelihood ratio test (LRT) has attained a significant attention. In [12], a linear test statistic is applied based on an LRT detector at several PU conditions. In [18], the focus is on maximum eigenvalue-based LRT against different noise behaviors of the PU signal. In [19], the authors have investigated a distributed LRT detector for sensing the spectrum of the PU spectrum where the channels are considered having random and Nakagami-lognormal mixture distribution. Similarly, in [20], some inspections against frequency-selective Nakagami channels using correlation in the frequency domain are investigated. In [21], authors have presented a collusion pattern of the attackers. These attackers usually form a collusive group that can boost the spectrum sensing data falsification (SSDF) attack power, resulting in falsification of the spectrum sensing data. These attackers are prevented by applying a trust mechanism technique, in which the reports of the SUs are examined by their historical sensing behaviors [22]. The less trusted SUs are given low weights, or even their reports are deleted during final decision. The collusive attackers are riskier as they improve their trust value, which results in increasing their attack power. The main contributions of this paper are as follows.

(i) In this paper, a new behavior of MU, i.e., a lazy malicious user (LMU), is introduced in the CSS environment. The LMU reports PU information to the FC in two operating modes, i.e., an awakened phase and sleeping phase. The user acts as normal SU during the awakened phase with accurate sensing reports in this phase, while in the sleeping phase, the LU acts maliciously by reporting false sensing data randomly selected as AYMU and ANMU probabilistically. The OMU category of MUs senses the PU channel and repots sensing data to the FC that negate the channel actual status

(ii) The proposed techniques in the paper detect LMUs and OMUs by applying outlier detection tests while reporting to the FC. During the sleeping phase of the LMUs, received sensing reports are detected as abnormal and discarded while making global decision at the FC. Similarly, as the awakened phase sensing reports of the LMUs are accurate, therefore, the outlier detection tests declare their sensing reports as normal and suggest for consideration in the global decision

(iii) Simulation verifies that, out of the many outlier detection tests, the median test shows better detection results of MUs in CSS and produces minimum error probability. The results are further compared at low SNR values with those of the mean test, quartile test, Grubbs test, and generalized extreme studentized deviate (GESD) test

The proposed work limitation lies in the parameter selection of statistical tests. It is noticeable that whenever univariate data samples are selected less than a certain limit, outlier values near the upper and lower fence of the data distribution cannot be detected reliably. Hence, the number of SUs should be sufficient enough to get better sensing results.

The rest of the paper is organized as follows: Section 2 presents the system model. Section 3 gives a detailed description of the proposed MU detection model. Section 4 discusses the simulation results. The paper is concluded in Section 5.

2. System Model

All the participating SUs sense the PU status and report their decisions to the FC. The SUs decide the PU activity locally and inform FC about their binary decision findings for making a global decision. FC collects individual hard binary decisions of the cooperative users and employs HDF schemes to recommend the final decision about licensed user activity as shown in Figure 1.

The binary hypothesis about the presence and the absence of the PU channel is given as

$$x_j(l) = \begin{Bmatrix} H_0: & n_j(l) \\ H_1: & h_j s(l) + n_j(l) \end{Bmatrix}, \tag{1}$$

where $x_j(l)$ denotes the energy received by the j^{th} SU in the l^{th} time slot. H_0 and H_1 represent the absence and presence hypothesis of the PU signal. $n_j(l)$ is the AWGN and h_j is the channel gain between the PU and the j^{th} SU. $s(l)$ is the PU transmission at the l^{th} time slot [23, 24]. The energy statistic of the PU received by the j^{th} SU in the i^{th} time interval is given as

$$W_j(i) = \begin{Bmatrix} \sum_{i=l_i}^{l_i+(b-1)} \left|n_j(l)\right|^2, & H_0 \\ \sum_{i=l_i}^{l_i+(b-1)} \left|h_j s(l) + n_j(l)\right|^2, & H_1 \end{Bmatrix}, \tag{2}$$

In (2), b is the number of samples at the i^{th} time interval. The central limit theorem (CLT) shows that for binary hypothesis and large sample size, the energy reported by the participating SUs resembles Gaussian random variables. The normalized energy is written as

$$W_j \sim \begin{Bmatrix} N\left(\mu_0 = b, \sigma_0^2 = 2b\right), & H_0 \\ N\left(\mu_1 = b\left(n_j + 1\right), \sigma_1^2 = 2b\left(n_j + 1\right)\right), & H_1 \end{Bmatrix}. \tag{3}$$

In (3), n_j represents the noise received by the j^{th} SU. The mean and variance of the energy statistics are μ_0 and σ_0^2, respectively, for the hypothesis H_0. Similarly, for H_1, the mean and variance of the energy statistics are μ_1 and σ_1^2. The energy statistics collected at each SU locally decide the existence of the PU status. These statistics are further compared with the predefined threshold value to send the hard decisions 1 or 0 to the FC [15] as

$$Z_j(i) = \begin{Bmatrix} 1, & W_j(i) \geq \gamma_j \\ 0, & \text{otherwise} \end{Bmatrix}, \tag{4}$$

where $W_j(i)$ is the energy statistic of the PU received by the j^{th} SU in the i^{th} interval. γ_j denotes the threshold value for the j^{th} reporting user.

2.1. Proposed MU Detection Model. A flow chart of the proposed CSS model is shown in Figure 2, where multiple SUs sense a spectrum band of the PU and report their observations to the FC. In the flow chart, simple AND, OR, and majority voting are the schemes where outlier tests are not applied and reports are collected from all SUs about PU activity. Similarly, the modified AND, OR, and majority voting are those schemes where outlier tests were used for the MU identification based on all users' reported information. The global decision is calculated under both simple HDF and modified HDF schemes separately, and results are compared.

Pseudocode 1 of the proposed Algorithm 1 is shown in Section 3.

3. Pseudocode 1 of Algorithm 1

A pseudocode of the proposed algorithm to solve the given problem in a stepwise manner is shown. Here, the users take their hard binary decisions and report the same information as 1 or 0 to the FC. FC tries to collect and stores user reports during the N sensing intervals and stores the same in its local database in Z. The FC takes its final decision normally using hard decision schemes before collection of enough reports from the sensing users. At the end of a required number of iterations, the results in Z are accumulated by finding each user total sensing data to form vector $\mathbf{z}$. The algorithm calls statistical outlier detection tests to detect any abnormality in $\mathbf{z}$ results as outlier or malicious data. After the identification of MUs, modified HDF schemes are allowed to take decision based on the sensing reports of the normally declared users in the subsequent sensing intervals.

3.1. Hard Decision Schemes. A centralized CSS allows SUs to forward their local sensing results to a central unit where the final decision of the PU activity is made based on sensing reports. To categorize the information provided to the FC, local sensing schemes are divided into HDF and SDF. In HDF, the SUs convert the sensing reports into binary digits 1 and 0 that represent the PU signal. HDF schemes reduce both the communication cost and implementation complexity of the system. In the SDF scheme, reports are in the form of energy values of the PU signal that are forwarded to the FC. LRT has attained a significant attention out of the different SDF schemes.

3.1.1. AND Scheme. In the AND scheme, all SUs have to be consistent about the reports of PU:

$$G_{\text{d}} = \begin{Bmatrix} H_1: & \sum_{j=1}^{n} Z_j(i) = n \\ H_0: & \text{otherwise} \end{Bmatrix}, \tag{5}$$

$Z_j(i)$ consists of reports in the i^{th} interval by the SUs. The channel is declared occupied when all SUs reports the PU availability where H_1 is generated by the FC as a global decision G_{d}; otherwise, decision H_0 is declared.

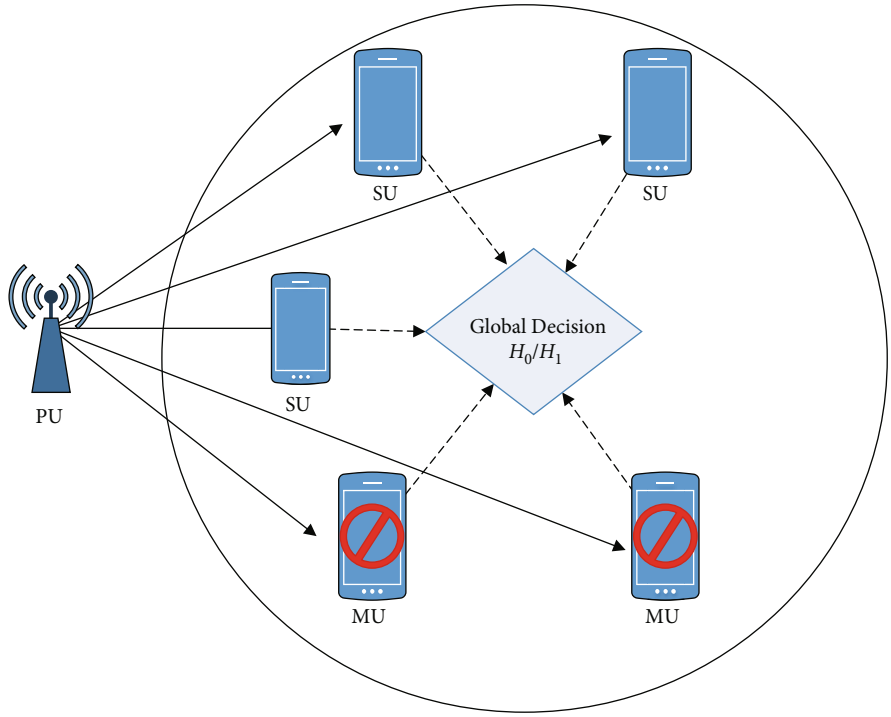


Figure 1: System model.

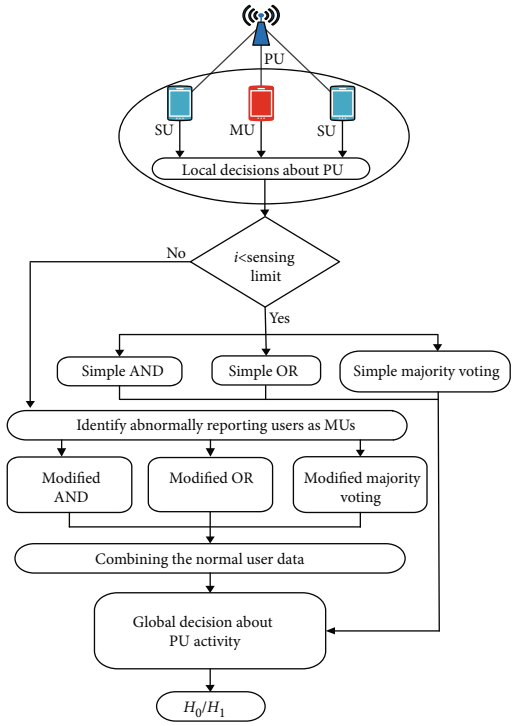


Figure 2: Flow chart for the proposed methodology.

3.1.2. OR Scheme. In the OR scheme, if any SU detects the PU signal, then FC takes it as a global decision and generates H_1; otherwise, the global decision is H_0:

$$G_{\mathrm{d}} = \left\{ \begin{array}{ll} H_1: & \sum_{j=1}^{n} Z_j(i) \geq 1 \\ H_0: & \text{otherwise} \end{array} \right\}. \tag{6}$$

3.1.3. Majority Voting Scheme. The majority voting scheme is based on the voting of SUs. If majority users declare the PU availability, the decision is made in favor of majority voters:

$$G_{\mathrm{d}} = \left\{ \begin{array}{ll} H_1: & \sum_{j=1}^{n} Z_j(i) \geq k \\ H_0: & \text{otherwise} \end{array} \right\}, \tag{7}$$

where k is the number of SUs, declaring that PU has occupied the channel, and n is the total number of participating SUs. The majority voting scheme is the special case of global decision when $k = n/2$. The FC applies statistical analysis by combining the reports of all participating users to remove the nasty data from MUs in the local sensing.

3.2. Statistical Outlier Tests. The outliers in the data are dissimilar values to the rest of the data set. They are generated through different mechanisms in the CSS [25]. The outliers can also be defined as those observations that deviate from their members in the data sample [26]. In this work, the reports of the MUs are outliers because from the definition, outliers are the data samples generated by another

(1) For i= 1 to total iterations
(2) For$j$=1 to total SUs
(3) if PU is available
(4) j^{th} user sensing in i^{th} interval $E(i, j)$
(5) if $E_{(i,j)}$> threshold
(6) local decision $Z(i, j) = 1$ (Reporting 1 as hard decision),
(7) else
(8) local decision $Z(i, j) = 0$ (Reporting 0 as hard decision)
(9) end
(10) end
(11) end
(12) if $i <$ Sensing limit
(13) Compile the results for simple HDF schemes
(14) Simple AND
(15) Simple OR
(16) Simple Majority voting
(17) else
(18) Identify**z**
(19) Run Outlier Tests using **z**to detect outliers
(20) Compile results for Modified HDF schemes
(21) end
(22) if $\sum_{\substack{j=1 \\ j\neq MU}}^{n} Z_j(i) = n$
(23) AND global Decision as H_1
(24) else
(25) AND global Decision as H_0
(26) end
(27) Compile results for Modified HDF schemes if $\sum_{\substack{j=1 \\ j\neq MU}}^{n} Z_j(i) \geq 1$
(28) OR global Decision as H_1
(29) else
(30) OR global Decision as H_0
(31) end
(32) if $\sum_{\substack{j=1 \\ j\neq MU}}^{n} Z_j(i) \geq n/2$
(33) Voting global Decision as H_1
(34) else
(35) Voting global Decision as H_0
(36) end
(37) end of iterations

Pseudocode 1

mechanism; hence, the reports from the MUs deviate from those reports which are generated by normal SUs [27].

In the proposed model, SUs sense the PU channel and report their hard binary findings to the FC, where it stores n SU sensing data reported in N sensing iterations to form matrix $\mathbf{Z}$ as shown in

$$\mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \cdots & : \\ \vdots & \vdots & \cdots & : \\ z_{N1} & z_{N2} & \cdots & z_{Nn} \end{bmatrix}. \quad (8)$$

At the end of the required number of iterations, each user contribution in sensing is determined by adding total hard decisions of the SUs to form vector $\mathbf{z}$ as given in

$$\mathbf{Z} = \sum_{i=1}^{N} \left(Z_j(i)\right), \quad i \in 1, \cdots, N. \quad (9)$$

Outlier detection techniques are called by giving the result in equation (9) as an argument to declare the users as normal or abnormal using various detection tests:

$$\mathbf{Z} = [z_1 \quad z_2 \quad \cdots \quad z_n], \quad (10)$$

Finally, the detected outlier is declared as malicious and

taken out of the hard combination scheme in the following sensing intervals.

3.2.1. Proposed Outlier Median Test Scheme. This outlier detection scheme searches for anomaly in the normally distributed sensing data as in [28]. In the case of univariate data, the median absolute deviation (MAD) is the robust dispersion measure against outliers [28]. Therefore, outlier presence in the data needs to be properly detected and removed. Automatic analysis for the detection of these anomalies in the normally distributed data is mandatory. The traditional method of the mean plus-minus 3 test based on the standard deviation of the data follows normal distribution of the data, where 99.87% of the data type occurs within this range. Similarly, taking decision to remove the values occurring in 0.13% of all cases is not too conservative [28]. There are three problems when the mean is considered the central tendency in the data set. First, the data set has to be normally distributed when outliers are included. Secondly, the outliers in the data have a strong impact on the mean and standard deviation. At last, for any small data sample values, the outlier detection is not guaranteed. Due to these drawbacks, the mean test failed to detect outliers in data distributions when the data sample is limited in size.

Therefore, Miller proposed an outlier indication test using the median of the data set. This outlier test detects anomaly for the value of c: it is most conservative when c is 3, medium conservative when c is 2.5, and less conservative when c is 2 [29]. The constant value of 3 is used in this work. The limiting point against the users' total sensing reports in the $\mathbf{z}$ vector is determined in

$$M - (c \times \mathrm{MAD}) < \mathbf{z}_i < M + (c \times \mathrm{MAD}). \tag{11}$$

The result in (11) is written in a more simplified form as

$$\frac{\mathbf{z}_i - M}{\mathrm{MAD}} \geq |\pm 3|. \tag{12}$$

The results of the median test are compared further with those of other outlier tests such as the Grubbs test, GESD test, and quartile test such as the box and whisker plot and mean plus-minus 3 test.

3.2.2. Grubbs Test. Frank Grubbs in 1969 proposed an outlier test to verify some univariate data [30]. The Grubbs test is used to detect a single outlier in sampled data. This test analyzes the minimum/maximum values of the sample data and applies statistics to search outliers. The test statistic is

$$G = \frac{|\max \text{ value} - M|}{\sigma}, \tag{13}$$

where M is the sample mean and σ is the standard deviation given by

$$\sigma = \sqrt{\frac{\sum (\mathbf{z}_i - \bar{\mathbf{z}})^2}{S - 1}}. \tag{14}$$

In (14), S is the number of data values, $\mathbf{z}_i$ is the maximum value of the row vector, and $\bar{\mathbf{z}}$ is the mean value of the vector $\mathbf{z}$. The following steps are used in the Grubs test to detect suspicious report as an outlier.

(i) Find the G test statistics using (13) for the users' sensing reports in $\mathbf{z}$

(ii) State the null and alternative hypotheses about the existence of outliers in $\mathbf{z}$

(iii) Find the G critical value from the table and select the confidence level. The default confidence level is 95%

(iv) Compare the tested G statistic with the G critical value

(v) The maximum value in $\mathbf{z}$ is an outlier if the test statistics are greater than the critical value

The Grubbs tests can be used to detect and remove outlier values from the minimum values of the data sample as given in

$$G = \frac{|M - \min \text{ value}|}{\sigma}. \tag{15}$$

3.2.3. Generalized Extreme Studentized Deviate (GESD) Test. GESD is an iterative hypothesis test proposed by Rosner in 1983. It can spot one or more outliers in a data set. In this test, the upper bound or the total number of outlier values is given in the null hypothesis. After that, a separate test is performed by using the Grubbs statistics as given in [31]

$$T_k = \frac{\max |\mathbf{z}_i - M|}{\sigma}, \tag{16}$$

where M and σ denote the mean and standard deviations in the data. The observation corresponding to $\max |\mathbf{z}_i - M|$ is removed using Grubbs statistics, and T_2 is computed from the remaining sample. A sample mean and standard deviation are computed for the remaining $n - 1$ data values. This process is repeated until T_k is determined for a prespecified k. Here, k represents the number of outliers in the data set known as the upper bound specified in the null hypothesis [32].

3.2.4. Mean Test. This method is based on the characteristics of normal distribution of data. It is necessary for the outlier test to detect the presence of the outlier's data. In [33], the mean plus-minus 3 standard deviation scheme is formulated as

$$\bar{\mathbf{z}} - (a \times \sigma) < \mathbf{z}_i < \bar{\mathbf{z}} + (a \times \sigma), \tag{17}$$

where $\bar{\mathbf{z}}$ denotes the sample mean and σ denotes the standard deviation of z. The constant parameter a is carefully selected which is 3 here to produce accurate results. The value $\mathbf{z}_i$ is an outlier in the z if it exceeds the upper boundary of the data sample such as $\mathbf{z}_i < \bar{\mathbf{z}} + (a \times \sigma)$ or if the data value exceeds the lower boundary, i.e., $\bar{\mathbf{z}} - (a \times \sigma) < \mathbf{z}_i$ [28]. It is guiding

the outlier detection test where the indicator itself is altered by the existence of outlying values in the data.

3.2.5. Quartile and Percentile Test (Box-Whisker Plot). Tukey in 1977 proposed a graphical outlier indication test to identify the skewness and unusual data points in the data distribution [34]. It can detect one or more outlier values in the data set and can also detect outliers in the upper and lower boundaries of the data samples [35]. The following are the steps of the quartile and percentile test.

(i) Determine the first quartile of the data as the 25^{th} percentile (Q_1) in $\mathbf{z}$

(ii) Identify the third quartile of the data as the 75^{th} percentile (Q_3) in $\mathbf{z}$

(iii) Determine the interquartile range (IQR) of the $\mathbf{z}$ vector as

$$\text{IQR} = Q_3 - Q_1. \tag{18}$$

(iv) A data value is considered an outlier in the lower fence if it exceeds the results in

$$Q_1 - 1.5(\text{IQR}). \tag{19}$$

(v) Similarly, a value is considered an outlier in the upper fence of the data set if it exceeds the results in

$$Q_3 + 1.5(\text{IQR}). \tag{20}$$

Figure 3 shows the diagram of the box-whisker plot. All parameters are indicated in the figure.

3.3. Modified HDF Schemes. After detecting the reports of MUs at the FC, global decision is made by FC in the modified form in the subsequent sensing intervals. In the modified AND scheme, sensing reports of the normally declared cooperative users are considered in the global decision. Similarly, the reports received from the detected outliers such as MUs are deleted in this combination. Hence, the modified equation of the AND decision scheme now takes the following form:

$$G_{\text{d}} = \begin{Bmatrix} H_1: & \sum_{\substack{j=1 \\ j\neq \text{MU}}}^{n} Z_j(i) = n_{\text{Modified}} \\ H_0: & \text{otherwise} \end{Bmatrix}. \tag{21}$$

In the modified AND scheme, only normal SU n_{Modified} sensing reports about the presence of PU are considered, whereas the reports of OMUs and LMUs are discarded.

The criteria for the decision of the OR scheme are modified as

$$G_{\text{d}} = \begin{Bmatrix} H_1: & \sum_{\substack{j=1 \\ j\neq \text{MU}}}^{n} Z_j(i) \geq 1 \\ H_0: & \text{otherwise} \end{Bmatrix}. \tag{22}$$

The modified majority voting scheme takes its global decision based on the reports of normally declared users. After the filtration and elimination of the MUs, the modified majority voting scheme decision is given as

$$G_{\text{d}} = \begin{Bmatrix} H_1: & \sum_{\substack{j=1 \\ j\neq \text{MU}}}^{n} Z_j(i) \geq k_{\text{Modified}} \\ H_0: & \text{otherwise} \end{Bmatrix}, \tag{23}$$

where k_{Modified} is the number of sensing reports received from the normal SUs that declare the presence and absence of the PU signal by H_1 andH_0, respectively.

4. Simulation Results

In this section, we present simulation results of the proposed outlier detection-based HDF schemes and compared them with other statistical outlier schemes. In the simulation, the number of MUs varied in the cooperating environment to investigate the overall effect in the CSS. The simulation parameters are defined in Table 1.

4.1. Case 1: Median Test Results. In case 1, the results for the median test are plotted using HDF schemes. The median test results in Figures 4 and 5 are compared with those of the simple HDF schemes. Figure 4 shows the simulation results when there are no MUs in the cooperative environment. It is observed that when the proposed median test is applied, the error probability reduces than when the traditional HDF scheme is applied. In Figure 4, an increase in SNR from -30 dB to -15 dB results in an abrupt change in the error probability for proposed OR and AND schemes, where error probability reduces from 0.47 to 0.25 for the OR scheme. In the AND decision scheme, error probability starts at 0.50 that gradually reduces to 0.23 when SNR exceeds from -30 dB to -10 dB for the proposed outlier test. Similarly, the proposed majority voting error probability is reduced from 0.26 to 0.25, when SNR is increased from -40 dB to -10 dB. The proposed HDF schemes show better sensing results with minimum error probabilities, while the simple HDF schemes result in maximum error probability.

In Figure 5, the number of LMUs is increased to five with one OMU reporting with normal SUs. In this case, error probability remains high for the traditional HDF schemes while the proposed test has a reduced error probability. It can be observed that for the simple HDF schemes when the number of MUs is increased, the error probability remains high at all SNR values, i.e., 0.53 approximately for the simple

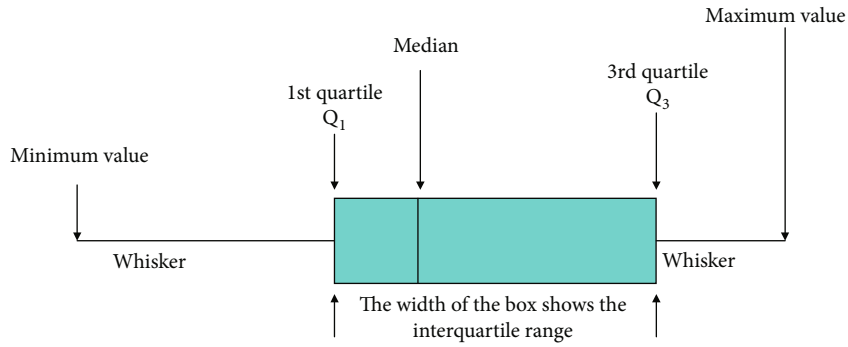


Figure 3: Box-whisker plot.

Table 1: Simulation parameters.

Parameter	Value
Total number of users	40
Malicious users	2
SNR range	-40 dB to -10 dB
Number of samples in each interval	270
Sensing iterations	1000
Iteration range for simple HDF	1-500
Iteration range for modified HDF	501-1000
Time consumption	1 msec

OR scheme and 0.5 for the simple AND scheme. Similarly, for the simple majority voting scheme, the error probability starts at 0.28 approximately and gradually reduces to 0.26. The proposed AND scheme has an error probability of 0.5 at the SNR value of -40 dB that sharply reduces to 0.26 when SNR is increased from -25 dB to -10 dB. The proposed OR scheme has an error probability of 0.48 at the SNR value of -40 dB that reduces to 0.25 at the SNR value of -10 dB. Similarly, in the case of the proposed majority voting scheme, error probability starts at 0.26 approximately and remains lower than that in the simple majority voting scheme.

The results of percent decrease in error probability of the modified and traditional HDF schemes at different SNRs values are illustrated in Table 2 in the presence of LMUs and OMU. The table result shows that when SNRs = −26 dB, the proposed voting scheme results in better sensing performance with 9.1% minimum sensing error probability compared with the simple voting scheme. Similarly, the proposed OR scheme obtained 8.9% reduction in error probability as compared with the simple OR scheme, while the proposed AND decision scheme has 0.4% reduction in the error probability results compared with the simple AND combination scheme. As the SNRs are increased to -10 dB, the percent decrease in error probability of the proposed HDF schemes is further improved for the proposed voting (9.8%), proposed OR (1.7%), and proposed AND (47.1%) schemes compared with the simple voting, simple OR, and simple AND decision schemes.

4.2. Case 2: Performance Comparison of the Proposed Scheme with Other Statistical Outlier Test Schemes. In this case, we present the performance comparison of the proposed median test with the other statistical outlier tests in Figures 6–11. HDF schemes are plotted separately and compared with outlier tests. The comparison is made for HDF schemes in the following scenario in CSS.

(1) When no MU exists in the network

(2) When five LMUs and one OMU exist in the network

4.2.1. Scenario 1: OR Scheme. A global decision of the OR scheme is made when a single SU detects the presence of a PU signal; hence, there is a chance of error in the sensing report. In Figures 6 and 7, the results for the OR scheme is investigated for all outlier tests along with the results of the simple HDF (OR) scheme in the global decision. These figures show that the proposed test scheme is outperforming other statistical outlier test schemes, when there is no MU in the network. The SNR varies from -40 dB to -10 dB. In the simple OR scheme, error probability starts at 0.51 that reduces after -20 dB and reaches a value of 0.38 approximately at -10 dB. All the statistical outlier test schemes have a starting error probability of 0.49 which is lower than that of the simple HDF OR scheme and abruptly reduces after -30 dB. The mean test results are with maximum error probability among all other outlier detection tests as SNR ranges from -30db to -10 dB which is followed by the GESD, Grubbs, and quartile tests. The proposed median outlier test scheme has minimum error probability from -30 dB to -10 dB.

In Figure 7, there are five LMUs with one OMU and 34 normal SUs reporting to FC for a global decision. The simple HDF scheme has an error probability of 0.51 approximately at -40 dB that is slightly reduced to 0.49 at -10 dB. Similarly, all the statistical outlier tests have the same probability of error up to -25 dB which is slightly reduced to 0.47 approximately for the mean test. For the GESD test, the P_e is 0.45 at -10 dB, and for the Grubbs test, the P_e is 0.43 approximately at -10 dB. These results further reduce to 0.26 approximately for the quartile test at the SNR value of -10 dB. Similarly, when the SNR value exceeds -25 dB, the proposed test scheme curve is skewed down to the error probability of

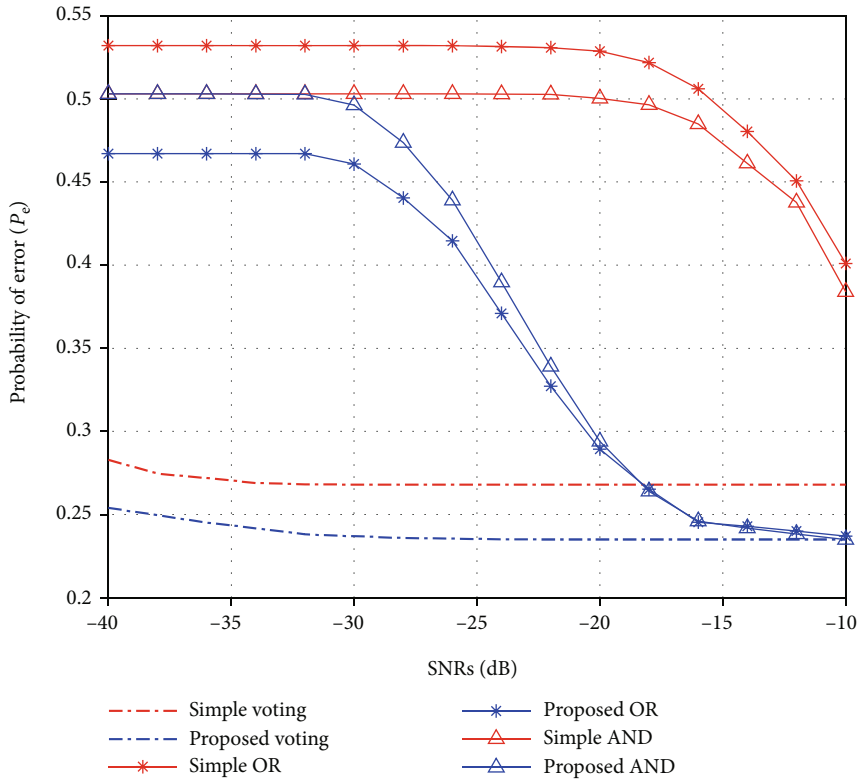


Figure 4: No MUs with 40 normal SUs in the network.

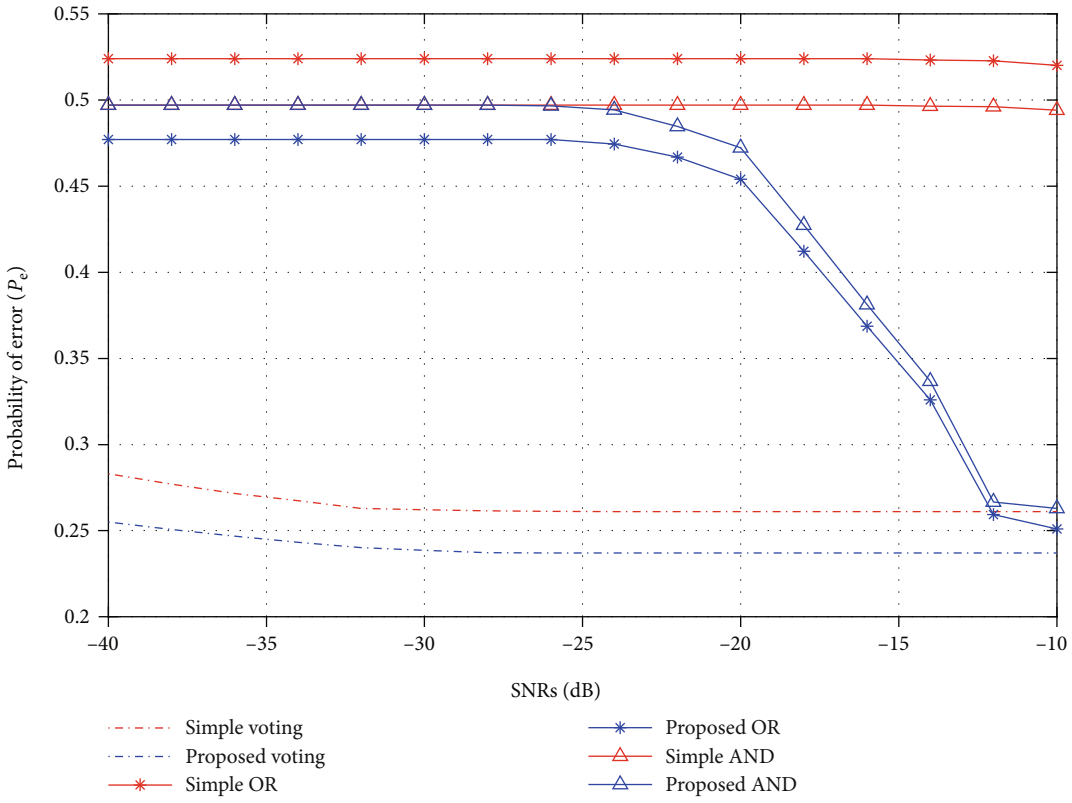


Figure 5: Five LMUs and one OMU with 34 normal SUs in the network.

Table 2: Percent decrease in the error probabilities for modified HDF schemes.

Decision schemes	SNR values				
	-26 dB	-22 dB	-18 dB	-14 dB	-10 dB
Simple vs. proposed voting	9.1%	9.1%	9.1%	9.1%	9.8%
Simple vs. proposed OR	8.9%	11%	21.3%	38.7%	51.7%
Simple vs. proposed AND	0.4%	2.6%	14%	32.2%	47.1%

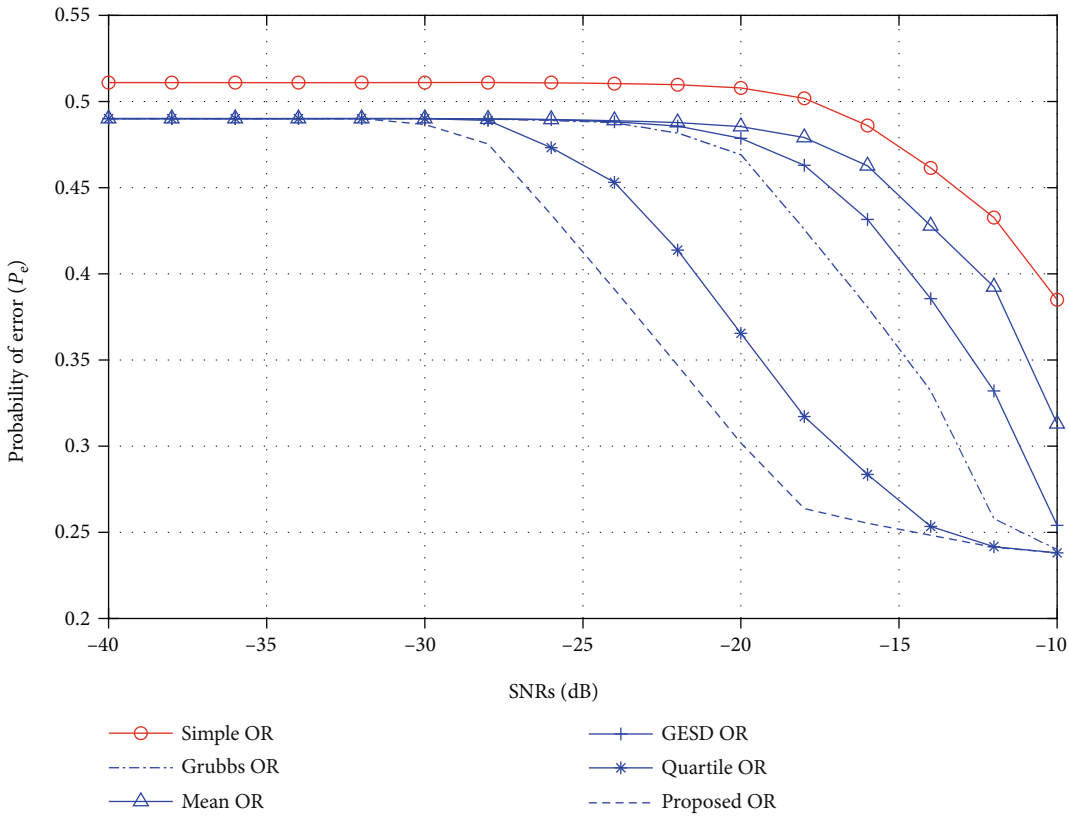


Figure 6: No MUs with 40 normal SUs in the network.

0.25 at -10 dB, which is the minimum error probability of all outlier tests.

The results in Table 3 illustrate the percent decrease in the error probability of the proposed median test-based OR HDF combination scheme as compared with other outlier detection tests in the presence of LMUs and OMU at various SNRs. The table shows that at -26 dB, the proposed OR HDF scheme results in better sensing performance with 1.67% reduction in sensing error probability compared with the mean OR, GESD OR, Grubbs OR, and quartile OR test schemes. Similarly, as SNRs are increased to -10 dB, the percent decrease in error probability of the proposed OR HDF scheme is further improved as compared with that of the mean OR (91.4%), GESD OR (81.7%), Grubs OR (72.7%), and quartile OR (4.87%) test schemes.

4.2.2. Scenario 2: AND Scheme. When all the SUs confirm the presence of the PU signal, a global decision is made in favor of the AND scheme. This scheme is tested at the same SNR values of -40 dB to -10 dB. In Figure 8, no MUs are included in sensing. Therefore, the simple HDF scheme performed moderately with an error probability of 0.50 which proceeded to 0.37 approximately at -10 dB. Similarly, other outlier tests have minimum error probabilities compared with the simple HDF scheme at -10 dB. The mean test has $P_e = 0.36$, the GESD test has $P_e = 0.3$, and the Grubbs test, quartile test, and proposed test have $P_e = 0.27$ approximately. It is observed that for the SNR values between -30 dB and -15 dB, the proposed median test scheme has the best performance with minimum error probability.

The simple HDF scheme has poor sensing performance when MUs appear in sensing with constant $P_e = 0.52$ approximately from -10 dB to -40 dB as observed in Figure 9. The same maximum P_e of 0.52 is observed for all outlier tests from -40 dB to -20 dB. These results are followed by the mean test that has $P_e = 0.46$, Grubbs and GESD tests with $P_e = 0.45$, and quartile test with $P_e = 0.34$ approximately at -10 dB. The proposed median test has $P_e = 0.28$ at -10 dB, which is

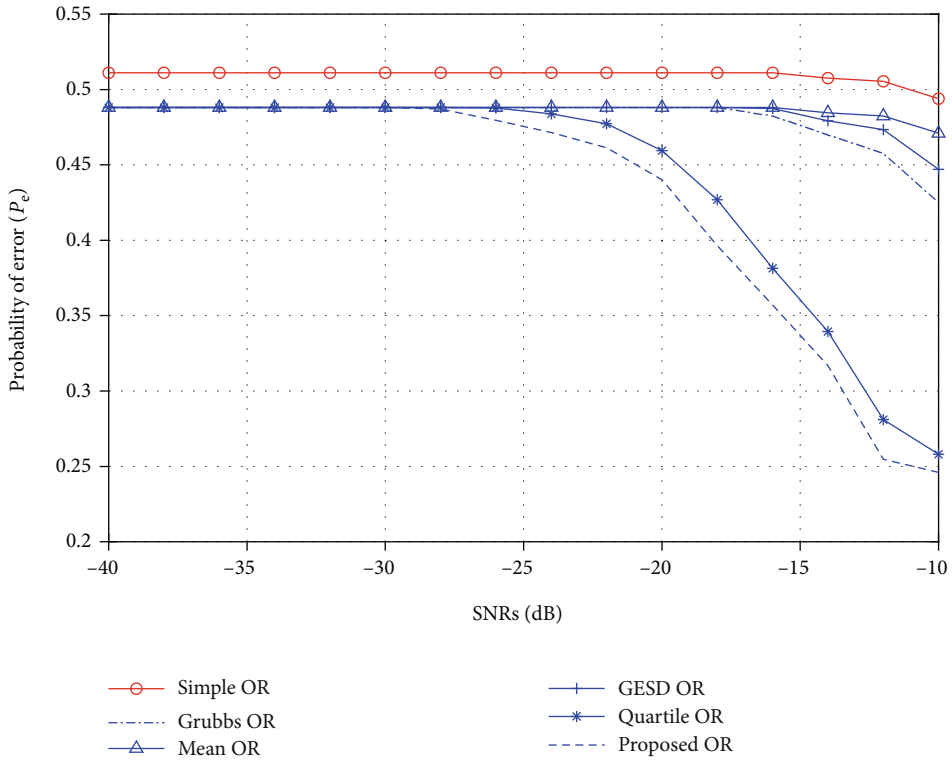


Figure 7: Five LMUs and one OMU with 34 normal SUs in the network.

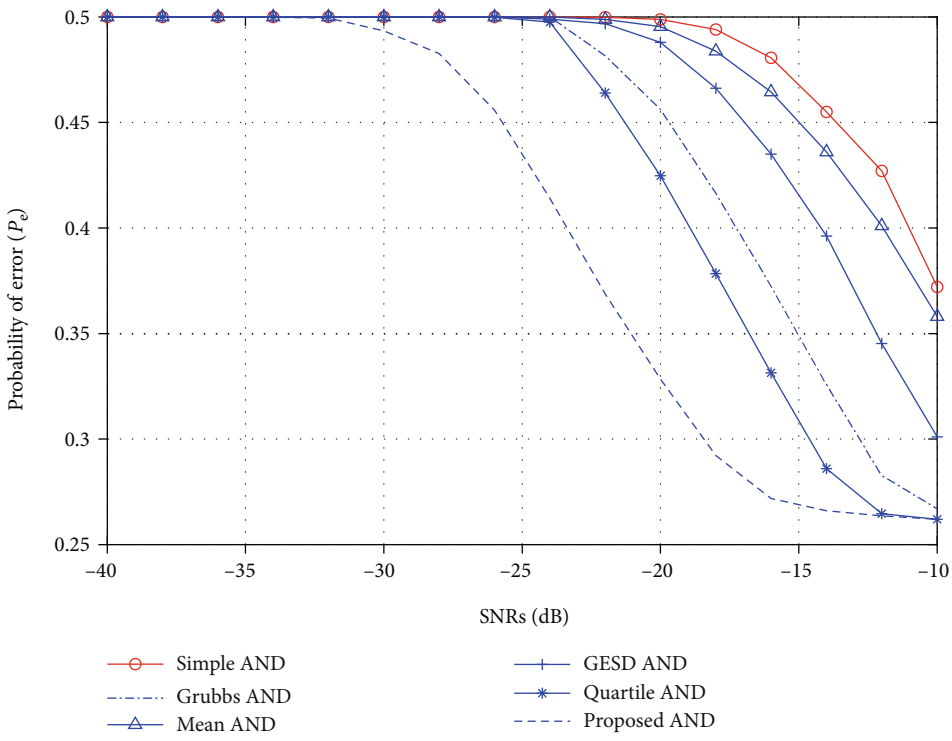


Figure 8: No MUs with 40 normal SUs in the network.

lowest in all the outlier test schemes. The proposed median test surpasses all other outlier tests at SNR values from -15 dB to-10 dB with the minimum P_e of 0.28 at -10 dB.

Table 4 shows the performance gain in terms of percent decrease in error probabilities at different SNR values for the proposed outlier detection test using the AND HDF

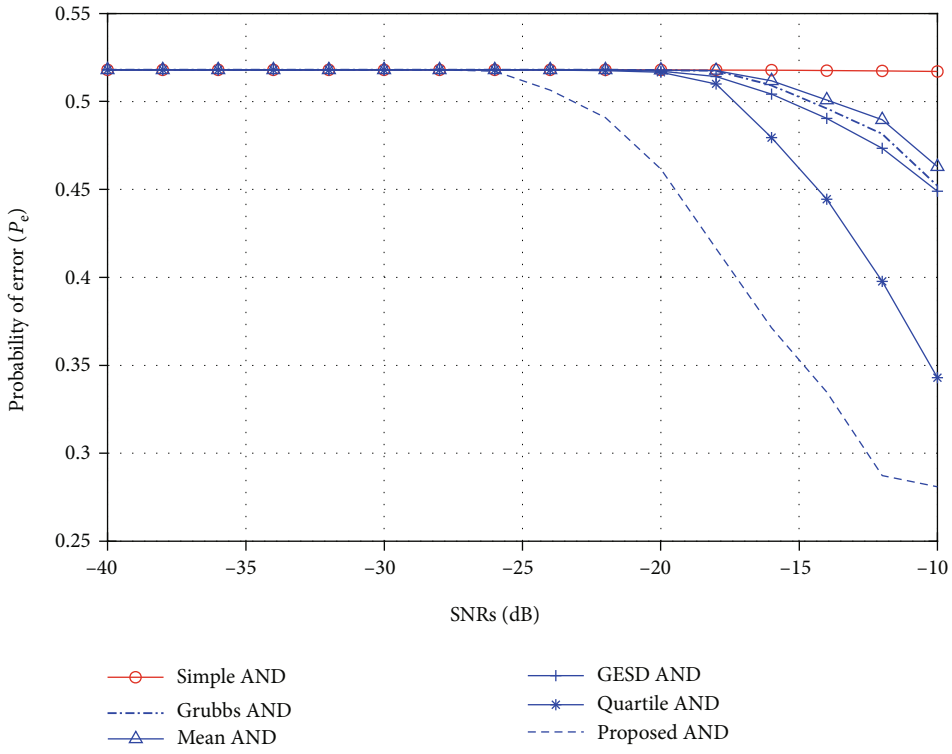


Figure 9: Five LMUs and one OMU with 34 normal SUs in the network.

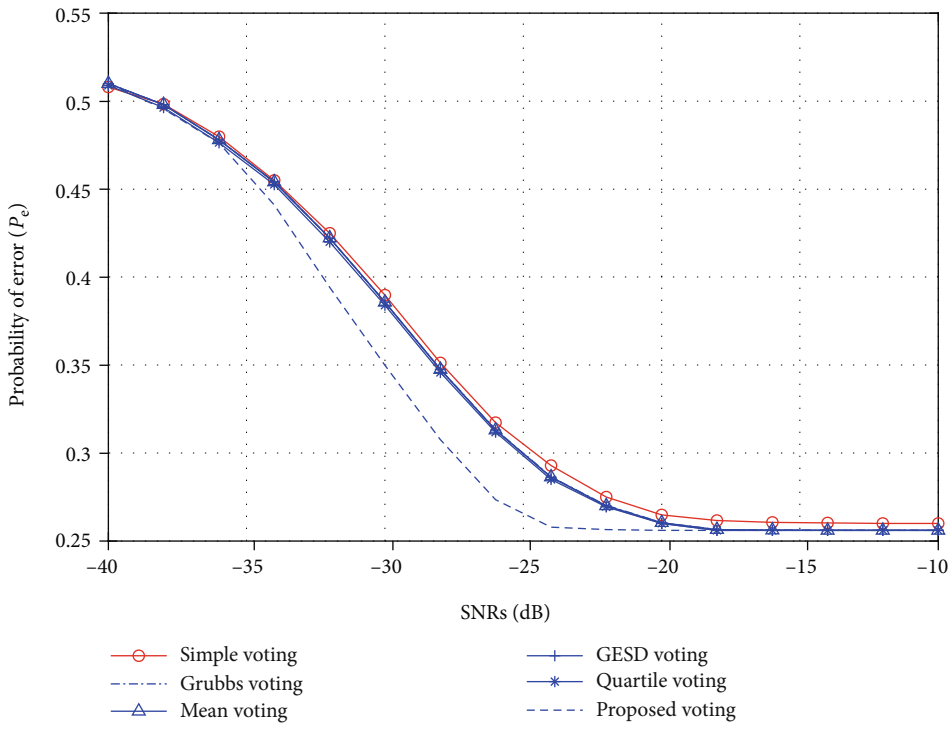


Figure 10: No MUs with 40 normal SUs in the network.

scheme containing LMUs and OMU. The table shows that at -26 dB, the proposed AND HDF scheme results in better performance gain with 0.19% decrease in sensing error probability compared with the mean AND, GESD AND, Grubbs AND, and quartile AND schemes. Similarly, as the SNRs are increased to -10 dB, the percent decrease in error probability of the proposed AND HDF scheme is further improved as compared with the mean AND (64.7%), GESD AND

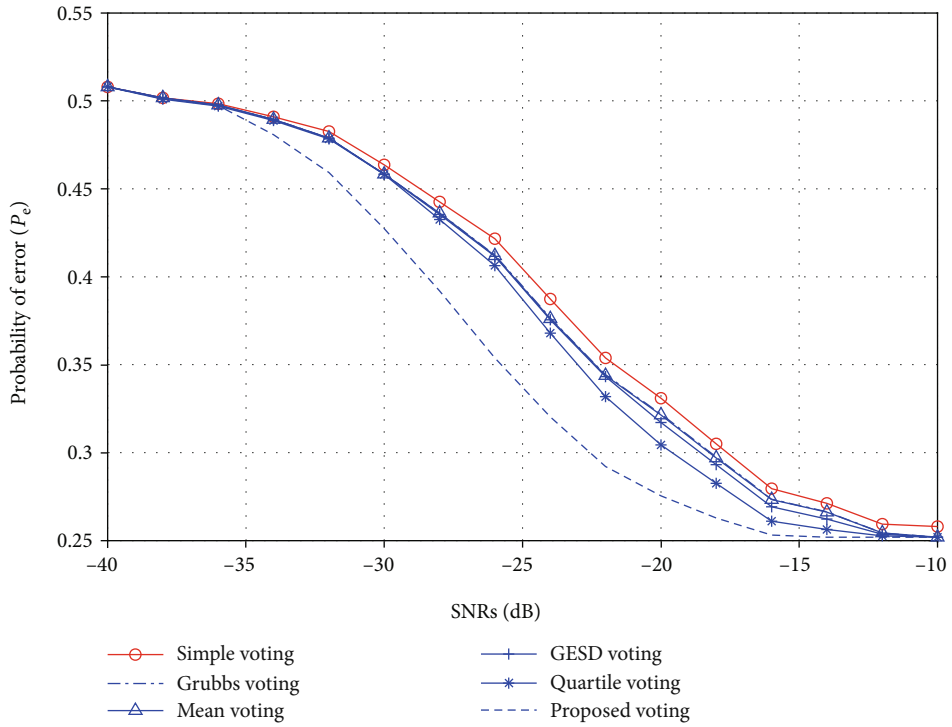


Figure 11: Five LMUs and one OMU with 34 normal SUs in the network.

Table 3: Percent decrease in the error probabilities for the proposed OR HDF scheme.

Decision schemes	SNR values				
	-26 dB	-22 dB	-18 dB	-14 dB	-10 dB
Proposed OR vs. mean OR	1.67%	5.8%	23.2%	53.1%	91.4%
Proposed OR vs. GESD OR	1.67%	5.8%	23.2%	51.5%	81.7%
Proposed OR vs. Grubbs OR	1.67%	5.8%	23.2%	48.4%	72.7%
Proposed OR vs. quartile OR	1.67%	3.47%	16.6%	7.27%	4.87%

Table 4: Percent decrease in the error probabilities for the proposed AND HDF scheme.

Decision schemes	SNR values				
	-26 dB	-22 dB	-18 dB	-14 dB	-10 dB
Proposed AND vs. mean AND	0.19%	5.71%	24.2%	49.7%	64.7%
Proposed AND vs. GESD AND	0.19%	5.71%	23.5%	46.7%	59.7%
Proposed AND vs. Grubbs AND	0.19%	5.5%	24.2%	48.5%	60.8%
Proposed AND vs. quartile AND	0.19%	5.7%	22.5%	32.9%	22%

(59.7%), Grubbs AND (60.8%), and quartile AND (22%) schemes.

4.2.3. Scenario 3: Majority Voting Scheme. For simulation purposes, $k = n/2$ is selected for the majority voting scheme, where more than one SU has to declare the PU channel available to make its decision about the presence and absence of PU; otherwise, PU absence is declared. Figure 10 shows the results without any misbehaving users. The simple voting scheme has maximum P_e with a starting value of 0.51 at the SNR value of -40 dB and reduces to a value of 0.26 at the SNR value of -10 dB. All outlier tests give similar P_e results at all SNR levels. The proposed median test results are slightly improved giving low P_e values from SNR values of -35 dB to -20 dB which are significantly lower than those of the simple HDF of the majority voting scheme.

The value of P_e is increased for the simple majority voting scheme when MUs transfer reports in the sensing interval. From Figure 11, it is observed that P_e of the simple majority voting scheme is 0.51 approximately at -40 dB and after

Table 5: Percent decrease in the error probabilities of the proposed voting scheme.

Decision schemes	SNR values				
	-26 dB	-22 dB	-18 dB	-14 dB	-10 dB
Proposed voting vs. mean voting	16.7%	17.4%	13.3%	5.5%	0.4%
Proposed voting vs. GESD voting	16.4%	17.4%	11.8%	3.9%	0.4%
Proposed voting vs. Grubbs voting	16.4%	17.4%	13.3%	5.5%	0.4%
Proposed voting vs. quartile voting	15%	13.3%	7.6%	1.6%	0.4%

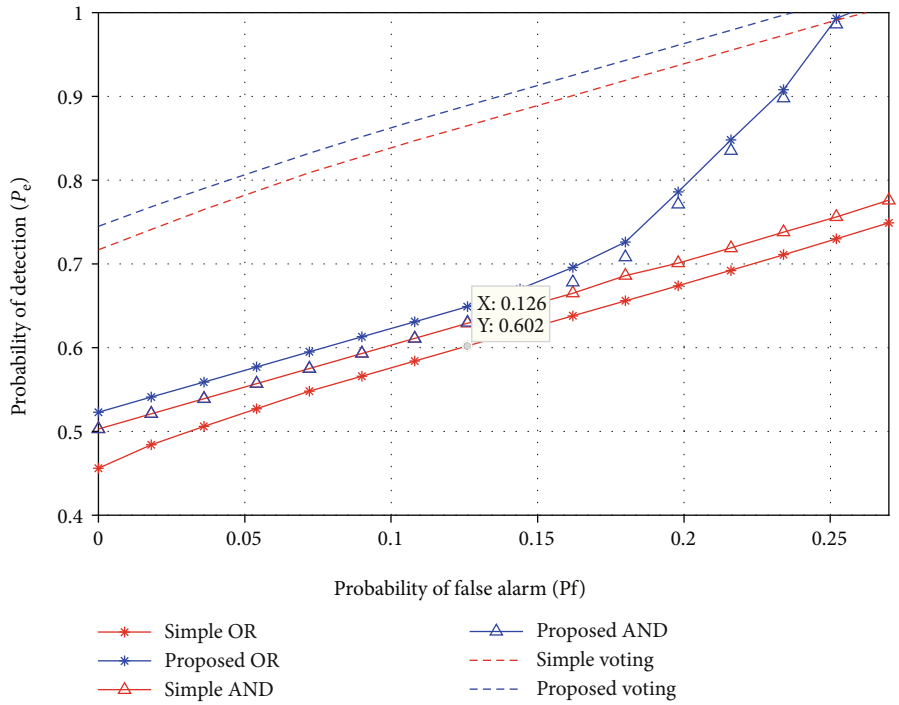


Figure 12: Probability of detection (P_d) vs. probability of false alarm (P_f) for the simple and modified HDF schemes with normal SUs and MUs.

-30 dB it reduces further to 0.26. Outlier test schemes show that P_e is between -35 dB and -15 dB and is minimized slightly compared with that of the simple majority voting scheme. On the other hand, the proposed median test scheme shows significance upon all outlier test schemes from SNR values of -35 dB to -15 dB. It is observable from the simulation results that the outlier test scheme detects falsifying reports of MUs and removes them in the final decision that decreases the error probability.

The overall performance gain of the voting scheme is better than that of the AND HDF and OR HDF schemes by establishing minimum error probability. Table 5 illustrates the results of percent decrease in the error probabilities obtained by the proposed outlier detection test using the voting scheme in the presence of LMUs and OMU at different SNR values. At -26 dB, the proposed voting scheme results in better sensing performance with 16.7% decrease in sensing error probability compared with the mean voting scheme, 16.4% decrease compared with the GESD voting and Grubbs voting schemes, and 15% improvement compared with the quartile voting scheme. As the SNRs are i5ncreased to -10 dB, the percent decrease in error probability of the proposed voting scheme is 0.4% as compared with that of mean voting, GESD voting, Grubs voting, and quartile voting schemes.

The receiver operating characteristics (ROC) with probability of detection P_d vs. probability of false alarm P_f are collected in the presence of LMUs and OMUs in Figures 12–15. In Figure 12, the results are plotted for simple HDF and proposed (modified) HDF schemes. The simple OR HDF results are highly deteriorated by producing high P_f values with the contributions of MUs, whereas the proposed (modified) OR decision scheme gives better detection results. Similarly, AND combination scheme detection probability with the employment of the proposed median test has better detection results with minimum false alarm than the detection probability of the traditional AND decision scheme. Likewise, detection results of the proposed voting scheme remain superior and surpass those of all other HDF schemes in Figure 12.

The modified scheme ROC results are further compared with those of the other outlier detection tests to investigate the proposed test superiority. Figure 13 shows the result illustrations for the OR HDF scheme which was compared with the other outlier detection tests. It is observable from the

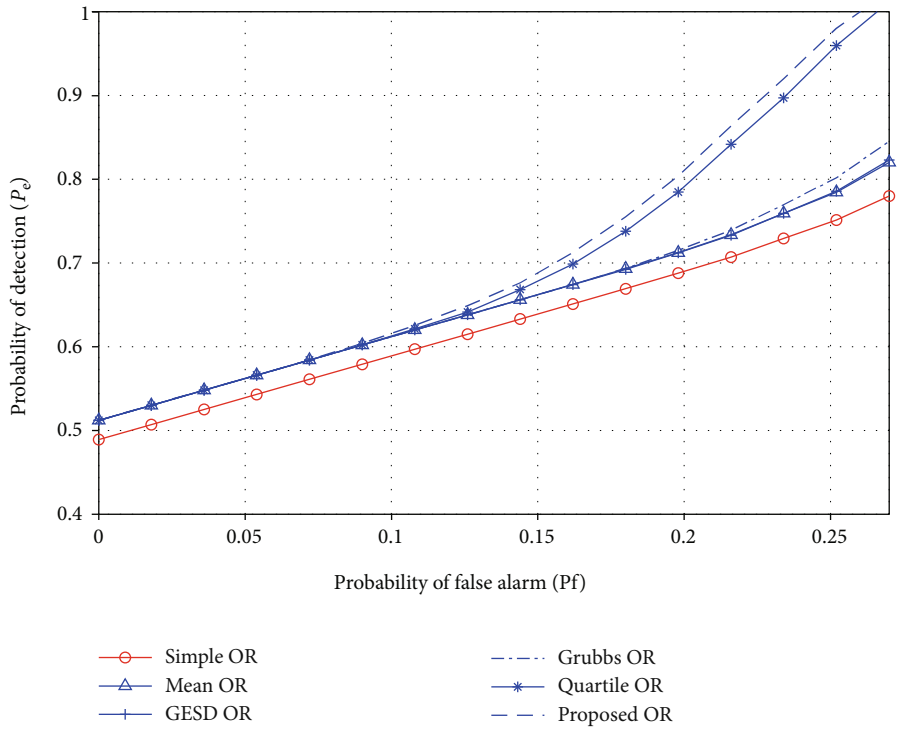


Figure 13: Probability of detection (P_d) vs. probability of false alarm (P_f) of the OR HDF schemes with the normal SUs and MUs.

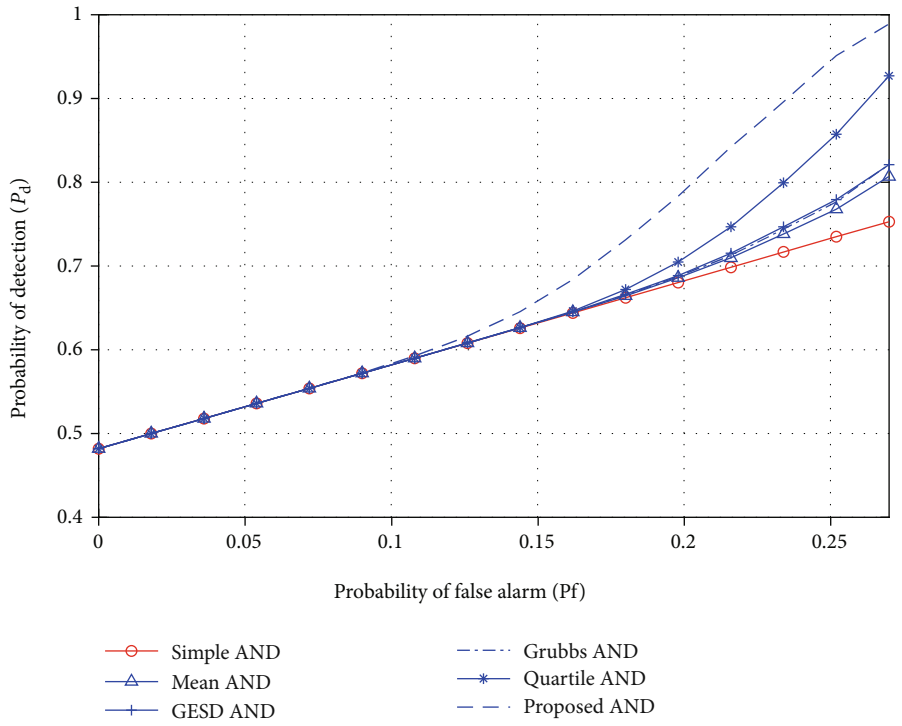


Figure 14: Probability of detection (P_d) vs. probability of false alarm (P_f) of the AND HDF schemes with the normal SUs and MUs.

results in Figure 13 that the proposed median outlier test has attained maximum detection probability, whereas the simple OR HDF scheme shows minimum detection probability. The median test performance is next followed by the quartile and corresponding Grubbs tests. The mean and GESD tests produce similar detection results with their detection probabilities comparatively limited as compared with the detection probability of the proposed test.

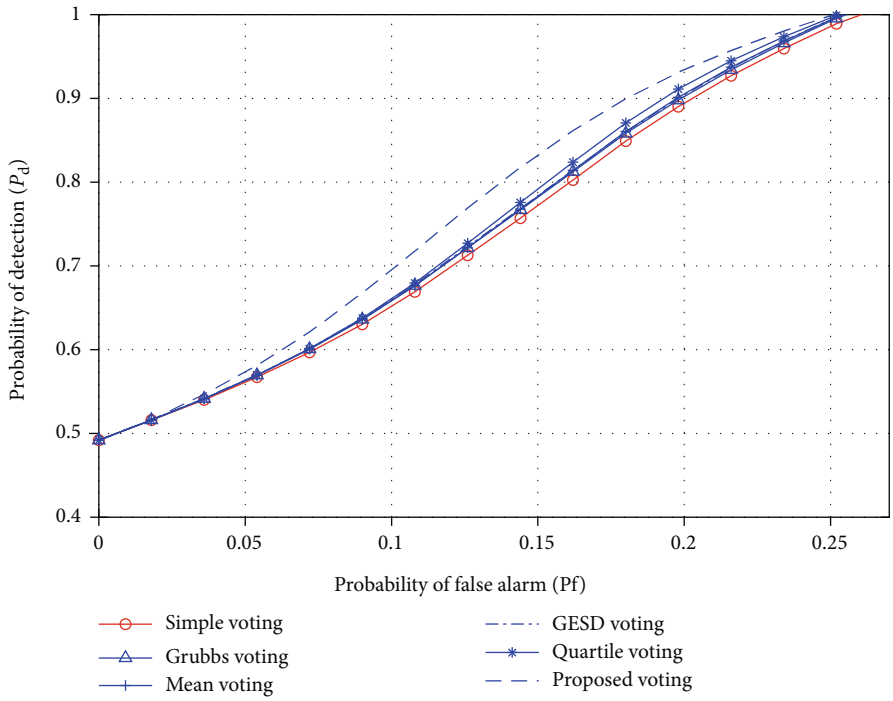


Figure 15: Probability of detection (P_d) vs. probability of false alarm (P_f) results of the voting HDF schemes with the normal SUs and MUs.

In Figure 14, it is noticeable that the proposed median test-based AND HDF scheme has high detection probability with minimum false alarm probability in comparison with the other outlier-based detection results. The proposed median test-based ROC results are followed by the quartile test, GESD test, and Grubbs test. The mean test has the minimum detection probability as compared with all other outlier detection tests.

The voting HDF scheme ROC results are shown in Figure 15 to compare the proposed and various outlier detection tests. In Figure 15, all other outlier detection tests give similar detection results, while the proposed median test is able to achieve significant improvement over all other outlier detection tests.

5. Conclusion

The CSS is reliable in detecting the presence and absence of the PU signal; however, the participation of the MUs in the CSS results in false report collection at the FC. This research work considered the involvement of the MUs in the CSS. An improved statistical analysis is employed for spectrum sensing in the CRN. The focus in this research work is to boost the performance of the traditional HDF schemes with some statistical analysis. The false reports of the MUs can be efficiently detected using different outlier tests. The results of the four outlier tests are compared and concluded that the median plus-minus 3 test outperforms other statistical outlier detection tests. The proposed outlier test is accurately detecting the behavior of the LMU and OMU in the CSS.

For future work, it is recommended that these outlier statistics should be further investigated by applying them to detect the MU behavior of always yes, always no, and random opposite categories of the MUs. Similarly, other categories of outlier detection techniques such as density-based, depth-based, and cluster-based schemes can be employed.

Data Availability

The data used to support the finding of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01426) supervised by the IITP (Institute for Information and Communication Technology Planning & Evaluation) and in part by the National Research Foundation (NRF) funded by the Korea Government (MSIT) (No. 2019R1F1A1059125).

References

[1] K. Ben Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, 2009.

[2] S. M. Diamond and M. G. Ceruti, "Application of wireless sensor network to military information integration," in *2007 5th*

IEEE International Conference on Industrial Informatics, pp. 316–322, Vienna, Austria, June 2007.

[3] N. Gul, I. M. Qureshi, A. Omar, A. Elahi, and S. Khan, "History based forward and feedback mechanism in cooperative spectrum sensing including malicious users in cognitive radio network," *PLOS ONE*, vol. 12, no. 8, article e0183387, 2017.

[4] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications," *Mobile Networks and Applications*, vol. 6, no. 5, pp. 435–441, 2001.

[5] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Malicious user detection in a cognitive radio cooperative sensing system," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2488–2497, 2010.

[6] T. S. Sundara and N. Padmaja, "Performance analysis of cognitive radio based on cooperative spectrum sensing," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 4, pp. 821–827, 2013.

[7] N. Gul, I. M. Qureshi, A. Naveed, A. Elahi, and I. Rasool, "Secured soft combination schemes against malicious-users in cooperative spectrum sensing," *Wireless Personal Communications*, vol. 108, no. 1, pp. 389–408, 2019.

[8] P. Pandya, A. Durvesh, and N. Parekh, "Energy detection based spectrum sensing for cognitive radio network," in *2015 Fifth International Conference on Communication Systems and Network Technologies*, pp. 201–206, Gwalior, India, April 2015.

[9] O. León and K. P. Subbalakshmi, *Cognitive Radio Network Security BT - Handbook of Cognitive Radio*, W. Zhang, Ed., Springer Singapore, Singapore, 2017.

[10] M. Khasawneh and A. Agarwal, "A collaborative approach for monitoring nodes behavior during spectrum sensing to mitigate multiple attacks in cognitive radio networks," *Security and Communication Networks*, vol. 2017, Article ID 3261058, 16 pages, 2017.

[11] S. Shobana, R. Saravanan, and R. Muthaiah, "Matched filter based spectrum sensing on cognitive radio for OFDM WLANs," *International Journal of Engineering and Technology*, vol. 5, no. 1, pp. 142–146, 2013.

[12] N. Do and B. An, "A soft-hard combination-based cooperative spectrum sensing scheme for cognitive radio networks," *Sensors*, vol. 15, no. 2, pp. 4388–4407, 2015.

[13] Y. L. Lee, W. K. Saad, A. A. El-Saleh, and M. Ismail, "Improved detection performance of cognitive radio networks in AWGN and Rayleigh fading environments," *Journal of Applied Research and Technology*, vol. 11, no. 3, pp. 437–446, 2013.

[14] W. Han, J. Li, Z. Li, J. Si, and Y. Zhang, "Efficient soft decision fusion rule in cooperative spectrum sensing," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 1931–1943, 2013.

[15] N. Gul, I. M. Qureshi, A. Elahi, and I. Rasool, "Defense against malicious users in cooperative spectrum sensing using genetic algorithm," *International Journal of Antennas and Propagation*, vol. 2018, Article ID 2346317, 11 pages, 2018.

[16] A. Haldorai and U. Kandaswamy, "Secure distributed spectrum sensing in cognitive radio networks," *Intelligent Spectrum Handovers in Cognitive Radio Networks*, pp. 175–191, 2019.

[17] J. Luo and X. He, "A soft–hard combination decision fusion scheme for a clustered distributed detection system with multiple sensors," *Sensors*, vol. 18, no. 12, p. 4370, 2018.

[18] R. Ujjinimatad and S. R. Patil, "Spectrum sensing in cognitive radio networks with known and unknown noise levels," *IET Communications*, vol. 7, no. 15, pp. 1708–1714, 2013.

[19] N. Reisi, S. Gazor, and M. Ahmadian, "Distributed cooperative spectrum sensing in mixture of large and small scale fading channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5406–5412, 2013.

[20] N. Reisi, S. Gazor, and M. Ahmadian, "A distributed average likelihood ratio detector for detection of signals in frequency-selective Nakagami channels," *IEEE Wireless Communications Letters*, vol. 3, no. 3, pp. 245–248, 2014.

[21] S. Shrivastava, S. John, A. Rajesh, and P. K. Bora, "Preventing collusion attacks in cooperative spectrum sensing," in *2018 International Conference on Signal Processing and Communications (SPCOM)*, pp. 90–94, Bangalore, India, July 2018.

[22] A. G. Fragkiadakis, E. Z. Tragos, and I. G. Askoxylakis, "A survey on security threats and detection techniques in cognitive radio networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 428–445, 2013.

[23] M. Abdelhakim, L. E. Lightfoot, J. Ren, and T. Li, "Distributed detection in mobile access wireless sensor networks under Byzantine attacks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 950–959, 2014.

[24] N. Gul, A. Naveed, A. Elahi, T. Saleemkhattak, and I. M. Qureshi, "A combination of double sided neighbor distance and genetic algorithm in cooperative spectrum sensing against malicious users," in *2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pp. 746–753, Islamabad, Pakistan, January 2017.

[25] D. Divya and S. S. Babu, "Methods to detect different types of outliers," in *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, pp. 23–28, Ernakulam, India, March 2016.

[26] M.-j. Zhou and X.-j. Chen, "An outlier mining algorithm based on dissimilarity," *Procedia Environmental Sciences*, vol. 12, pp. 810–814, 2012.

[27] M. Gupta, J. Gao, C. C. Aggarwal, and J. Han, "Outlier detection for temporal data: a survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 9, pp. 2250–2267, 2014.

[28] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median," *Journal of Experimental Social Psychology*, vol. 49, no. 4, pp. 764–766, 2013.

[29] C. C. Agarwal, *Outlier Analysis Second Edition*, vol. 24, no. 2, 2017.

[30] R. Kumar and A. Khadar, "A survey on outlier detection techniques in dynamic data stream," *International Journal of Latest Engineering and Management Research*, vol. 2, no. 8, pp. 23–30, 2017.

[31] T. A. Cohn, J. F. England, C. E. Berenbrock, R. R. Mason, J. R. Stedinger, and J. R. Lamontagne, "A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series," *Water Resources Research*, vol. 49, no. 8, pp. 5047–5058, 2013.

[32] Y. Guo, Q. Xu, S. Sun, X. Luo, and M. Sbert, "Selecting video key frames based on relative entropy and the extreme studentized deviate test," *Entropy*, vol. 18, no. 3, p. 73, 2016.

[33] D. Cousineau and S. Chartier, "Outliers detection and treatment: a review," *International Journal of Psychological Research*, vol. 3, no. 1, 2010.

[34] A. Li, M. Feng, Y. Li, and Z. Liu, "Application of outlier mining in insider identification based on boxplot method," *Procedia Computer Science*, vol. 91, pp. 245–251, 2016.

[35] V. Praveen, T. Delhi Narendran, R. Pavithran, and C. Thirumalai, "Data analysis using box plot and control chart for air quality," in *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, pp. 1082–1085, Tirunelveli, India, May 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8845371, 17 pages
https://doi.org/10.1155/2020/8845371

WILEY | Hindawi

Review Article

Nonorthogonal Multiple Access for Next-Generation Mobile Networks: A Technical Aspect for Research Direction

Muhammad Hussain and Haroon Rasheed

Electrical Engineering Department, Bahria University Karachi Campus, Karachi 75260, Pakistan

Correspondence should be addressed to Muhammad Hussain; engr.m.hussain.bukc@bahria.edu.pk

Received 22 March 2020; Revised 14 July 2020; Accepted 29 October 2020; Published 30 November 2020

Academic Editor: Farman Ullah

5G mobile communications offer several benefits, which include providing extremely low latency, very high data rates, significant improvement in the number of users, and increase in base station capacity and perceived quality of service. This may be achieved at the cost of an increased receiver complexity by nonorthogonal access of users. Nonorthogonal multiple access (NOMA) is one of the capable contenders to achieve the vision of 5G wireless communications. Supporting a higher number of users than available orthogonal resources is the key feather of NOMA. In this article, the basic principle of NOMA has been reviewed and compared with other orthogonal multiple access (OMA). A comprehensive survey is presented in the latest NOMA scheme. The distinguished NOMA schemes design principle features, and recent deployments are discussed. Furthermore, the performance is compared in terms of the bit error rate, system capacity, and energy efficiency. The performance results show that NOMA can achieve the required goals, in terms of the user data rate, system capacity, interference cancellation scheme, and reception complexity.

1. Introduction

Multiple access schemes have been a landmark technology from 1G to 4G for the growth of mobile communications. As a design aspect, these multiple access technologies are mostly from the orthogonal multiple access (OMA) category; they are in the time domain, code domain, frequency domain, and time-frequency domain. OMA can easily detect the user information signal by utilizing a simple receiver. However, the entire number of users that the system can accommodate is firmly restricted by the number of available orthogonal resources. Also, the system requirements for synchronization are highly limited in order to guarantee the orthogonality of resource allocation among users. Therefore, it is very difficult for OMA to meet the data rate and other requirements of the next-generation mobile network. The 5G structure demands an innovative multiple access scheme to counter this challenge and recently proposed nonorthogonal multiple access (NOMA) technology which is accepted as a 5G multiple access scheme [1, 2].

Within the common physical layer using the code domain or power domain multiple access, NOMA permits numerous users to utilize frequency and time resources [3]. In recent times, various NOMA topologies have received a lot of attention due to attractive features. We can generally categorize into two types. These types are code domain multiple access and the power domain multiple access. NOMA achieved its goals by a combination of multiple access techniques like sparse code multiple access (SCMA) [4], multiuser shared access (MUSA) [5] with Low-Density Spreading (LDS) [6], and Pattern Division Multiple Access (PDMA) [7].

1.1. Motivation. In September 2014, the 3rd generation partnership project (3GPP) started the study on NOMA in Release 14 (Rel-14). NOMA may be combined with upcoming wireless communication systems in order to achieve the

requirements, including massive connectivity, high spectral and energy efficiency, significant achievable data rate, low latency, exceptional user fairness, large throughput, ultrahigh reliability, and upholding different quality of services (QoS).

Some previous impressive survey work on NOMA is followed. In [8], the transceiver block diagram of each category of NOMA is explained by the authors, regarding detailed key features, basic principles, and algorithms of the transmission-reception. In [9], characteristics and working principles of different NOMA schemes are summarized by the authors. In [10], NOMA schemes are compared and analysed by the authors. The authors focus on the future research directions of NOMA, prototype development, recent progress, standardization, and challenges. In [11], some promising nonorthogonal schemes were discussed which include sparse code multiple access (SCMA), Power Domain Non-orthogonal Multiple Access (PD-NOMA), Pattern Division Multiple Access (PDMA), multiuser shared access (MUSA), and some key modern waveforms including Generalized Frequency Division Multiplexing (GFDM), Universal Filtered Multicarrier (UFMC), and filter bank-based multicarrier (FBMC). The authors provided a future research path for 5G waveform and multiple access schemes by comparing and analysing the characteristics of these technologies.

However, in [8], achievable sum rate performance was presented, based on average mutual information rather than actual theoretical analysis. In [9], without mathematical justification, the performance of the NOMA schemes is assessed. In [10, 11], performance evaluation has not been examined by the authors. Furthermore, most of the previous work may just focus on one scheme, and no comprehensive work has been published to examine the performance of major NOMA schemes.

The objective of this research is to fill in the gap by presenting the basic principles, key features, and recent application of major categories of NOMA. Moreover, we present actual theoretical analysis and mathematical justification of the NOMA schemes. The major contributions are summarized as follows.

1.2. Contribution. In this article, a comprehensive and comparative survey on NOMA is presented.

(i) The survey includes different popular categories of NOMA, their basic model, working principles, technical aspect, key performance indicators (KPIs), advantages, and disadvantages

(ii) The article presents the state-of-the-art review of NOMA in enabling the 5G network, the applicability aspect of each category of the NOMA scheme, and the associated enablers

(iii) Moreover, in this article, we present important and recent deployments, potential challenges, and future trend work for researchers in the field of NOMA

(iv) The survey also includes the performance comparison of major categories of NOMA prototype in terms of achievable data rate, system capacity, energy efficiency, and bit error rate with mathematical justifications

Furthermore, this article is planned as follows: Section 2 is a recall history of mobile communication and their technology aspect. Section3 explains and investigates important nonorthogonal multiple access schemes with their principle of implementation, followed by a review of every scheme's key features and advantages and disadvantages. A summary of the NOMA scheme is presented in Section 4, and discussion of the results is presented in Section 5. Section 6 presents a review of recent developments in NOMA schemes, Section 7 presents the future research challenges of NOMA, and in Section 8, a conclusion is made.

2. Background

In the third generation mobile system, the Wideband Code Division Multiple Access (WCDMA) scheme was launched. As a result, movies can be transmitted due to improved speed of data communication. Furthermore, 3G presented an improved technology, i.e., High-Speed Packet Access (HSPA) and HSPA+ (3.5G), with which the user data experience was improved. However, in comparison to Wi-Fi and wireless LANs, high data rate applications like streaming of moving images were slower. Today, network operators provide services of 4G networks based on Long-Term Evolution (LTE). The achievable communication speed rises up to 5 to 6 times in comparison to 3G, and data throughput is also expressively enhanced in LTE than HSPA+. In LTE-Advanced (LTE-A), the available bandwidth is twice as LTE; therefore, several 4G network service providers are also transferred to LTE-Advanced (4.5G). With LTE and LTE-Advanced, communication technology has improved, at a level close to Wi-Fi with respect to user data experience. 4G network and LTE and LTE-A technology are saturated in terms of further improvement. The wireless data requirement is increasing day by day. Therefore, there is a need for new technology to speed up data access. However, for wireless communication, improvement in the data capacity and the data transmission rate is essential. Therefore, for the mobile Internet extension and modernization, researchers all over the world started investigating ways to improve data capacity and data transfer rates.

Meanwhile, from the beginning of digital communications in the 1990s, cellular phone technology has been on the track in terms of progress, focused on increasing the data rate and capacity. In the current world communication trends, mobile Internet and video calling have become a reality, and its new version has been launched, i.e., 5G mobile communication. Now, at any emergency condition such as online medical imaging or smart vehicles in congestion, more data needs to be delivered to the specific user. Thus, 5G networks will respond accordingly. Researchers also recognize 5G as an opportunity to redefine not only the network enable connecting a wide variety of new devices but also the networks that realize exceptional data rates. The next version of 5G wireless mobile technology is 6G, which means 6th generation wireless mobile technology. Satellite networks

for global coverage will be efficiently used in 6G which was not used before [12]. The 6G wireless mobile technology maximizes data throughput and improves system performance. The 6G technology is responsible for more data transfer and data security. It also increases data configuration choices. In 6G technology, devices connected to the Internet by using wireless broadband receive 10 GB or even more data speed. 6G is a satellite-based network; roaming and handover from one satellite to another satellite are still an issue which will be solved soon. The combination of fiber optics and the latest radio technology is used in 6G, to provide a very fast data experience. The 6G wireless mobile technology will change the way of thinking about wireless communication and will perform beyond the expectation of the users [12]. Moreover, this performance depends on technology use in next-generation networks.

Numerous proposals have been presented by researchers to establish the performance of NOMA in both downlink and uplink. The basic principle of downlink NOMA is presented in [13], power division is used for multiple user access at BS, and SIC is used for signal detection at the receiver. In [14], researchers proposed a two-user model for NOMA. Researchers presented link-level simulations and system-level simulations for the NOMA downlink system. Results provided in [14] showed that NOMA performance is better than OMA in terms of overall system throughput and individual user throughputs. The authors in [14] derived the closed-form expressions for outage probability and ergodic sum rate for the NOMA downlink system. In [15], to find the effect of user pairing for the two-user model of the NOMA system, the authors employed statistically allocated transmit powers among NOMA users. Moreover, the authors proposed fixed and opportunistic user pairing schemes. In [16], the authors consider the consequence of power allocation on fairness. To ensure that users are getting an equal share of system resources, the fairness index should be close to 1. The authors proposed a power allocation scheme to maintain the fairness index. In [17], the authors used the concept of user pairing; the authors paired strong channel users along with weak channel users for the cooperative NOMA system through imperfect CSI and perfect CSI feedback. The authors in [18] presented NOMA-aided precoded spatial modulation (NOMA-PSM) in which researchers combined NOMA with Multiple Input Multiple Output (MIMO). Researchers also presented a comparison with OMA in terms of implementation cost, multiuser interference, spectral efficiency, and performance gain of the system. In [19], the authors proposed full-duplex NOMA relaying-based Device-to-Device (D2D) communication. The authors in [19] proposed the solution for the D2D power allocating problem by presenting a linear fractional programming-based power allocation scheme.

The basic principle of uplink NOMA is presented in [20], the SIC signal detection scheme is utilized at BS, and the power control scheme is used at the user side. The authors investigated the challenges of joint power allocation and subcarrier assignment, and the authors designed a suboptimal solution to increase the sum rate of the NOMA cluster. In [21], the researchers derived the closed-form expressions for outage probability and system capacity for the two-user model of the NOMA uplink system. The researchers investigated the static powers for several users and recognized that a user could be in outage without proper selection of the required data rate. In [22], for the uplink PD-NOMA system, the authors presented an adaptive power control scheme which is based on Evolutionary Game Theory (EGT). To enhance users' throughput or payoffs, the proposed power control scheme allows users to adaptively adjust their transmit power level. SIC is used for signal detection at the receiver. In [23], researchers provided the advantages and challenges of NOMA as a contender scheme in dense networks. The authors compared the performance of NOMA in UL systems. To compare the performance of WSMA-based NOMA and MU-MIMO, researchers presented link-level evaluation results. In [24], the authors provided a foundation to investigate multicell uplink NOMA systems. The authors considered the coverage probability of a NOMA user with high interference at the BS due to a large number of cochannel NOMA transmitting users. The authors in [24] provided closed-form expression of the rate of coverage by characterizing the Laplace transform of the intercluster interference in different SIC scenarios. Afterward, the authors characterized the Laplace transform of the intercluster interference through distance distribution from geometric probability. To evaluate the benefits of NOMA, in 2018, 3GPP considered NOMA as a research icon and provided guidelines to support NOMA, in comparison to the OMA [25]. Table 1 summarizes the review of nonorthogonal multiple access.

3. Nonorthogonal Multiple Access

NOMA is a diverse multiple user access scheme with respect to other established and existing multiple access schemes, i.e., orthogonal multiple access. At the transmitter side, NOMA deliberately introduces intercell and/or intracell interference; therefore, it can utilize nonorthogonal transmission. At the receiver side, successive interference cancellation (SIC) technique is used to decode the desired signal. In comparison with orthogonal multiple access, the complexity of the receiver is increased, but better spectral efficiency can be achieved. Therefore, the fundamental concept of nonorthogonal access is to utilize a receiver with a complex design in trade-off for high spectral efficiency. Therefore, the enhancement in chip processing technology makes the nonorthogonal access scheme possible.

3.1. Power Domain Multiple Access. The NOMA scheme consists of two key technologies. One is power domain NOMA (PD-NOMA), which utilizes efficiently the SIC scheme in order to perform multiuser detection. SIC is a famous physical layer interference cancellation scheme which is used to receive two or more users' signals simultaneously [46]. SIC is sufficiently used in comparison to the existing scheme which causes degradation of the signal. In the SIC scheme, the strongest signals are subtracted from the received combined signal one after another by the SIC receiver; finally, the SIC receiver extracts the desired signal. It is a gradual

Table 1: State-of-the-art review of nonorthogonal multiple access.

Ref.	Objective	Solution approach	Category	Tech.
[26]	Improve reliable detection, maximum diversity gain, and reduce system complexity.	The highest diversity gain with minimum outage probability achieved by cooperative PD-NOMA. User pairing is used as a promising solution to reduce system complexity.	Single carrier power domain	Co-PD-NOMA
[27]	Achieve the fairness performance of the NOMA scheme better than TDMA under perfect and average CSI.	Investigated power allocation techniques that ensure fairness by formulating the research problems as nonconvex optimization.	Single carrier power domain	PD-NOMA
[28]	Further improve the outage performance of MIMO-NOMA.	Improvement achieved by implementing detection and precoding matrices for MIMO-NOMA.	Single carrier power domain	MIMO-NOMA
[29]	Resource allocation algorithm design for multicarrier NOMA systems. Multiple half-duplex uplink and downlink users simultaneously served by a full-duplex base station.	An algorithm is designed for multiple half-duplex uplink and downlink users simultaneously served by a full-duplex base station. Used weighted sum system throughput maximization from the solution of a nonconvex optimization problem.	Multicarrier power domain	MC-NOMA
[30]	For the downlink NOMA system, optimized power allocation and subchannel assignment to increase energy efficiency.	For subchannel multiplexed users, a low-complexity suboptimal algorithm is presented, which comprises power proportional factor determination and energy-efficient subchannel assignment.	Single carrier power domain	PD-NOMA
[31]	Improve the link-level performance of SCMA in highly overloaded scenarios.	Proposed an iterative multiuser SCMA receiver by employing channel coding which uses the coding gain and diversity gain.	Multicarrier code domain	SCMA
[32]	Maximize the mutual information in sparse code multiple access (SCMA).	Maximize the mutual information between continuous output and discrete input using an iterative codebook optimization algorithm.	Multicarrier code domain	SCMA
[33]	Substantially minimize the hurdles of the message passing algorithm (MPA) scheme.	For uplink SCMA systems, a shuffled-message passing algorithm (S-MPA) scheme is proposed, based on a serial message update strategy.	Multicarrier code domain	S-MPA
[34]	Reduce the decoding hurdles of the current message passing algorithm.	Based on list sphere decoding (LSD), a low-complexity decoding algorithm is proposed. The LSD only works with signals inside a hypersphere by evading the extensive search for all possible hypotheses.	Multicarrier code domain	LDS
[35]	Minimizing the hurdles of the SCMA decoding.	Proposed a Monte Carlo Markov Chain- (MCMC-) based SCMA decoder. Benefiting from the linearly increasing complexity of the MCMC method.	Multicarrier code domain	MCMC
[36]	Maximize the sum rate subject to QoS and system-level constraints like power constraints.	Multiple users utilized the same SCMA codebook, and for user signal nonorthogonality, the PD-NOMA scheme is utilized.	Power & code domain	PD-SCMA
[37]	For random signature selection, allowed grant-free transmission to achieve high overloading.	Introduced a blind multiple user detection for MUSA systems by using a special blind detection algorithm.	Single carrier code domain	MUSA
[38]	For the paired users, optimized the modulated symbol mapping.	Performance of MUSA with SIC has been considered by using mirror constellation bit error ratio (BER).	Single carrier code domain	MUSA

Table 1: Continued.

Ref.	Objective	Solution approach	Category	Tech.
[39]	A family of short length complex sequences is selected to permit an easy multiuser interference cancellation.	Successive/parallel interference cancellation with minimum mean square error (MMSE-SIC/PIC) has been investigated for appropriate MUSA receivers.	Single carrier code domain	MMSE-SIC/PIC
[40]	Increase user overloading and minimize multiuser interference.	Enlarge the pool of the spreading sequences by using nonorthogonal dense spreading sequence to increase user overloading and reduce multiuser interference.	Single carrier code domain	MUSA
[41]	To further enlarge the coverage area and improve transmission reliability.	With forward relay and half-duplex decode, an uplink cooperative PDMA (co-PDMA) scheme is suggested.	Single carrier code domain	Co-PDMA
[42]	Increase the performance of PDMA uplink system by using diversity gains and coding potentials.	By using diversity gains and coding potentials, an iterative detection and decoding (IDD) algorithm is developed for an advanced PDMA receiver.	Single carrier code domain	IDD
[43]	Using the cyclic redundancy check (CRC) to avoid the error propagation.	Based on the MMSE channel decoding and detection, a novel iterative decoding and detection algorithm is proposed, called the SIC iterative processing algorithm.	Single carrier code domain	SIC-MMSE
[44]	Proposed the power allocation and pattern assignment in the downlink PDMA system.	To optimize the overall throughput of total users based on the optimum Iterative Water-Filling (IWF) algorithm, a joint pattern assignment and power allocation (JPPA) scheme is offered.	Single carrier code domain	JPPA & IWF
[45]	Improve security by changing the signal's identity.	Physical layer security system is suggested based on constellation scrambling (CS) and multiple parameter weighted fractional Fourier transform (MP-WFRFT).	Single carrier code domain	MP-WFRFT

interference elimination strategy. This type of technique is also used in CDMA to eliminate Multiple Access Interference (MAI). First, the MAI introduced by the user might be eliminated with the help of a signal amplitude recovery process by subtracting the individual user's amplitude one at a time from the received signal. The same process is carried out repetitively to subtract remaining users and to decode the desired signal [47]. Secondly, the PD-NOMA multiplexing scheme uses the power domain technology, that is, power domain multiple access (PDM), which was not used efficiently in previous schemes as used in the PD-NOMA scheme. In the power domain, multiplexing non-orthogonality is deliberately introduced. In fact, power dissimilarity among paired users and implementation of SIC within the power domain ensure that user demultiplexing is concurrent. It is different from the other common methods used previously to control power. Also, an algorithm is needed to be used for power distribution at the base station [48].

Figure 1 illustrates the PD-NOMA system with an SIC computation unit. User Equipment (UE) is uniformly distributed in every cell. With different transmitted power of multiple users in each subband, the base station (BS) performed downlink transmission for multiple users simultaneously.

Multiple single users can be scheduled at the same time for the same subband by implementing the Proportional Fair (PF) scheduling scheme at BS in the PD-NOMA system. The scheduling procedures for users are described in Figure 2 [49]. First, the BS selects a set of users known as the "NOMA candidate user sets," in which total users cannot exceed $N_{\max}$. The selected user set is prepared by using the total number of possible combinations of users within one single cell. Secondly, for every user set, BS allocates the transmission power by using a power allocation scheme. The scheduling metric for the corresponding user set is estimated on behalf of power assignment ratios. Thirdly, with the help of the maximum scheduling metric, the scheduler decides the candidate user sets on each subband for data transmission. Finally, for every allocated subband, the scheduler estimates equivalent signal-to-interference-plus-noise ratios (SINRs) for every single scheduled user. The Coding and Modulation Scheme (CMS) determines the SNR for each user [49].

In PD-NOMA, the total transmit power "P" is divided among multiple users. Let a group of "k" user equipment

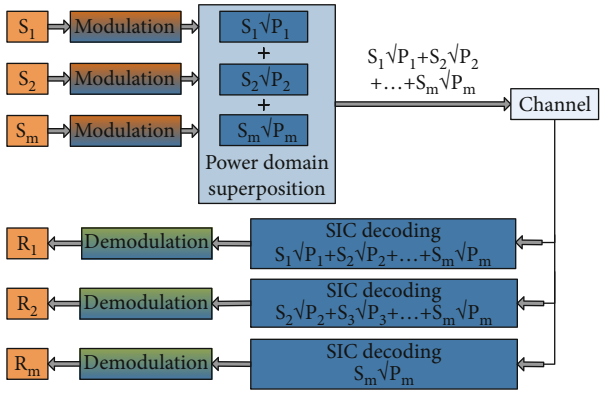


Figure 1: PD-NOMA system with SIC.

(UE) be located under the same base station (BS). Therefore, the fraction of power allocated to the k_{th} user by BS is p_k, where $p_k + \sum_{i=1}^{k-1} p_i = P$. A typical NOMA system model is shown in Figure 1. The received signal at the k_{th} receiver can be written as

$$y_k = s_k \sqrt{p_k}\, g_k + \sum_{i=1}^{k-1} \sqrt{p_i}\, g_k, \tag{1}$$

where $s_k\sqrt{p_k}g_k$ is the received vector of the k_{th} user and $\sum_{i=1}^{k-1}\sqrt{p_i}g_k$ is the interference due to other users.

For PD-NOMA of the downlink system [50], the SINR of the k_{th} user can be written as

$$\text{SINR}_k = \frac{p_k |g_k|^2}{N_0 W + \sum_{i=1}^{k-1} p_i |g_k|^2}. \tag{2}$$

Also, throughput for the k_{th} user can be written as

$$R_k = \text{W}\ \log_2\left(1 + \frac{p_k |g_k|^2}{N_0 W + \sum_{i=1}^{k-1} p_i |g_k|^2}\right), \tag{3}$$

where p_k and p_i are power allocated to the k_{th} and i_{th} users, g_k is the channel gain coefficient of the k_{th} user, N_0 is the noise density, and W is the bandwidth.

3.2. Sparse Code Multiple Access. By means of a typical NOMA technology, SCMA is conceived as the most promising next-generation multiple access scheme for communication networks. SCMA combines Low-Density Spreading (LDS) and multidimensional modulation (MDM) through the SCMA encoding process [51]. In MDM, the numbers of propagating modes have been scaled to the number of available carrier dimensions, as it is considered coded modulation. For a small set of subcarriers, each user spreads its data via a distinguished LDS. Therefore, more than one user can share each subcarrier because there is no exclusivity in the subcarrier allocation. Compared to the total number of users at every subcarrier, a user will have a relatively small number of interferes [52]. In SCMA uplink scenarios, codebook sets are assigned to every user and users select the random codewords from the dedicated codebook sets. All of the users' codewords are multiplexed and shared at the same orthogonal medium, that is, the OFDM subcarrier, illustrated in Figure 3 [53]. Therefore, the multidimensional codebook plays a crucial role in SCMA systems.

In the SCMA system model, a map is defined as an SCMA encoder in which from $\log_2(M)$ to M bits of K-dimensional complex codebooks are available. K represents the spreading factor of the system, which is the length of an SCMA codeword. Sparse vectors are a value of $N(N < K)$ which is the nonzero entries of K-dimensional complex codewords from the codebook. If $N = 2$, two-dimension constellation points can be mapped over $k > 2$ resources. A user could be configured with a codebook by using a contention-based multiple access scheme for uplink transmission [54]. A K-dimensional codeword is carefully chosen from the codebook which is used for mapping a user's data bits for transmission on K radio resources (Figure 4 [55]) which are subcarriers of the OFDMA scheme. Each block of SCMA is carried over K number of OFDMA tones.

Let an SCMA uplink system with "M" numbers of users or codebooks, where "K" is the length of the codeword and "N" number of the nonzero elements are present in each codeword. "d_m" is the distance between m_{th} user "U_m" and BS. Over K subcarriers, M users are multiplexed. The received signal over all subcarriers $y = [y_1, y_2, y_3, \cdots, y_k]^T$ at BS can be written as

$$y = \sum_{m=1}^{M} \sqrt{\frac{p_m}{N}}\ \text{diag}\,(f_m)\ \text{diag}\,(h_m) x_m + w, \tag{4}$$

where "P_m" is the transmission power of user U_m. $x_m = [x_{m1}, x_{m2}, x_{m3}, \cdots, x_{mk}]^T$ is the codeword or transmit symbols of user U_m. The channel coefficient vector for user U_m is $h_m = [h_{m1}, h_{m2}, h_{m3}, \cdots, h_{mk}]^T$.

For the SCMA uplink system [56], the average SNR can be written as

$$\text{SNR} = \sum_{m=1}^{M} \frac{f_{mk} p_m |g_{mk}|^2}{N\, d_m^{\alpha}}. \tag{5}$$

For the SCMA uplink system [56], the average sum rate can be written as

$$R = \sum_{k=1}^{K} E\left(\log_2\left(1 + \sum_{m=1}^{M} \frac{f_{mk} p_m |g_{mk}|^2}{N\, d_m^{\alpha}}\right)\right), \tag{6}$$

where α is the path loss exponent, g_{mk} is the channel gain for the m_{th} user on the k_{th} subcarrier, f_{mk} is the subcarrier index, and p_m is the power of the m_{th} user.

3.3. Multiuser Shared Access. Multiuser shared access (MUSA) uses the advanced successive interference cancellation (A-SIC) scheme and the advantages of good spreading sequences (SS). In SS, the data bit sequence is encoded per

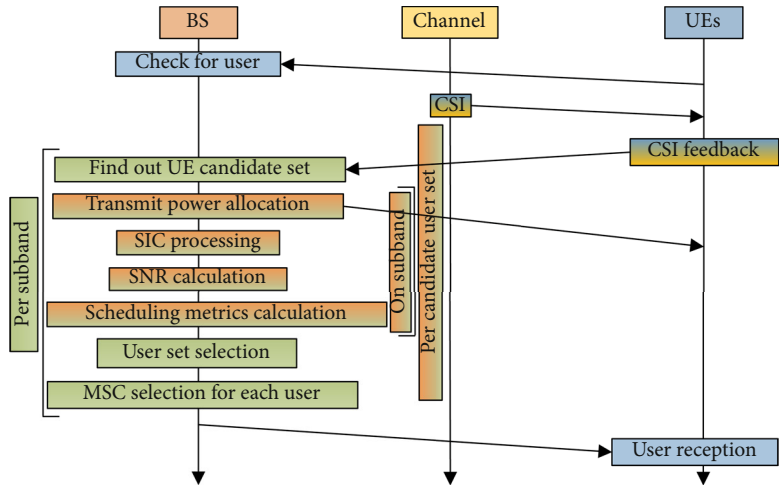


Figure 2: Scheduling algorithm for PD-NOMA.

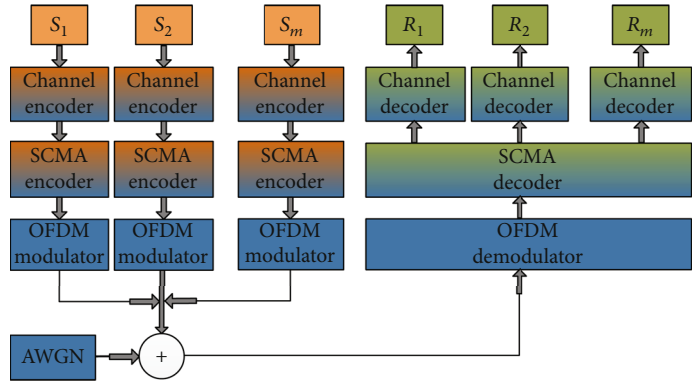


Figure 3: SCMA uplink system with m users.

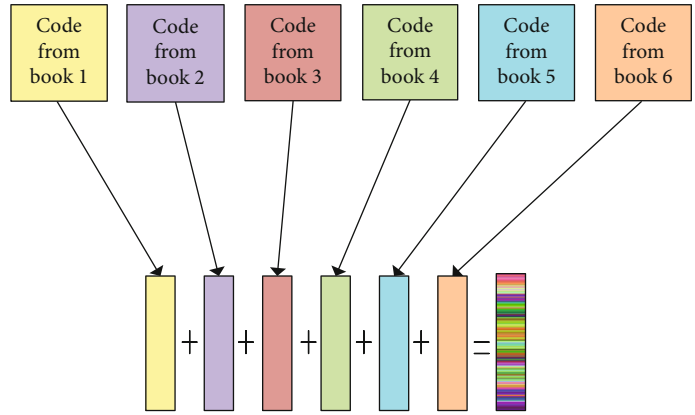


Figure 4: SCMA code multiplexing.

codeword. Therefore, at the same time, the identical number of codewords could be encoded by utilizing the encoder. Afterward, the coded bits are permuted (arranged in all possible ways) through random interleaving patterns. If a large number of interleaving patterns are used, the permuted sequence would be statistically independent. The coded sequence is distributed to each subcarrier after modulation on the quadrature amplitude modulation scheme [57].

MUSA uses special spreading sequences, for spreading multiple users' individual data. After that, the user's spread data is overlapped and transmitted. For recovering and demodulating the data of individual users at reception,

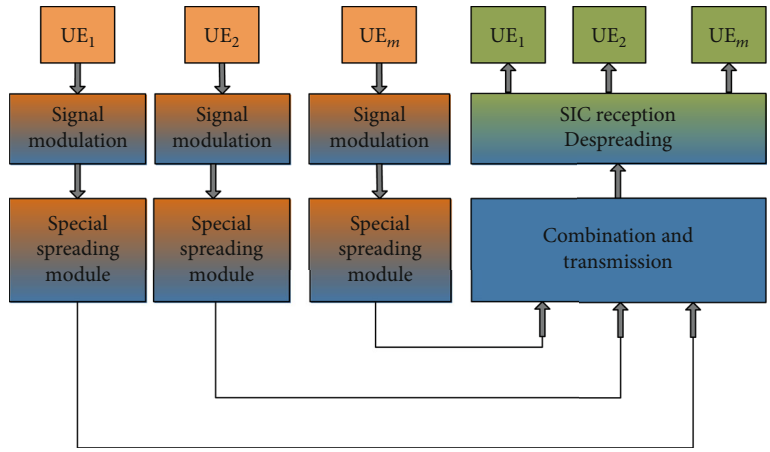


Figure 5: Multiuser shared access.

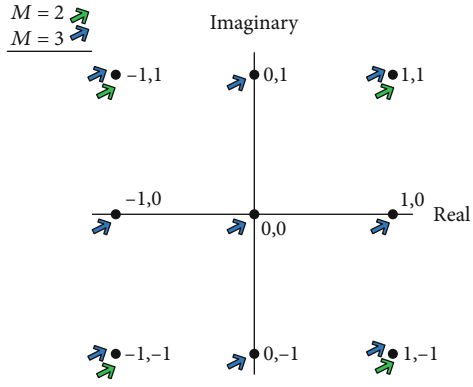


Figure 6: Elements of complex spreading code.

MUSA uses A-SIC receiver. The basic idea is illustrated in Figure 5 [58]. To allow grant-free transmission and maintain a higher overloading factor of users, the nonbinary complex spreading codes at the SIC receiver can be used.

The SIC algorithm is chosen by the SIC receiver to achieve the nonorthogonality between users. It is designed to reduce the power, delay, and complexity if there is large user overloading, and a requirement arises by short spreading codes. A good choice is to use one of the types of the Multicode Complex Domain (MCCD). Due to the design flexibility with the imaginary part and real part, the length of multi-MCCD could be shortened. The sequence with component ± 1 is a type of complex spreading code that might be created as shown in Figure 6 [58]. Moreover, Figure 6 shows the real and imaginary parts of the code for $M = 2$ and $M = 3$. Therefore, before normalization, all elements of the complex spreading codes are just elements of the set $\{1 - i, -1 - i, -1 + i, 1 + i\}$ because the values of the imaginary part and real part contain 1 and -1. The maximum number of existing codes is 4^{L} for the code length L. In the current scenario, the maximum existing code is 256 for the code length of 4, which is not sufficient. Therefore, the existing elements of the set which include the imaginary part and real part need to increase, to improve the number of existing codes, which should be M-ary with $M > 2$.

A preferred selected value of M is 3; the sequence with components 0 and ± 1 is a type of complex spreading code that might be created as shown in Figure 6. Therefore, before normalization, all elements of the complex spreading codes are just elements of the set $\{-1 + i, -1 + 0i, -1 - i, 0 - i, 1 - i, 0 + 0i, 1 + 0i, 1 + i, 0 + i\}$ because the values of the imaginary part and real part contain 0, 1, and -1, that is, a 3-ary [58]. With the help of the new set, 9^L codes could be created, which bring considerable improvement in the number of user access.

For multiuser detection at the receiver, SIC is used in MUSA. Linear conjunction of the received signal detects symbols of multiple users. For the linear system, MMSE is used for detecting users. The received signal can be written as

$$\begin{aligned} y &= hx + n, \\ \tilde{x} &= h^{-1}y - \tilde{n}, \end{aligned} \tag{7}$$

where "x" is the composite transmitted signal, "h" is the channel coefficient matrix, and "n" is a complex noise sample of Gaussian noise with zero mean and variance "σ."

To detect the signal of each user at the receiver, compute the inverse of channel matrix "h^{-1}"; by this inverse, we get the estimated signal "$\tilde{x}$." The MMSE weight matrix can be written as

$$W_{\mathrm{MMSE}} = \left(h^H h + \sigma^2 I\right)^{-1} h^H, \tag{8}$$

where "I" is the identity matrix. Now, we get

$$\tilde{x} = W_{\mathrm{MMSE}}\, y. \tag{9}$$

The SINR at the i_{th} antenna of the MUSA uplink system [59] can be formulated as

$$\mathrm{SINR}_i = \frac{E_x |w_i h_i|^2}{E_x \sum_{l \neq i} |w_i h_l| + \sigma |w_i|^2}. \tag{10}$$

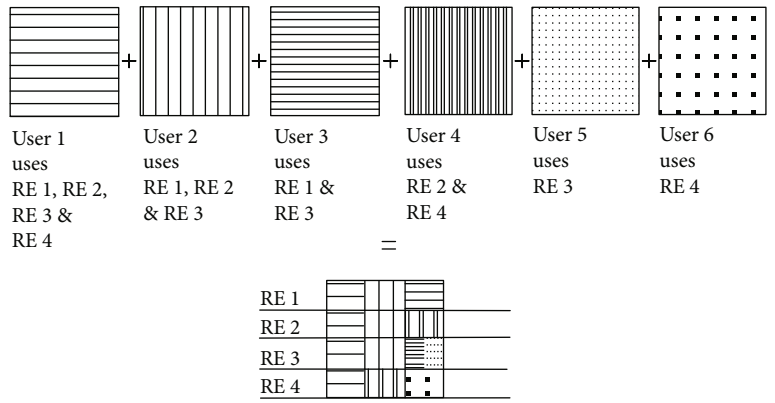


Figure 7: PDMA pattern for 4 REs used by 6 users.

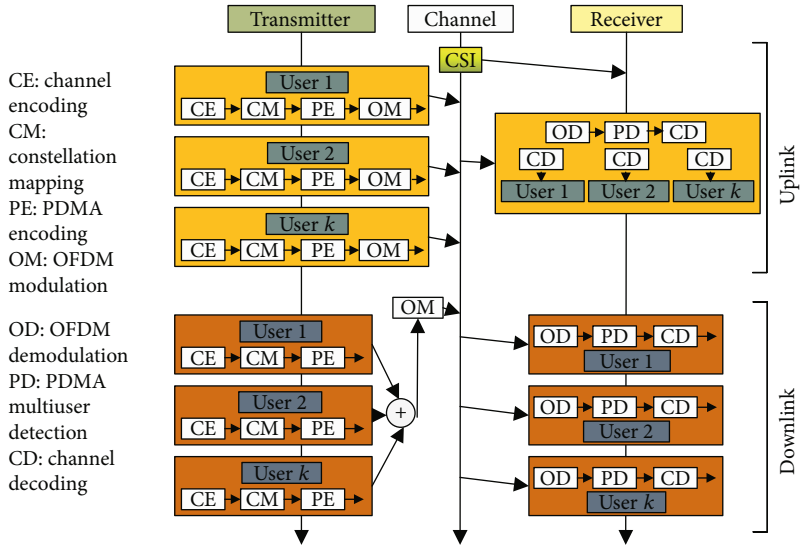


Figure 8: PDMA system model.

Also, throughput at the i_{th} antenna can be written as

$$R_i = \log_2\left(1 + \frac{E_x|w_i h_i|^2}{E_x\sum_{l\neq i}|w_i h_l| + \sigma|w_i|^2}\right), \tag{11}$$

where E_x is the transmitted signal energy, i is the number of transmitted antennas, h_i is the i_{th} column of the channel matrix, and w_i is the i_{th} row of the weight matrix. The weight matrix is constructed by using MMSE technique.

3.4. Pattern Division Multiple Access. Pattern Division Multiple Access (PDMA) is well known as an emerging nonorthogonal multiple access technique based on SIC Amenable Multiple Access (SAMA) technology [60]. PDMA utilizes Low Complexity Quasi-ML (LCQ-ML) SIC detection [61] at the reception and holistic/combined scheme of SIC-Amenable (SIC-A) pattern at the transmission side. An example of the PDMA pattern with resource mapping is shown in Figure 7 [62]. On four resource elements (REs), six users are multiplexed. First of all, the single PDMA pattern is allotted to a single user. All four REs in the cluster are used for user1's data mapping, the first three REs are used for user2, the first and third REs are used for user3, the second and fourth REs are used for user4, the third RE is used for user5, and the fourth RE is used for user6. For all six users, the order of transmission diversity is 4, 3, 2, 2, 1, and 1 [62].

Different users in PDMA are separated at the transmitter through a nonorthogonal character of pattern with various domains, for example, code, space, and power domain. Particularly, at the receiver side, multiple users consist of an irregular diversity degree in order to perform SIC amenable detection. After SIC amenable detection, users can acquire an equivalent diversity degree (Figure 8 [62]). Therefore, in PDMA, steadiness between multiplexing and diversity degree can be achieved [62].

At the receiver side, BS receive signal of "N" resource blocks. The received signal at the m_{th} antenna of BS is $y_m = [x_{1,m}, x_{2,m}, x_{3,m}, \cdots, x_{N,m}]^T$. At the m_{th} antenna, the received signal of the nth resource block can be written as

$$y_{n,m} = \sum_{k=1}^{K} H_{\text{PDMA}}(n,k) h_{nk,m} \sqrt{P_{nk}} x_k + w_{n,m}, \tag{12}$$

Table 2: Feature of different NOMA schemes.

Type	Advantage	Disadvantage	Key feature
PD-NOMA	(i) Is not affected by apparent near-far (ii) 20% uplink spectral efficiency increase (iii) 30% downlink throughput increase [64]	(i) High receiver complexity needs improvement in chip technology (ii) Power domain multiplexing is in the research phase (iii) SIC increases the system signalling overhead	(i) For user multiplexing, PD-NOMA utilized the power domain multiple access (ii) At receiver, SIC scheme is used (iii) Take advantage of different channel conditions
SCMA	(i) Three times increase in spectral efficiency (ii) 2.8 times uplink system capacity upgrade (iii) 8% and 5% increase in coverage gain and downlink throughput, respectively [64]	(i) Optimization and design of the code are difficult (ii) Increased interference between users (iii) High-dimensional modulation (HDM) required	(i) SCMA utilizes sparse spreading sequence, based on LDS-OFDM (ii) Spreading with low-density signatures and bit-to-constellation mapping are combined in SCMA (iii) Codebooks are created by multidimensional constellation. Users' codewords are taken from codebooks
MUSA	(i) Block Error Rate (BLER) is low (ii) Huge number of users' access is supported (iii) Spectral efficiency increased by 1.5 times [64]	(i) Interuser interference is increased (ii) Spread symbol design is challenging	(i) MUSA is an upgraded scheme of CDMA via code domain multiplexing (ii) At the transmitter, MUSA achieved higher overloading through low-correlation spreading sequences (iii) SIC is performed at the receiver side, to decode superimposed symbols
PDMA	(i) 2-3 times uplink system capacity increased (ii) 1.5 times spectral efficiency increase in downlink system	(i) The pattern optimization and design are challenging (ii) Increase interference between users	(i) Nonorthogonal patterns are used in PDMA (ii) Multiplexing is achieved in space domain, power domain, code domain, and their composite domain (iii) Code domain multiplexing is similar to SCMA (iv) Low Complexity Quasi-ML SIC detection is utilized in PDMA

where "H_{PDMA}" is the PDMA pattern matrix, "P_{nk}" is the transmitted power of the k_{th} user at the n_{th} resource block, and "x_k" is the transmitted signal from the k_{th} user to BS. "$w_{n,m}$" is complex additive white Gaussian noise at the n_{th} resource block in m_{th} receiving antenna.

Using [63], the SINR at the m_{th} receiving antenna in n_{th} resource block of the k_{th} user for the PDMA system can be written as

$$\mathrm{SINR}_{nk,r} = \frac{P_{nk} H_{\mathrm{PDMA}(n,k)} \left| h_{nk,r} \right|^2}{\sigma^2 + \sum_{j \neq k}^{K} P_{nj} H_{\mathrm{PDMA}(n,j)} \left| h_{nj,r} \right|^2}. \quad (13)$$

Also, throughput for k_{th} user can be written as

$$R_k = \sum_{r=1}^{N_r} \sum_{n=1}^{N} \log_2 \left(1 + \frac{P_{nk} H_{\mathrm{PDMA}(n,k)} \left| h_{nk,r} \right|^2}{\sigma^2 + \sum_{j \neq k}^{K} P_{nj} H_{\mathrm{PDMA}(n,j)} \left| h_{nj,r} \right|^2} \right), \quad (14)$$

where P_{nk} is the power of k_{th} user in n_{th} resource block, H is the PDMA pattern matrix, N_r indicates the number of receiving antennas, $h_{nk,r}$ is the channel gain coefficient, and σ^2 is the AWGN density. Moreover, Table 2 highlighted the main key features, advantages, and disadvantages of major categories of NOMA schemes.

4. Summary

PD-NOMA utilizes nonorthogonal transmission among the users as compared to CDMA and OFDMA. PD-NOMA does not have apparent near-far problem compared to 3G. Similarly, the MAI complications are not challenging in PD-NOMA. PD-NOMA has a simple way to respond to multiple links and changing conditions of link by applying Adaptive Modulation and Coding (AMC) particularly in high-speed mobile environments [65]. Therefore, PD-NOMA does not need a highly accurate feedback signal or channel state information (CSI) from the user end. In PD-NOMA, multiple users share the same channel; therefore, spectral efficiency is increased at the unchanged transmission rate compared to 4G [66, 67]. In contrast, from a technical implementation aspect, PD-NOMA is still facing several challenges. Initially, implementation needs enhancement in chip technology from signal processing aspect because the nonorthogonal decoder is complex in design. Furthermore, the power domain multiplexing scheme is under the research phase and has a long way to go [64]. Technologies used in different types of nonorthogonal access schemes are presented in Table 3.

As an innovative multiple-access modulation technique, SCMA offered several improvements, for example, multidimensional constellation shaping gain along with benefits of CDMA and LDS. The link-level performance of SCMA in

Table 3: Technology used by NOMA.

Type	PD-NOMA	SCMA	MUSA	PDMA
PDM	●			
SIC	●		●	●
LDS		●		
HDM		●		
MPA		●		
MCCD			●	
MLD				●

highly overloaded scenarios can be achieved by employing channel coding which uses the coding gain and diversity gain. Although the structure of the code is well defined, optimization and design of the code are problematic [64]. To reduce the decoding hurdles of the message passing algorithm, LSD based on a low-complexity decoding algorithm is used in SCMA, in which the LSD only works with signals inside a hypersphere by evading the extensive search for all possible hypotheses.

Uplink access in MUSA utilizes an advanced complex multidomain code structure and multiuser decoding on the basis of SIC. In order to confirm that unlimited reliable access at the same frequency-time slot for multiple users, MUSA makes the procedure of resource allocation simpler in the access scheme. So that MUSA significantly cuts the access time, makes the system implementation simpler, and minimizes energy utilization. MUSA downlink access utilizes superposition symbol expansion and superposition coding scheme, to offer better capacity as compared to downlink transmission provided by the OMA. Also, uplink access in MUSA offers to decrease the energy consumption and make the implementation of user terminal simpler which is the same as MUSA downlink [64].

PDMA can increase the performance of the spectrum utilization for the downlink system by 1.5 times and increases capacity in the uplink system by 2-3 times [68]. To improve the security in the PDMA system, constellation scrambling with MP-WFRFT is utilized. To avoid error propagation, MMSE channel decoding and detection-based SIC iterative processing is used in the PDMA system. Cooperative PDMA is used with forward relay and half-duplex decode in order to further enlarge the coverage area and improve transmission reliability. On the other side, PDMA needs to encounter some important challenges to be resolved in upcoming applications. These include designing simpler receivers, design patterns at the transmission end to discriminate users without difficulty, and combine MIMO with PDMA in order to develop space domain coding design, etc.

5. Discussion

System sum rate performance versus the total number of users of PD-NOMA, SCMA, MUSA, and PDMA is illustrated in Figure 9 from [69]. SCMA has been confirmed by theory and in lab tests that SCMA has a better sum rate among all major four categories while the complexity of SCMA is bigger than the PD-NOMA due to the code domain. SCMA is capable of achieving coding gains and improved shaping. SCMA allows a fixed number of resource blocks to each user while PDMA allows a changeable number of resource blocks to each user, since, in PDMA, the user data rate is different, which results in degradation of the system sum rate. PD-NOMA and MUSA both utilized SIC, but PD-NOMA utilized power domain multiple access and MUSA utilized a special spread sequence to spread the user's data symbols. Therefore, MUSA has a better sum rate than PD-NOMA.

A comparison of the average aggregate energy efficiency of PD-NOMA and SCMA schemes against the number of its users is illustrated in Figure 10 from [70]. SCMA outperforms its counterparts due to nonorthogonality with high overloading. Therefore, in SCMA, further access of users is achieved with low energy consumption. PD-NOMA also utilized nonorthogonal access; more users can be employed on less numbers of resources, but due to power domain access, overloading cannot be achieved. On the other hand, OFDMA underperforms because OFDMA is an orthogonal scheme in which users are restricted by orthogonal resources.

Figure 11 illustrated the bit error rate of PD-NOMA, SCMA, PDMA, and MUSA uplink systems in the Rayleigh fading channel from [71]. For performance comparison of the PDMA and SCMA, the same factor graph is used with QPSK modulation. The number of orthogonal resources is 4, and the number of symbols which are transmitted is 6. Therefore, the resulting overloading factor becomes 150%. With the help of [72], the codebooks are designed in SCMA. Pseudorandom sequences whose image and real values are obtained from set {-1, 0, 1} are used to generate spreading sequences for MUSA, and nonorthogonal patterns are designed accordingly [73] for PDMA.

The SCMA uplink system has high-quality BER performance among all code domains, as shown in Figure 11. However, the BER performance of MUSA and PDMA is very similar and lesser than SCMA. The effect of error propagation of the SIC receiver on the system performance is the main reason for performance degradation of MUSA and PDMA. Nevertheless, when PDMA utilized the unchanged factor graph as used in SCMA, SCMA still has better BER performance than PDMA. The reason for this performance improvement is because of the near-optimal strategy of sparse codewords. On the other hand, the code domain NOMA scheme achieves a better system sum rate than the power domain NOMA scheme; therefore, PD-NOMA managed poor performance among all.

6. Recent Development

The race for developing 5G technology has integrated NOMA with different communication technologies, such as NOMA-based communication for the Tactile Internet, NOMA for D2D communication, cognitive radio nonorthogonal multiple access, and SWIPT-NOMA-based HetNets. A brief review of recent developments in NOMA is as follows.

The authors in [74] presented NOMA-based application-specific communication for the Tactile Internet, by which

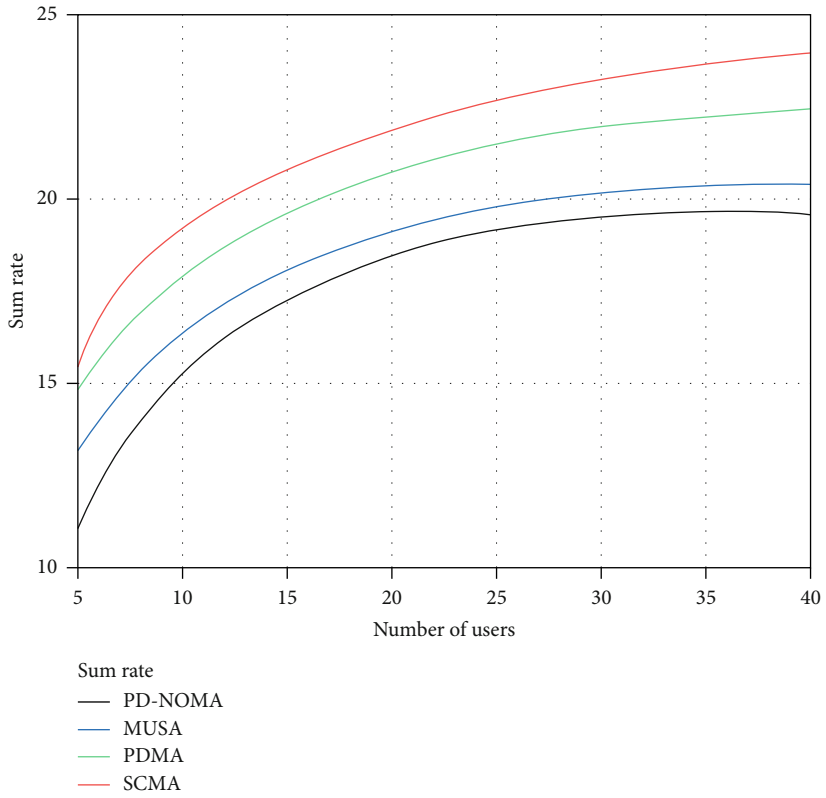


Figure 9: Sum rate of different non-orthogonal schemes.

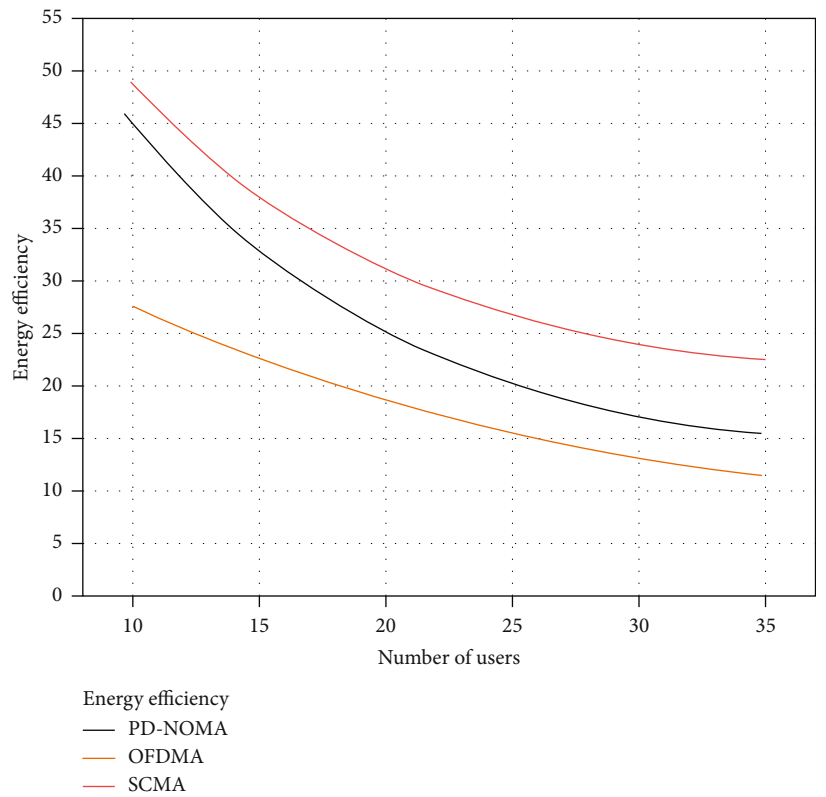


Figure 10: Energy efficiency of different orthogonal and non-orthogonal schemes.

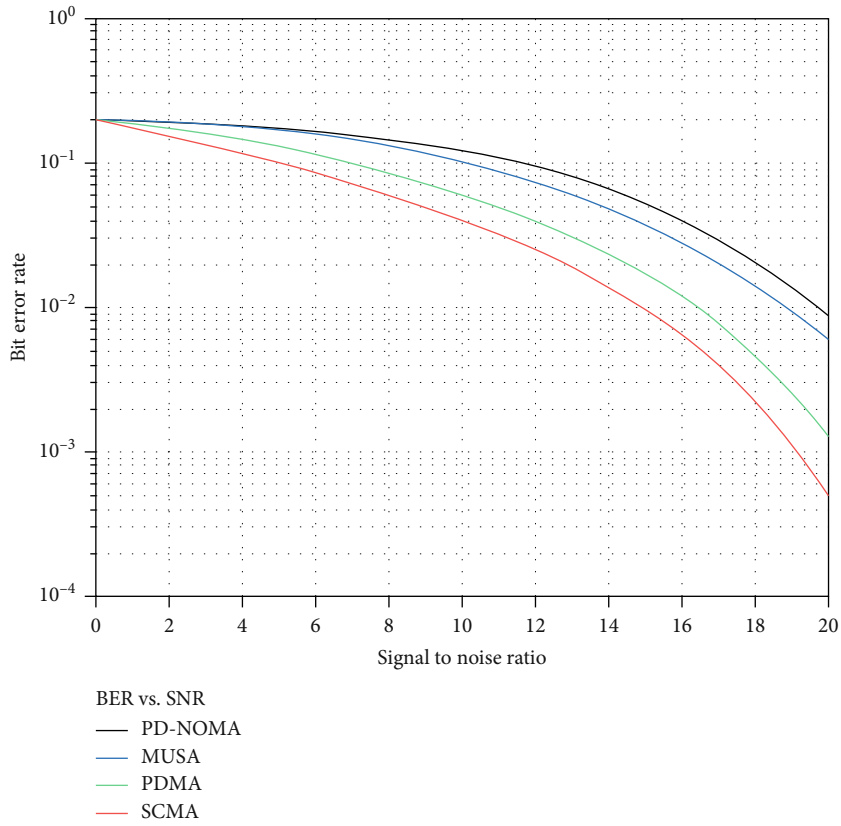


Figure 11: BER performance of different non-orthogonal schemes.

heterogeneity can be achieved in 5G networks. Tactile Internet allows nonorthogonal resource sharing from a pool of massive machine-type communications (mMTC), ultrareliable low latency communications (URLLC), enhanced mobile broadband (eMBB), and critical machine-type communication (cMTC) devices to a shared the same base station. The authors in [74] summarized many different types of NOMA and their appropriateness for low latency Tactile Internet-based applications. Additionally, the authors in [74] presented a sample case of a healthcare-based network and explained how in the healthcare domain NOMA-based architecture can be utilized for low latency networks.

In [75], the authors used NOMA at the D2D transmitter to improve the spectral efficiency of the network. The authors proposed in [75] the Tactile Internet Driven Delay Assessment for D2D communication (DIYA) scheme to resolve the issue of interference and delay from the neighboring nodes in two-hop transmission. In the first phase at relays (intermediate nodes), a full duplex communication is used for the first and second hop transmission concurrently, at the same time interval. Afterward, at D2D transmitter transmission rate is improved using Tactile Internet-based communication. In the second phase, to reduce the cochannel interference and increase the throughput of the cell edge users, pricing-based 3D matching is proposed by authors. Furthermore, authors in [75] used successive convex approximation (SCA) with less complexity in order to optimize the power of the D2D transmitter. SCA converts the nonconvex optimization problem of power control and subchannel allocation into the convex problem.

In [76], to improve the sum rate of the femtocell users, researchers proposed a joint power control and channel allocation algorithm by utilizing cognitive radio nonorthogonal multiple access (CR-NOMA) at the femtobase station. The authors used the channel gain difference among weak and strong users' pairs in the proposed algorithm. This reduces the interference between NOMA users and improves channel utilization. Furthermore, to provide the QoS for weak users, the authors differentiated the odd and even numbers of users in a femtocell. The aforementioned scheme, OMA, is utilized to obtain a preset data rate by a greedy channel allocation algorithm.

The authors in [77] presented a subchannel assignment scheme for SWIPT-NOMA-based HetNets with imperfect CSI for the downlink system. Furthermore, the many-to-many matching theory is presented by authors to formulate the subchannel assignment. Considering imperfect CSI, the authors in [77] presented the energy-efficient subchannel assignment as a nonconvex probabilistic optimization problem. The many-to-many matching theory is utilized by authors to deal with this problem. The authors used SWIPT and NOMA with macrousers and pico-/femtobase station, in which multiple users served by NOMA simultaneously and SWIPT harvest energy from the radio frequency signals. Both techniques increase the energy efficiency of the network.

In [78], to maximize the sum rate and spectral efficiency of femtousers with guaranteed QoS, the authors investigated the NOMA transmission with 5G enabled cognitive femtocell. To reduce the NOMA interference among multiple femtousers, a pairing algorithm between weak and strong users has been presented by authors. The authors also calculated the sum rate for an even/odd number of femtousers in order to achieve a higher data rate.

7. Research Challenges

In this article, we investigated major categories of NOMA, contributions of NOMA in enabling the 5G network, integration with different communication technologies, and recent research trends. Conversely, there are still many challenges which should be solved further to improve the performance of NOMA systems. We present various significant challenges of NOMA and specify potential research.

7.1. Imperfect SIC Cancellation. In practical circumstances during SIC processing, some residual interference left; the successive interference cancellation is mostly imperfect. Therefore, in theoretical analysis, we have to consider this imperfect cancellation aspect. Furthermore, error propagation in SIC is also a major problem. This indicates that when the higher-order user has been decoded erroneously, the error will sequentially propagate to lower-order users.

7.2. Imperfect CSI. The current research works on NOMA presume a perfect CSI to implement multiuser interference cancellation at the user receiver or resource allocation at BS. However, perfect CSI is impossible in practical scenarios. Therefore, real-time NOMA systems work with channel estimation errors. In theoretical analysis of NOMA, there is a need to consider channel estimation errors and imperfect CSI.

7.3. Design of Spreading Sequences or Codebooks. In SCMA, codebook design is still an issue particularly for outsized higher-dimensional codebooks. For further performance improvement of SCMA, the joint design of the factor graph matrix and constellation construction is required. For this, advance multidimension constellation is needed. Furthermore, to improve link adaptation, the design scheme for the case that all the overloaded users have different codebook sizes (transmission rate) needs to be investigated. Moreover, to determine the performance and capacity under practical scenarios, researchers have to consider error propagation in codebook allocation in theoretical analysis.

7.4. Receiver Complexity. As compared to OMA schemes, in NOMA, SIC needs additional implementation complexity, because the SIC receiver has to detect and cancel other users' signals prior to detecting its own signal. Moreover, as the number of users in the cell increases, the receiving complexity also increases. Therefore, a high-performance nonlinear detection algorithm is required at each stage of SIC for error-free propagation.

7.5. Heterogeneous Networks. A wireless network containing nodes with different coverage sizes and transmission powers is known as a heterogeneous network (HetNet). The HetNet has capability in terms of coverage and capacity with reduced energy consumption for future wireless networks. Real-time NOMA allows sharing of resources for different types of networks. To improve system throughput of heterogeneous networks, heterogeneous collaborative communication schemes with NOMA can be investigated.

7.6. Further Challenges. Several further challenges of NOMA systems must also be addressed, including signal design and channel estimation, maintaining system scalability, for multicarrier NOMA the reduction of the PAPR, the difficulties of channel-quality feedback design, and flexible configuration of multiple access schemes. It is accepted by researchers that by addressing these challenges NOMA will further improve.

7.7. Future Trends. As the expected new round of developments, NOMA has received huge attention and active input from researchers, and its development is very rapid. Some future research trends are NOMA in large-scale heterogeneous networks, full-/half-duplex user relaying in NOMA systems, NOMA for wireless powered IoT networks, NOMA-based massive MTC networks, adaptive NOMA/OMA mode-switching, NOMA systems over κ-μ shadowed fading channels, in large-scale underlay cognitive radio networks, and NOMA with spatial modulation.

8. Conclusion

Currently, nonorthogonal awareness has been significantly useful in the modern developments in the 5G multiple access scheme. Herein, favorable nonorthogonal multiple access technique for 5G mobile communications is reassessed and compared on the basis of their advantages, disadvantages, and key features and their future development. Furthermore, we considered the performance of significant NOMA schemes, i.e., PD-NOMA, SCMA, MUSA, and PDMA in Rayleigh fading channels. This comparison research reveals the performance of different NOMA schemes. With the in-depth review of their basic working principle, system model, and performance, 5G key multiple access technique will be progressively understood, enabling us to arrive at the essential stage of formulation and standardization.

Conflicts of Interest

The authors declared that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work would be funded by authors.

References

[1] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE*

Communications Surveys & Tutorials, vol. 20, no. 3, pp. 2294–2323, 2018.

[2] Y. Chen, A. Bayesteh, Y. Wu et al., "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 19–27, 2018.

[3] A. Benjebbour, K. Saito, and Y. Kishiyama, "Experimental trials on non-orthogonal multiple access," in *In Multiple Access Techniques for 5G Wireless Networks and Beyond*, pp. 587–607, Springer,, 2019.

[4] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 332–336, London, UK, 2013.

[5] Z. Yuan, G. Yu, and W. Li, "Multi-user shared access for 5G," *Telecommunication Network Technology*, vol. 5, no. 5, pp. 28–30, 2015.

[6] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616–1626, 2008.

[7] X. Dai, Z. Zhang, B. Bai, S. Chen, and S. Sun, "Pattern division multiple access: a new multiple access technology for 5G," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 54–60, 2018.

[8] Z. Wu, K. Lu, C. Jiang, and X. Shao, "Comprehensive study and comparison on 5G NOMA schemes," *IEEE Access*, vol. 6, pp. 18511–18519, 2018.

[9] L. Zhang, M. Xiao, G. Wu, M. Alam, Y. C. Liang, and S. Li, "A survey of advanced techniques for spectrum sharing in 5G networks," *IEEE Wireless Communications*, vol. 24, no. 5, pp. 44–51, 2017.

[10] Y. Wang, B. Ren, S. Sun, S. Kang, and X. Yue, "Analysis of non-orthogonal multiple access for 5G," *China Communications*, vol. 13, no. 2, pp. 52–66, 2016.

[11] Y. El Gholb, N. El Amrani, E. Idrissi, and H. Ghennioui, "5G: an idea whose time has come," *International Journal of Scientific and Engineering Research*, vol. 8, no. 3, 2017.

[12] A. P. Singh, S. Nigam, and N. K. Gupta, "A study of next generation wireless network 6G," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 1, 2007.

[13] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama, and T. Nakamura, "Non-orthogonal multiple access (NOMA): concept, performance evaluation and experimental trials," in *2015 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, pp. 1–6, Marrakech, Morocco, 2015.

[14] J. Kim, J. Lee, D. Kim, and Y. Choi, "System-level performance evaluation for non-orthogonal multiple access in coordinated direct and relay transmission," in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1296–1298, Jeju, South Korea, 2017.

[15] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, 2016.

[16] H. Xing, Y. Liu, A. Nallanathan, Z. Ding, and H. V. Poor, "Optimal throughput fairness tradeoffs for downlink non-orthogonal multiple access over fading channels," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3556–3571, 2018.

[17] S. Arzykulov, T. A. Tsiftsis, G. Nauryzbayev, and M. Abdallah, "Outage performance of cooperative underlay CR-NOMA with imperfect CSI," *IEEE Communications Letters*, vol. 23, no. 1, pp. 176–179, 2019.

[18] P. Yang, Y. Xiao, M. Xiao, and Z. Ma, "NOMA-Aided precoded spatial modulation for downlink MIMO transmissions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 729–738, 2019.

[19] L. Song, S. Li, and Y. Sun, "Power allocation for full-duplex NOMA relaying based underlay D2D communications," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 1, pp. 16–33, 2019.

[20] L. Anxin, A. Benjebbour, X. Chen, H. Jiang, and H. Kayama, "Uplink non-orthogonal multiple access (NOMA) with single-carrier frequency division multiple access (SC-FDMA) for 5G systems," *IEICE Transactions on Communications.*, vol. E98.B, no. 8, pp. 1426–1435, 2015.

[21] B. Xia, J. Wang, K. Xiao, Y. Gao, Y. Yao, and S. Ma, "Outage performance analysis for the advanced SIC receiver in wireless NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6711–6715, 2018.

[22] S. Riaz, J. Kim, and U. Park, "Evolutionary game theory-based power control for uplink NOMA," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 6, pp. 2697–2710, 2018.

[23] B. Makki, K. Chitti, A. Behravan, and M. S. Alouini, "A survey of NOMA: current status and open research challenges," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, 2020.

[24] H. Tabassum, E. Hossain, and M. J. Hossain, "Modeling and analysis of uplink non-orthogonal multiple access (NOMA) in large-scale cellular networks using Poisson cluster processes," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3555–3570, 2017.

[25] P. Merias and J. M. Meredith, *Study on non-orthogonal multiple access (NOMA) for NR*, 3GPP, Sophia Antipolis, France, 2018, Rep. TR 38.812.

[26] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1462–1465, 2015.

[27] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1647–1651, 2015.

[28] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 537–552, 2016.

[29] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1077–1091, 2017.

[30] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, 2016.

[31] Y. Wu, S. Zhang, and Y. Chen, "Iterative multiuser receiver in sparse code multiple access systems," in *2015 IEEE International Conference on Communications (ICC)*, pp. 2918–2923, London, UK, 2015.

[32] C. Dong, G. Gao, K. Niu, and J. Lin, "An efficient SCMA codebook optimization algorithm based on mutual information

maximization," *Wireless Communications and Mobile Computing*, vol. 2018, 13 pages, 2018.

[33] Y. Du, B. Dong, Z. Chen, J. Fang, and L. Yang, "Shuffled multiuser detection schemes for uplink sparse code multiple access systems," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1231–1234, 2016.

[34] F. Wei and W. Chen, "Low complexity iterative receiver design for sparse code multiple access," *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 621–634, 2017.

[35] J. Chen, Z. Zhang, S. He, J. Hu, and G. E. Sobelman, "Sparse code multiple access decoding based on a Monte Carlo Markov chain method," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 639–643, 2016.

[36] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi, and H. Pishro-Nik, "A new multiple access technique for 5G: power domain sparse code multiple access (PSMA)," *IEEE Access*, vol. 6, pp. 747–759, 2018.

[37] Z. Yuan, C. Yan, Y. Yuan, and W. Li, "Blind Multiple User Detection for Grant-Free MUSA without Reference Signal," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Toronto, ON, USA, 2017.

[38] Y. Xu, G. Wang, L. Zheng, R. Liu, and D. Zhao, "BER performance evaluation of downlink MUSA over Rayleigh fading channel," in *International Conference on Machine Learning and Intelligent Communications*, pp. 85–94, Springer, 2017.

[39] 3GPP Document R1-166404, "Receiver details and link performance for MUSA," in *3GPP TSG RAN WG1 Meeting No. 86*, Gothenburg, Sweden, 2016.

[40] N. Ye, H. Han, L. Zhao, and A. H. Wang, "Uplink non-orthogonal multiple access technologies toward 5G: a survey," *Wireless Communications and Mobile Computing*, vol. 2018, 26 pages, 2018.

[41] W. Tang, S. Kang, and B. Ren, "Performance analysis of cooperative pattern division multiple access (co-PDMA) in uplink network," *IEEE Access*, vol. 5, pp. 3860–3868, 2017.

[42] B. Ren, X. Yue, W. Tang et al., "Advanced IDD receiver for PDMA uplink system," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–6, Chengdu, China, 2016.

[43] D. Kong, J. Zeng, X. Su, L. Rong, and X. Xu, "Multiuser detection algorithm for PDMA uplink system based on SIC and MMSE," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–5, Chengdu, China, 2016.

[44] J. Zeng, B. Liu, and X. Su, "Joint pattern assignment and power allocation in PDMA," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Toronto, ON, USA, 2017.

[45] B. Ren, Y. Wang, X. Dai, K. Niu, and W. Tang, "Pattern matrix design of PDMA for 5G UL applications," *China Communications*, vol. 13, Supplement2, pp. 159–173, 2016.

[46] G. Mazzini, "Power division multiple access," in *ICUPC '98. IEEE 1998 International Conference on Universal Personal Communications. Conference Proceedings (Cat. No.98TH8384)*, pp. 543–546, Florence, Italy, Italy, 1998.

[47] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access. Vehicular Technology Conference (VTC Spring)," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*,, pp. 1–5, Dresden, Germany, 2013.

[48] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 770–774, Naha, Japan, 2013.

[49] X. Chen, A. Beiijebbour, A. Li, H. Jiang, and H. Kayama, "Consideration on successive interference canceller (SIC) receiver at cell-edge users for non-orthogonal multiple access (NOMA) with SU-MIMO," in *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 522–526, Hong Kong, China, 2015.

[50] G. Gui, H. Sari, and E. Biglieri, "A new definition of fairness for non-orthogonal multiple access," *in IEEE Communications Letters*, vol. 23, no. 7, pp. 1267–1271, 2019.

[51] B. Di, L. Song, and Y. Li, "Radio resource allocation for uplink sparse code multiple access (SCMA) networks using matching game," in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kuala Lumpur, 2016.

[52] M. Hussain and H. Rasheed, "A Computational Power Allocation Scheme for Fair NOMA Downlink System," *Journal of Information Communication Technologies and Robotic Applications (JICTRA)*, vol. 9, no. 1, pp. 73–79, 2018.

[53] S. Zhang, K. Xiao, B. Xiao et al., "A capacity-based codebook design method for sparse code multiple access systems," in *2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)*,, pp. 1–5, Yangzhou, China, 2016.

[54] Y. Zhou, H. Luo, R. Li, and J. Wang, "A dynamic states reduction message passing algorithm for sparse code multiple access," in *2016 Wireless Telecommunications Symposium (WTS)*, pp. 1–5, London, UK, 2016.

[55] K. Au, L. Zhang, H. Nikopour et al., "Uplink contention based scma for 5G radio access," in *2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 900–905, 2014.

[56] Z. Yang, J. Cui, X. Lei, Z. Ding, P. Fan, and D. Chen, "Impact of factor graph on average sum rate for uplink sparse code multiple access systems," *IEEE Access*, vol. 4, pp. 6585–6590, 2016.

[57] K. Hyukjoon, L. Jungwon, and K. Inyup, "Successive interference cancellation via rank-reduced maximum a posteriori detection," *Communications, IEEE Transactions on*, vol. 61, no. 2, pp. 628–637, 2013.

[58] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-user shared access for Internet of things," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Nanjing, China.

[59] E. M. Eid, M. M. Fouda, A. S. T. Eldien, and M. M. Tantawy, "Performance analysis of MUSA with different spreading codes using ordered SIC methods," in *2017 12th International Conference on Computer Engineering and Systems (ICCES)*, pp. 101–106, Cairo, Egypt, 2017.

[60] X. Dai, "Successive interference cancellation amenable space-time codes with good multiplexing-diversity Tradeoffs," *Wireless Personal Communication*, vol. 55, no. 4, pp. 645–654, 2010.

[61] X. Dai, S. Sun, and Y. Wang, "Reduced-complexity (quasi-) maximum- likelihood detectors with no performance degradation for S-QAM modulated MIMO systems," *Wireless Personal Communication*, vol. 66, no. 4, pp. 613–627, 2012.

[62] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access (PDMA) - a novel non-orthogonal multiple access for 5G radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, 2019.

[63] J. Zeng, D. Kong, X. Su, L. Rong, and X. Xu, "On the performance of pattern division multiple access in 5G systems," in

2016 8th International Conference on Wireless Communications & Signal Processing (WCSP), pp. 1–5, Yangzhou, China, 2016.

[64] Y. Tao, L. Liu, S. Liu, and Z. Zhang, "A survey: several technologies of non-orthogonal transmission for 5G," *China Communications*, vol. 12, no. 10, pp. 1–15, 2015.

[65] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1746–1755, 2004.

[66] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC).*, pp. 611–615, London, UK, 2013.

[67] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, pp. 781–785, Barcelona, Spain, 2014.

[68] M. Hussain and H. Rasheed, "Performance of Orthogonal Beamforming with NOMA for Smart Grid Communication in the Presence of Impulsive Noise," *Arabian Journal for Science and Engineering (AJSE)*, vol. 45, pp. 6331–6345, 2020.

[69] M. Moltafet, N. M. Yamchi, M. R. Javan, and P. Azmi, "Comparison study between PD-NOMA and SCMA," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1830–1834, 2018.

[70] Y. Dong, L. Qiu, and X. Liang, "Energy efficiency maximization for uplink SCMA system using CCPSO," in *2016 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–5, Washington, DC, USA, 2016.

[71] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, "Comparison study of non-orthogonal multiple access schemes for 5G," in *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, pp. 1–5, Ghent, Belguim, 2015.

[72] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall),*, pp. 1–5, Vancouver, BC, Canada, 2014.

[73] X. Dai, S. Chen, S. Sun et al., "Successive interference cancelation amenable multiple access (SAMA) for future wireless communications," in *2014 IEEE International Conference on Communication Systems*, pp. 222–226, Macau, China, 2014.

[74] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and J. J. P. C. Rodrigues, "Tactile internet for smart communities in 5G: an insight for NOMA-based solutions," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 3104–3112, 2019.

[75] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and J. J. P. C. Rodrigues, "DIYA: tactile internet driven delay assessment NOMA-based scheme for D2D communication," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6354–6366, 2019.

[76] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and M. Guizani, "Cross layer NOMA interference mitigation for femtocell users in 5G environment," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4721–4733, 2019.

[77] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and N. Guizani, "Subchannel assignment for SWIPT-NOMA based HetNet with imperfect channel state information," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 842–847, Tangier, Morocco, 2019.

[78] I. Budhiraja, S. Tyagi, S. Tanwar, N. Kumar, and M. Guizani, "CR-NOMA based interference mitigation scheme for 5G femtocells users," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Abu Dhabi, United Arab Emirates, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8833767, 11 pages
https://doi.org/10.1155/2020/8833767

Research Article

Bodacious-Instance Coverage Mechanism for Wireless Sensor Network

Shahzad Ashraf,[1] Omar Alfandi,[2] Arshad Ahmad,[3] Asad Masood Khattak,[2] Bashir Hayat,[4] Kyong Hoon Kim,[5] and Ayaz Ullah[6]

[1]*College of Internet of Things Engineering, Hohai University, Changzhou Jiangsu, China*
[2]*College of Technological Innovation at Zayed University, Abu Dhabi, UAE*
[3]*Department of IT & Computer Science, Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology, Mang Khanpur Road, Haripur 22620, Pakistan*
[4]*Institute of Management Sciences, Peshawar, Pakistan*
[5]*School of Computer Science & Engineering, Kyungpook National University, Daegu 41566, Republic of Korea*
[6]*Department of Computer Science, University of Swabi, Anbar 25000, Pakistan*

Correspondence should be addressed to Bashir Hayat; bashir.hayat@imsciences.edu.pk

Received 25 July 2020; Revised 22 September 2020; Accepted 29 October 2020; Published 28 November 2020

Academic Editor: Farman Ullah

Due to unavoidable environmental factors, wireless sensor networks are facing numerous tribulations regarding network coverage. These arose due to the uncouth deployment of the sensor nodes in the wireless coverage area that ultimately degrades the performance and confines the coverage range. In order to enhance the network coverage range, an instance (node) redeployment-based Bodacious-instance Coverage Mechanism (BiCM) is proposed. The proposed mechanism creates new instance positions in the coverage area. It operates in two stages; in the first stage, it locates the intended instance position through the Dissimilitude Enhancement Scheme (DES) and moves the instance to a new position, while the second stage is called the depuration, when the moving distance between the initial and intended instance positions is sagaciously reduced. Further, the variations of various parameters of BiCM such as loudness, pulse emission rate, maximum frequency, grid points, and sensing radius have been explored, and the optimized parameters are identified. The performance metric has been meticulously analyzed through simulation results and is compared with the state-of-the-art Fruit Fly Optimization Algorithm (FOA) and, one step above, the tuned BiCM algorithm in terms of mean coverage rate, computation time, and standard deviation. The coverage range curve for various numbers of iterations and sensor nodes is also presented for the tuned Bodacious-instance Coverage Mechanism (tuned BiCM), BiCM, and FOA. The performance metrics generated by the simulation have vouched for the effectiveness of tuned BiCM as it achieved more coverage range than BiCM and FOA.

1. Introduction

Wireless sensor networks (WSNs) have been widely considered as one of the most important technologies for the twenty-first century. The sensor nodes are deployed to observe the surrounding events for some phenomenon of interest and thereby process the sensed data and transmit it. These sensor nodes are typically smaller in size with inbuilt microcontrollers and radio transceivers. The fundamental issue in observing such an environment is the area coverage that reflects how well the region is being monitored. Coverage is usually defined as a measure of how well and how long the sensors are able to observe the physical space. The quality of coverage in static sensors is significantly affected by the initial deployment location of the sensor nodes [1]. Unfortunately, sensor deployment cannot be performed manually in most applications, for instance, the deployment in disaster areas, harsh environments, and toxic regions. Thus, sensors are usually deployed by scattering them from an aircraft; however, the actual landing position cannot be uniform due to the existence of obstacles like buildings, trees, and wind causing some areas of the sensing region to be denser than

others. Therefore, even if a large number of redundant nodes are deployed, the desired level of coverage still cannot be achieved [2]. Therefore, it is essential to make use of sagacious sensors that can move iteratively to a better location and can achieve the substantial coverage. In order to address the sensing coverage area, it is important to understand the attributes of the sensor node mobility control mechanism. Indeed, the sensor nodes have two types of mobility control attributes, i.e., centralized and distributed. For the centralized attribute, the bunch of nodes is centrally monitored by a sink node that overhears the sensing data from neighboring nodes, while in distributed networks, the sensors are self-controlled [3].

All sensor nodes have limited sensing and communication abilities which make the sensor nodes unable to obtain the entire network information. Due to that, sensors are being deployed randomly and allowed to move and communicate with respective neighbors by exchanging information among them. Miniaturized robotics has overcome some hurdles regarding sensor mobility. Thereby, mobile sensors have the same sensing capability as static sensors and can move freely to correct locations for providing the required coverage [4], but on the other hand, it is not a cost-effective solution. Considering all aforementioned challenges, we were motivated to design a sagacious sensor node deployment strategy which should enhance the coverage area by consuming the confine energy metrics. Considering the pattern of a hybrid sensor network [5], which has the dual mechanism of mobile and static sensors, we have proposed a Bodacious-instance Coverage Mechanism (BiCM) for wireless sensor networks. For this purpose, a BiCM algorithm has been designed which focuses on how to redeploy the sensor nodes to improve the network coverage area in the hybrid WSN environment. It is indeed a cost-effective solution for improving the coverage of unevenly deployed sensor nodes.

Initially, the proposed algorithm presages where the sensor nodes should be moved to while incurring the trivial moving cost. This will only result in a confined moving cost including the accumulated moving distance, total number of moves, and communication rounds. This algorithm can maintain a balance between coverage and resource consumption during the node redeployment process. The BiCM functions in two stages: In the first stage, the intended target positions of the instance (sensor node) are being computed through the Dissimilitude Enhancement Scheme (DES) [6]. The second stage is called the depuration [7], where the instance moving distance is sagaciously reduced; thereby, the final positions are attainable.

The strenuous contributions in regard to the objective of this study are given below.

(1) The proposed BiCM algorithm tends to overcome related issues with the network coverage range by shifting already deployed sensor nodes from previous to new positions

(2) In some cases, it makes substitutions of nodes to adjust the coverage hole

(3) The unnecessary sensor movement is also being monitored to reduce the movement distance between nodes which prevents the wastage of the energy resource

(4) The simulation results generated through MATLAB have vouched for the succulent performance of BiCM and tuned BiCM when compared with previous work such as FOA

(5) The proposed mechanism accomplished the operation in two junctures: During the first juncture, the intended target positions of the sensor node are computed through the Dissimilitude Enhancement Scheme (DES). The second juncture is referred to as depuration, where the moving distance between nodes is sagaciously reduced; thereby, the target positions are achieved

The rest of the findings are structured as follows: The previous work has been rummaged out in Section 2 and the proposed methodology has been explained in Section 3, while Section 4 renders the output performance and the discussion. Finally, overall achievements have been summarized in the form of a conclusion in Section 5.

2. Literature Review

Usually, the sensor nodes are deployed to cover the area between distinct boundaries; however, selection of the most suitable area has remained an ever present challenge. In order to achieve the sufficient coverage area, the distributed deployment strategy is commonly used to improve the coverage interest by moving the sensor nodes from one location to another. For this purpose, the distributed movement algorithms [8] are being used wherein the coverage area is allocated in multiple segments. If any sensor node was unable to detect the event happenings within the deployed segment, no other sensor node can detect it. Eventually, the monitoring of each segment area for the coverage gap (hole) [9] and calculation of a new instance location are the prime liabilities of the deployed sensor node.

All distributed movement algorithms are facing numerous tribulations regarding new instance calculations within the segment area while relocating the new location. No researcher could ever address overcoming the instance reallocation challenge in a hybrid environment. Therefore, no wireless network having coverage holes can successfully carry out its monitoring operation [10]. The researcher tried to incorporate more iterations in their designed model to address the new allocation issue, but it drastically increases the implications and causes higher energy consumption [11].

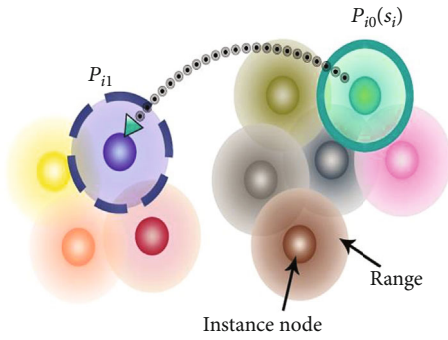
To some extent, numerous researchers have made substantial contributions to avoid such issues, for example, the motion capability of sensor nodes with relocation ability and dealing with sensor failure have been identified by Zhang and Fok [12]; they suggested a two-phase sensor relocation solution. The redundant sensors are first identified and then relocated to the target location. They proposed a grid-quorum solution to locate the closest redundant sensor and then use the cascaded movement to relocate the redundant sensors. In fact, the suggested model could not control the exorbitant energy drainage, and thereby, the entire network might die after the few transmission rounds. On the other

Table 1: Comparative analysis among various algorithms with the proposed BiCM.

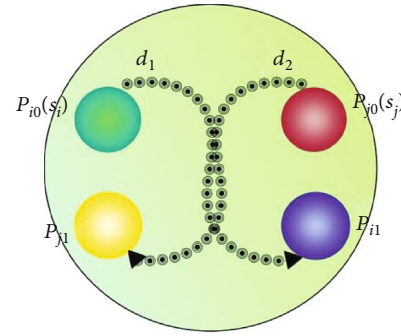
Algorithm	Working ground	Expediency	Impairments	Comparison with proposed BiCM
Genetic algorithm (GA)	Stochastic search methodology through generic system: within a population, it impels the recombination and mutation.	It is faster and has the ability to find the best quality solution in trivial time, possesses parallel capabilities, and easily discovers the global optimum.	It never guarantees an optimal solution. It is hard to choose parameters like number of generations and population size. It is expensive.	It functions in a hybrid environment and ensures relocation of the intended instance position within the coverage area; therefore, energy consumption remains confined.
Particle swarm optimization (PSO)	Inspired by bird flocking and fish schooling: the particles move in a multidimensional search space, and the single intersection of all dimensions forms a particle.	It can overcome the unconstrained minimization issue. Providing the derivative-free technique, it is less sensitive and less dependent on a set of initial points. It can generate high-quality solutions.	It can easily fall into the local optimum in high-dimensional space and has a low convergence rate in the iterative process. It is difficult to adopt the best topology.	At the beginning, it rummages where the sensor nodes should be moved; therefore, local minima can easily be avoided.
Tabu search (TS)	It works on the principle of adaptive memory and responsive exploration.	It has simple implementation and provides robust solution for complex issues.	It vanishes in a local minimum, requires large computing time, and cannot give an upper bound for the computation time	Within a trivial period, it maintains the network coverage range.
Bacterial foraging algorithm (BFA)	It works on search and optimal foraging decision-making capabilities; problems and movement take place either in clockwise or counterclockwise direction.	It is used for unconstrained numerical optimization, having dual movement, i.e., swimming and tumbling called chemotaxis.	It has a weak ability to perceive the environment and is vulnerable to perception of the local extreme; it is hard to deal with complex optimization problems.	As it operates in two stages, thereupon, no vulnerabilities can slow down the performance, and each stage performs independently.
Ant colony optimization (ACO)	Based on social behaviour of the insects, the optimization process is initialized by random solutions.	It allows rapid discovery of good solutions with guaranteed convergence.	It has dependent sequences of random decisions, a complicated theoretical analysis, and uncertain time to convergence.	The depuration technique in second stage reduces the moving distance, and there exists no uncertainty.
Harmony search (HS)	It is based on musical instrument harmony and is a process for better harmony movement.	No setting value is required; it can deal with discrete and continuous variables and can ignore the local optima.	It encounters a high-dimensional multimodal issue, causes unproductive iterations, and has poor local search.	Due to the hybrid environment, the local search is free of being followed by factors; thus, there are no impeaching hurdles.
Artificial bee colony (ABC)	Search optimization consists of three essential components: employed and unemployed foraging bees and food sources.	It minimizes the expense of deploying nodes inside the monitoring region, deals with local solution, and has broad applicability and complex functions.	It has a low process and a higher number of objective function evaluations; number of dimensions might change.	It maintains the network dimension by reducing the moving distance between instance nodes.
Jenga-inspired optimization algorithm (JOA)	Based on greedy fast convergence, it selects the minimum cost node subset through the roulette method and is a bridge between the optimal solution and a short computation time.	It addresses the energy-efficient coverage issues, having stochastic approach to conduct random exploration; if a sensor node cannot cover an area, the other node will avail of the chance.	The detection probability decreases exponentially as the distance becomes greater.	It has shrewd control over the moving distance; therefore, no uncouth movement can degrade the overall communication.

hand, Storn and Price [13] tried to address the coverage and load balancing issues by minimizing the moving distance and argued for a centralized movement solution, based on the Hungarian method. However, the centralized movement technique revealed that those sensor nodes already have appropriate positions when impelled to leave the position creating energy holes.

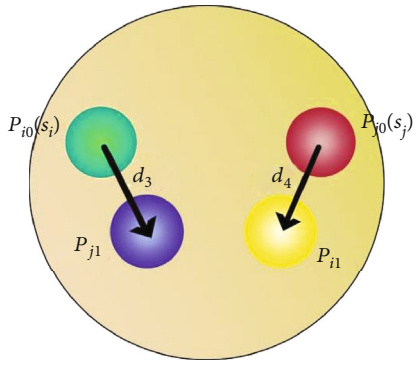
Wang et al. [14] proposed three different distributed movement-assisted sensor deployment algorithms, VEC, VOR, and Minimax, to improve the total area coverage.



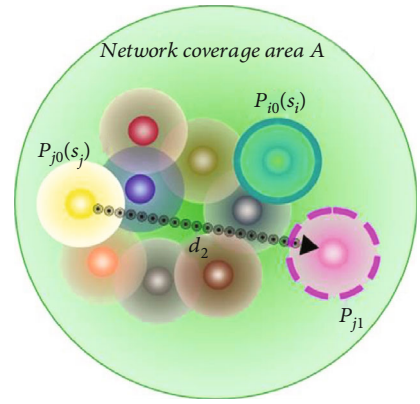
(a) Instance sensor node movement



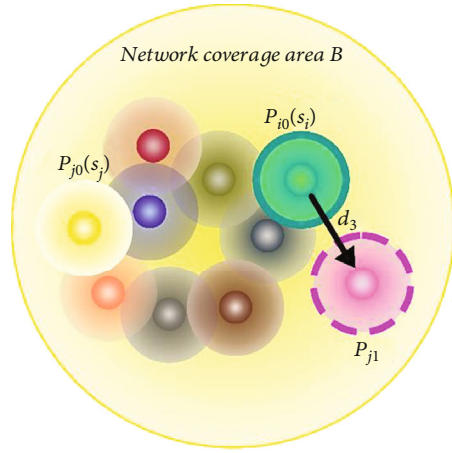
(b) Instance sensor node movement toward intended position



(c) Instance sensor node movement achieving intended position



(d) Moving distance reduction in coverage range A



(e) Moving distance reduction achieved in coverage range B

Figure 1: Instance sensor node movements.

Thereby, they used the Voronoi diagram to partition the monitoring area into n convex polygons where every polygon enclosed one sensor node only. This method utilizes the local polygon information [15], to calculate the new instance location to move the sensor node. The VEC approach uses virtual force between two nodes to push them away from each other at a certain distance. Minimax and VOR algorithms are greedy and try to fix the largest coverage hole by moving the sensor node towards the farthest polygon vertex. The nodes approaching the polygon do not need to move towards the farthest vertex. As a result, this movement may not reduce the coverage hole but might increase the complications.

The identification of a new instance location and its relative computation has been calculated through four local displacement conditions by Mahboubi and Aghdam [16], taking into account the circles having a centered position within the respective polygons. Some centers might lie out of the polygon, and thereby, sensor nodes locating around those circles may not have movement. Consequently, this issue demands more rounds to overcome the coverage tribulation. The more the rounds it demands, the more the resources are being consumed; as a result, the sensor nodes will cause the network to confine the lifespan before the specified time.

In order to increase the coverage rate of sensor nodes, various researchers have proposed different optimization

Table 2: Simulation parameters for BiCM.

Parameter identifiers	Values
Deployment area	$60 \times 60\,m^2$
Number of sensor nodes	60
Grid point	$0.4\,m * 0.4\,m$
Group size	20
Sensing radius	5 m
Maximum iterations	25
Loudness	0.5
Pulse emission rate	0.5
f_{min}	0
f_{max}	2

techniques. A sensing and perception-based Fruit Fly Optimization Algorithm (FOA) [17] was applied by Das et al. to address the position issue of the sensor node which is aimed at enhancing the coverage matter in ideal and obstacle environments. As the fruit flies can reach the food source by using their smell and vision organs, initially, they use osphresis organs to find all kinds of scents in the air. Then, they fly toward the food. When they get close to the food, they use their vision organs to get closer. Similar action is adopted for relocating the sensor positions. Despite its advantages, there are critical issues, for instance, the first pointing location remains poor. Further, the algorithm significantly traps into the local optimum, and the update strategy is limited.

In pursuit of a better coverage technique, a majority of scholars have tried to use intelligent algorithms, like Genetic Algorithm (GA) [18] and Particle Swarm Optimization (PSO) [19], to solve the issue. Though the Fruit Fly Optimization Algorithm is more simple and practicable than GA and PSO, but due to unavoidable limitations, the researchers are still exerting their efforts to develop a shrewder algorithm. Keeping the coverage phenomenon at a high level, Huang et al. [20] introduced a Multiworking Set Alternate Coverage (MWSAC) mechanism that claims to achieve a continuous partial coverage range. The author has achieved a maximum number of working sets by applying a distributed algorithm. The sleep and awakening mechanisms of nodes are adopted which separate the number of active and inactive nodes and keep them synchronous from time to time. Through this method, the nodes appear to work in shifts because the workload has been greatly reduced and the consumption of energy becomes trivial. The authors have however not addressed the false detection occurring in multiworking wireless sensor networks. Table 1 exhibits various comparisons among such algorithms and shows a significant improvement by the proposed algorithm.

3. Coverage Model

A coverage model explains the possible coverage range by the sensor nodes in a coverage area [21]. All sensor nodes have various coverage ranges characterized by area [22], where these sensors are being deployed, the accuracy, the environment factors, and the resolution. The coverage area depends on various factors such as the signal strength generated from the source, distance between the sensor node and the source, and the rate of attenuation in propagation [23]. For example, for an acoustic sensor network establishing the coverage range to detect the mobile vehicles, the sensor nearer to a vehicle can detect higher acoustic signal strength than the one farther away from the vehicle due to signal attenuation, and as a result, there is higher confidence of detecting vehicles [24].

3.1. Problem Formulation. For the proposed coverage model, a two-dimensional coverage area [25] has been considered. Further, the coverage area is divided into various segments each having unit size. When n number of sensor nodes have been deployed in the targeted area m, a full couplet of the sensor node can be defined as given in

$$S = \{S_1, S_2, \cdots..S_n\}. \tag{1}$$

The position of the i^{th} node is defined as $S_i = (x_i, y_i)$ where $i = (1, 2, \cdots n)$. The coverage range of sensor S_i can be expressed as a circle centered at its coordinates (x_i, y_i) with the radius of the sensing range R_s. Let E_i, be a random variable for an event where a sensor node S_i covers an area of segment $A(xA, yA)$. The presage factor for event E_i can be written as $P\{E_i\}$ which is equal to the coverage presage, i.e., $P(S_i, xA, yA)$. Thereupon, the happening of a presage event can be defined by the discrete coverage model expressed in

$$P(S_i, xA, yA) = \begin{cases} 1, d(S_i, xA, yA), & \leq R_s, \\ 0, & \text{other case.} \end{cases} \tag{2}$$

The Euclidean distance [26] of the i^{th} sensor node from segment area $A(x, y)$ can be computed by

$$P(S_i, xA, yA) = \sqrt{(x - x_i)^2 + (y - y_i)^2}. \tag{3}$$

All coverage pints within the coverage range are measured as unity covered by the particular sensor, whereas the points outside of this coverage range are regarded as 0. The shrewd objective of the coverage optimization issue is to provide a sufficient coverage range (CR) [27], by using less number of sensor nodes. The CR is used to estimate the performance of the sensor network. Generally, it is assumed that the segment area point can be covered by any sensor node only once.

3.2. BiCM Model. At present, among all optimization algorithms, the DES is considered as the fastest optimization scheme; therefore, we found it sagacious and were motivated to take full advantage of it for our proposed BiCM algorithm. Thus, the coverage range tribulations in WSN are being resolved by redeployment of sensor nodes through DES strategies, and therefore, the stages of the BiCM design model are explained one by one.

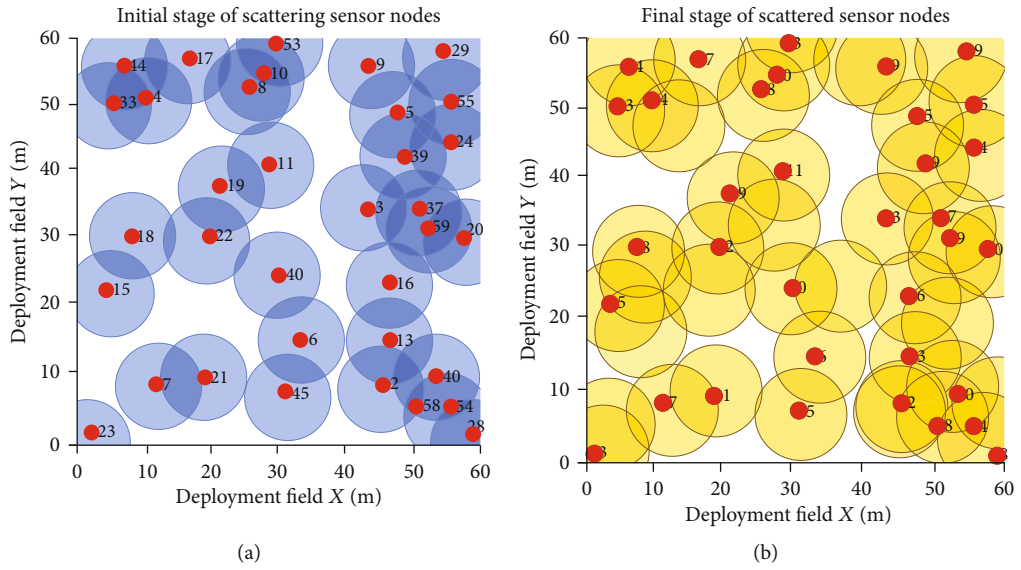


FIGURE 2: (a) The initial and (b) the final FOA sensor node deployment.

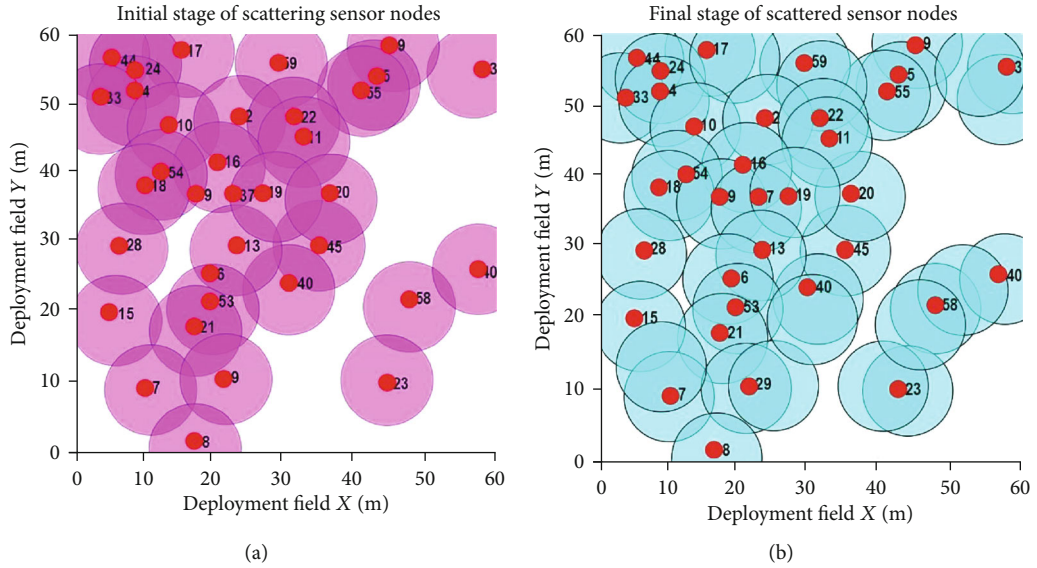


FIGURE 3: (a) The initial and (b) the final deployment of sensor nodes by BiCM.

3.2.1. Stage 1: Locating Intended Target Positions of the Instance. The Bodacious-instance Coverage Mechanism (BiCM) is an investigative search technique that utilizes the shrewd coverage mechanism. It exploits the instance of potential solutions and individuals, to probe the search range. It initializes the parameters while addressing the coverage area issue as depicted in

$$X_i = (x_{i1}, \cdots, x_{ii}, \cdots, x_{iD}), \tag{4}$$

considering $1 \leq i$, as the area range and $x_{ii} \in [\text{ai, bi}]$, where ai and bi denote the lower and upper bounds of the i^{th} node, respectively, and D represents the diameter of the sensor range accompanied with surrounding positions [28]. After every transmission round t, the corresponding reallocation round presages the new expected position of the bodacious instance node which is expressed as

$$V_i(t+1) = X_{\text{bodacious}} + F(X_{r2}(t) - X_{r3}(t)) + F(X_{r4}(t) - X_{r5}(t)). \tag{5}$$

The $X_{\text{bodacious}}$ indicates the appropriate position of the instance while r represents the transmission round and F points to a scaling factor that is a distance control parameter between the initial and the new instance position. To increase the sensing range, the position parameter $V_i(t+1)$

Table 3: Influence of pulse emission rate on coverage rate.

Pulse emission rate (r)	Initial coverage rate (%)	Final coverage rate (%)
0.1	0.8	0.8929
0.2	0.8124	0.905
0.3	0.787	0.9077
0.4	0.8281	0.9041
0.5	0.8097	0.908
0.6	0.8202	0.9025
0.7	0.8208	0.9218
0.8	0.8167	0.9108
0.9	0.8537	0.9354
1	0.8314	0.9153

Table 4: Effect of loudness on coverage rate.

Loudness, A_o (db)	Initial coverage rate (%)	Final coverage rate (%)
0.1	0.8052	0.8931
0.2	0.8375	0.9291
0.3	0.8491	0.9056
0.4	0.8281	0.9107
0.5	0.8276	0.9167
0.6	0.828	0.9219
0.7	0.8273	0.9048
0.8	0.8308	0.9259
0.9	0.8343	0.9281
1	0.8169	0.9179

incorporates the value of predicted instance $X_i(t)$, thereby yielding a temporal position $Q_i\,(t+1)$ as expressed in

$$Q_{i,j}\,(t+1) = \left\{ V_{i,j}(t+1), \text{if (rand } [0,1] < \text{CR or } j = J_{\text{rand}})X_i, j(t), \text{for other case.} \right. \tag{6}$$

The rand (0,1) represents a uniformly distributed random positions, while j_{rand} exhibits randomly predicted positions within the range $[1, D]$. The CR came up as a fractional control parameter $\epsilon[0, 1]$, which shows the inherited characters of previous instance position.

Proceeding towards the final position, the temporal position $Q_i(t+1)$ is being compared with predicted instance $X_i\,(t)$. The newly generated position that possessed a greater fitness metric among the rest of the positions is our intended position of the instance given in

$$X_i(t+1) = \begin{cases} Q_i(t+1), & \text{if } (f(Q_i(t+1)) \geq f(X_i(t))), \\ X_i(t), & \text{other case,} \end{cases} \tag{7}$$

Here, $f(X)$ represents the intended target position of the instance. In fact, the sensor network performs the virtual movement, and as long as it achieves the intended position of the instance sensor in accordance to the Equation (7), physical displacement has been performed accordingly.

Table 5: Effect of $f_{\max}$ on coverage rate.

$f_{\max}$ (f)	Initial coverage rate (%)	Final coverage rate (%)
0.1	0.8492	0.8698
0.2	0.819	0.8433
0.3	0.8135	0.8359
0.4	0.8115	0.8327
0.5	0.831	0.8602
0.6	0.8186	0.8507
0.7	0.8196	0.8414
0.8	0.8211	0.8417
0.9	0.8499	0.8712
1	0.8369	0.8549
1.1	0.8298	0.8888
1.2	0.822	0.9053
1.3	0.8134	0.9331
1.4	0.7965	0.898
1.5	0.8116	0.91
1.6	0.8367	0.9279
1.7	0.8145	0.9169
1.8	0.8267	0.9132
1.9	0.8296	0.9147
2	0.8127	0.9078

3.2.2. Stage 2: Depuration Process. The depuration process is performed to reduce the moving distance of the instance. This will reduce the number of instances (sensor nodes) that need to move, as well as reduce the average moving distance; however, it does not affect the network coverage. The moving distance reduction strategy can be understood as the following: consider the initial positions of an i^{th} instance node s_i is $P_{i0}(x_{i0}, y_{i0})$ and the j^{th} instance node s_j have $P_{j0}(x_{j0}, y_{j0})$. The length of the distance is defined as $d_1 = \overline{|p_{i0}p_{i1}|}$ and $d_2 = \overline{|p_{j0}p_{j1}|}$ and so on. The BiCM algorithm searches the new intended positions of all instance nodes in the coverage area and systematically reduces the number of instance nodes that are needed to be moved. The instance-sensing range may even fully overlap with other instance nodes [29]; these nodes are called redundant nodes and are illustrated in Figure 1(a). The instance sensor node s_i displaces from p_{i0} to p_{i1}; thereby, the coverage rate $R_{\text{area}}(S)$ shows that no substantial change has been recorded which confirms that no movement is required by the s_i instance node. Therefore, the substantial instance nodes can be removed from the queue which eventually decreases the distance.

The position of the instance nodes is being updated by changing the distance position of s_i and s_j that is $d_1 + d_2$ before and after the displacement has been occurred, and it will be updated to $d_3 + d_4$ accordingly as given in

Table 6: Influence of grid points on coverage rate.

Grid points (m ∗ m)	Initial coverage rate (%)	Final coverage rate (%)
0.1 ∗ 0.1	0.8306	0.9203
0.2 ∗ 0.2	0.7975	0.9006
0.3 ∗ 0.3	0.8006	0.9106
0.4 ∗ 0.4	0.8342	0.9132
0.5 ∗ 0.5	0.8012	0.9056
0.6 ∗ 0.6	0.8451	0.9341
0.7 ∗ 0.7	0.8052	0.9125
0.8 ∗ 0.8	0.8135	0.9181
0.9 ∗ 0.9	0.8142	0.9200
1 ∗ 1	0.8240	0.9212

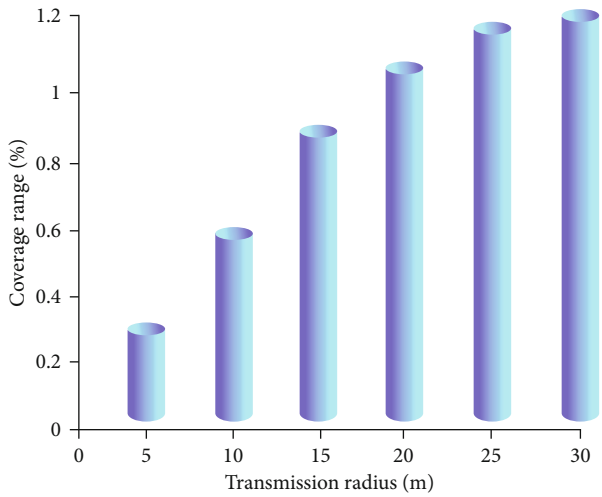


Figure 4: Coverage rate for varying sensing radii of sensor nodes by BiCM.

Figure 1(b). It is worth mentioning that $d_1 + d_2 > d_3 + d_4$; therefore, achieving the intended positions, the moving distance of s_i and s_j can be confined but no change will occur in the coverage area, but the coverage area distance rate will be extended. The instance nodes that are eager to update their moving position will be substituted with the moving position of the nodes which are stationary and do not require further movement. This step can prevent the instance nodes from making unnecessary and longer movement. In this case, the instance node does not possess sufficient energy while reaching the intended position; thereby, other surrounding nodes will surrogate the liability. We should consider Figure 1(c), where instance node s_i does not plan to leave its position while at the same time instance node s_j is eager to shift its position from P_{j0} to P_{j1}. Therefore, the instance node s_i is displaced from P_{i0} to P_{j1} but s_j remains in hiatus. The coverage range $B \geq A$ and $3 < d_2$, instead of sensor node s_j, and the algorithm smartly shifts the instance node s_i to the intended new position of node s_j while keeping the s_j node stationary. This change will not affect the coverage range of the network and does not impel the rest of the instance nodes to move in the queue. Eventually, an average moving distance of the instance node is reduced which enhances the coverage area distance rate. This moving distance reduction is illustrated in Figures 1(d) and 1(e).

4. Simulation Results and Discussion

In order to validate the efficiency of node deployment based on BiCM, the simulation trials are conducted using MATALAB R2016a [30]. The performance among BiCM, tuned BiCM, and FOA is carried out using the simulation setup parameters given in Table 2. To observe the performance of the aforementioned algorithms, nearabout 60 sensor nodes were deployed randomly in the monitoring area of size 60 $\times$ 60 m^2. To demonstrate the performances of FOA, BiCM, and tuned BiCM, the initial and final node deployments are presented in Figures 2 and 3.

These Figures 2 and 3 signify the initial and final node deployments after executing the FOA and BiCM algorithms. Thereupon, it can be clearly understood that node deployment based on BiCM has minimum redundancy and is most uniform compared to node deployment by the FOA mechanism. Table 3 signifies the influence of pulse emission rate (r) on the coverage of sensor nodes. The value of r changes from 0.1 to 1 whereas the value of other instance mechanism parameters such as loudness, maximum frequency, and sensing radius is kept constant to 0.5, 2, and 5, respectively. To beat the effect of arbitrariness [31], the instance mechanism is simulated 50 times, and greatest value of coverage is picked every time. The maximum value of coverage after performing BiCM is attained as 93.54% at a pulse emission rate of 0.9. As instances move towards their respective target (grid points), they emit a greater number of pulses [32]; therefore, the pulse emission rate will be high when sensor nodes move close to the grid points [33]. Thereupon, the value of the pulse emission rate is kept at 0.9. Further, to see the effect of the loudness parameter of the instance mechanism on the coverage rate of sensor nodes, the value of loudness (A_o) is varied from 0.1 to 1 while the pulse emission rate (r) is set to 0.9 and the value of other parameters is 0.5; the sensing radius (r_s) is fixed at 5 meters. Table 4 shows the variations of loudness and initial and final coverage rates of nodes after implementing BiCM. The BiCM is run 50 times, and the best value of the initial and final coverage rates is selected. The coverage rate after executing BiCM is obtained as the highest at about 93.1% at the 0.2 value of loudness. When sensor nodes (instance) get near to the grid point, the intensity of emitted pulses is low; therefore, the loudness parameter should be kept low [34]. Thereupon, the value of the loudness parameter is fixed at 0.2.

In addition to this, Table 5 demonstrates the effect of maximum frequency ($f_{\max}$) [35], on coverage; its value has been changed from 0.1 to 2. The constraints of the instance mechanism for instance pulse emission rate, loudness, and sensing radius are kept constant to 0.9, 0.2, and 5, respectively. For each variation of maximum frequency, the instance mechanism has been executed 50 times and supreme values of coverage before and after the execution

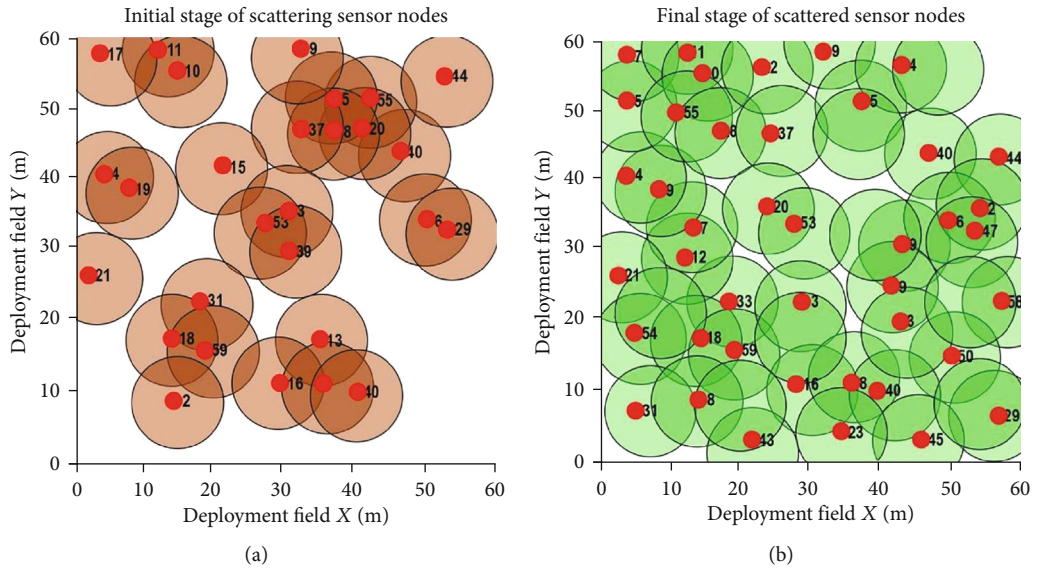


Figure 5: (a) Initial deployment of sensor nodes for tuned BiCM; (b) final deployment of sensor nodes by tuned BiCM.

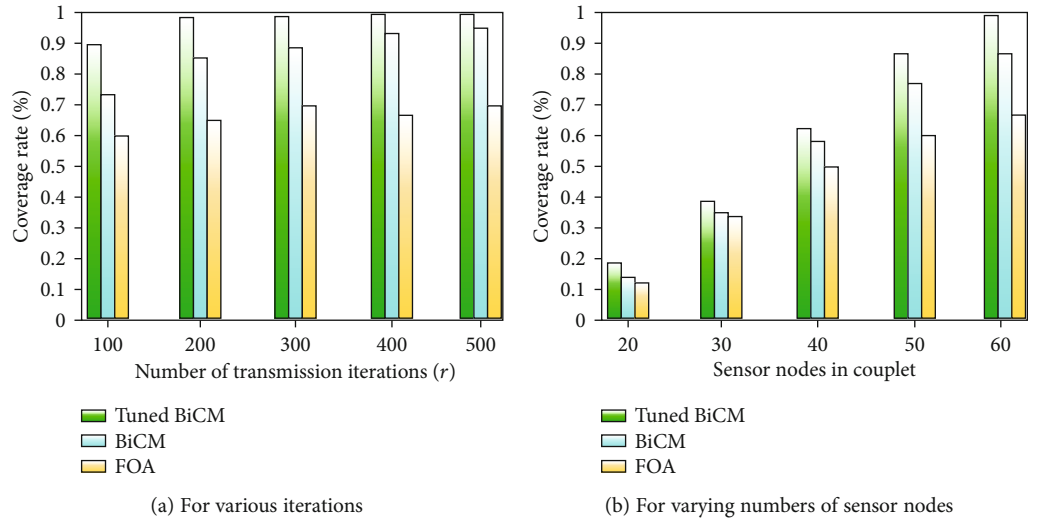


Figure 6: Coverage rate analysis by FOA, BiCM, and tuned BiCM.

Table 7: Deployment results for FOA, BiCM and Tuned BiCM.

Algorithms	FOA		BiCM		Tuned BiCM	
Parameters	Initial results	Final results after execution	Initial results	Final results after execution	Initial results	Final results after execution
Average coverage rate	75.56%	85.16%	82.72%	91.91%	91.54%	98.29%
Standard deviation	0.0286	0.0251	0.0187	0.0126	0.0126	0.0055
Best coverage value	78.92%	87.49%	87.10%	94.30%	93.45%	99.46%
Worst coverage value	68.40%	78.20%	79.38%	90.02%	89.55%	97.31%

of the instance mechanism have been chosen. The best value of coverage after implementing BiCM is 93.31% when $f_{\max}$ is 1.3. Thus, the value of $f_{\max}$ is set to 1.3. To observe the impact of grid points on the coverage rate of nodes, the value of the grid point has varied from 0.1 m ∗ 0.1 m to 1 m ∗ 1 m. The various simulation factors such as pulse emission rate,

Table 8: Comparison of computation time of BiCM, FOA, and tuned BiCM.

Algorithms	FOA	BiCM	Tuned BiCM
Computation time (s)	0.28	0.019	0.016

maximum frequency, sensing radius, and loudness are kept constant at 0.9, 1.3, 5, and 0.2, respectively. In Table 6, every value of grid point BiCM runs 50 times and the uppermost values of the coverage rate have been taken. The highest value of the coverage rate at about 93% is obtained after running the BiCM when grid points were set to 0.6 m ∗ 0.6 m. Further, the sensing radius is varied from 1 m to 10 m. Figure 4 signifies the variations of the coverage rate after applying BiCM w.r.t. changes in the sensing radius of the node. The parameters of BiCM, for example, grid points, loudness, pulse emission rate, and maximum frequency, are set as 0.6 m ∗ 0.6 m, 0.2, 0.9, and 1.3, respectively. It is clear from Figure 4, as the sensing radius has increased, that the coverage rate of sensor nodes is also increased, and its value is 100% when the sensing radius is increased beyond 7 m. But there is a trade-off between the sensing radius and cost: while the sensing radius of the node is increased, the cost of sensor nodes also increased.

The tuned values of various constraints of BiCM such as loudness, maximum frequency, sensing radius, pulse emission rate, and grid points are 0.2, 1.3, 6, 0.9, and 0.6 m ∗ 0.6 m, respectively. To validate the performance of node deployment based on BiCM after setting the above constraint values, the initial and final node deployments after executing the tuned BiCM are shown in Figure 5. Thereupon, it can be obviously seen that node deployment based on tuned BiCM has the lowest redundancy compared with BiCM and FOA. To further demonstrate the effectiveness of tuned BiCM, the coverage rates for the tuned BiCM, BiCM, and FOA for various iterations are shown in Figure 6. The iterations are varied from 0 to 500. The convergence speed of the tuned BiCM is more compared to FOA. The tuned BiCM converged around 150 iterations, whereas FOA converges around 350 iterations due to exploitation characteristics of the instances.

The tuned BiCM has achieved a higher coverage rate at about 99.46% compared to 93.37% and 88.33% of BiCM and FOA, respectively. In order to overwhelm the effect of randomness of tuned BiCM, instance mechanism optimization and Fruit Fly Optimization Algorithms are run 15 times. The deployment results in terms of average coverage rate, standard deviation, and best and worst coverage values for tuned BiCM and FOA are represented in Table 7. It can be obviously seen from Table 7 that tuned BiCM has achieved the average coverage rate of about 98.29% compared to 91.91% and 85.16% of BiCM and the Fruit Fly Optimization Algorithm. Further, the standard deviation for node deployment based on tuned BiCM is lowest, so tuned BiCM is more stable compared to FOA and BiCM. The best and worst coverage values for tuned BiCM are 99.46% and 97.31% compared to 94.30% and 90.02% and 87.49% and 78.20% for the BiCM- and FOA-based node deployments, respectively.

Further, the comparison of tuned BiCM, BiCM, and FOA in terms of computation time is represented in Table 8. The computation time for tuned BiCM is less, i.e., 0.016 seconds, compared to 0.019 seconds and 0.28 seconds for BiCM and FOA, respectively. The tuned BiCM and BiCM converge at 25 iterations whereas FOA converged at 500 iterations; therefore, the speeds of tuned BiCM and BiCM are more and converge faster at an earlier stage because of their exploitation feature compared to the Fruit Fly Optimization Algorithm.

5. Conclusion

In order to enhance the coverage rate of the sensor nodes, an innovative sensor deployment technique based on Bodacious-instance Coverage Mechanism (BiCM) has been purposed that accomplished the desired goal with limited energy consumption. The analysis of various factors of BiCM such as loudness, grid points, emission rate and radius of nodes, and frequency has been identified, and shrewd values of the above parameters are discovered. Node deployment based on tuned BiCM and BiCM shows that both algorithms converge at an earlier stage compared to the Fruit Fly Optimization Algorithm. The simulation results demonstrate that tuned BiCM has attained a mean coverage rate of about 98.29% which is higher compared to FOA and BiCM. Further, various simulations have been done by varying the number of sensor nodes and iterations, and a coverage rate curve is plotted for tuned BiCM, BiCM, and FOA. The comparison of the computation time is also represented in this paper. Tuned BiCM has a high coverage rate and less computation time compared to FOA and BiCM. In the future, the various evolutionary optimization algorithms can be applied to the node deployment problem to increase the coverage rate of sensor nodes.

Data Availability

The data to support the findings of this study is available inside the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Zayed University Research Fund # R19046.

References

[1] M. Abazeed, N. Faisal, S. Zubair, and A. Ali, "Routing protocols for wireless multimedia sensor network: a survey," *Journal of Sensors*, vol. 2013, 11 pages, 2013.

[2] S. Ashraf, M. Gao, Z. Chen, S. Kamran, and Z. Raza, "Efficient node monitoring mechanism in WSN using ContikiMAC protocol," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 11, 2017.

[3] F. Ait Aoudia, M. Gautier, M. Magno, O. Berder, and L. Benini, "A generic framework for modeling MAC protocols in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 3, pp. 1489–1500, 2017.

[4] S. Ashraf, M. Gao, Z. Mingchen, T. Ahmed, A. Raza, and H. Naeem, "USPF: underwater shrewd packet flooding mechanism through surrogate holding time," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 9625974, 12 pages, 2020.

[5] M. Li, X. Du, X. Liu, and C. Li, "Shortest path routing protocol based on the vertical angle for underwater acoustic networks," *Journal of Sensors*, vol. 2019, Article ID 9145675, 14 pages, 2019.

[6] S. Ashraf, T. Ahmed, A. Raza, and H. Naeem, "Design of shrewd underwater routing synergy using porous energy shells," *Smart Cities*, vol. 3, no. 1, pp. 74–92, 2020.

[7] M. S. Aliyu, A. H. Abdullah, H. Chizari, T. Sabbah, and A. Altameem, "Coverage enhancement algorithms for distributed mobile sensors deployment in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 12, no. 3, Article ID 9169236, 2016.

[8] S. Ashraf, Z. Aslam, A. Yahya, and A. Tahir, "Underwater routing protocols analysis of intrepid link selection mechanism, challenges and strategies," *International Journal of Scientific Research in Computer Science and Engineering*, vol. 8, no. 2, pp. 1–9, 2020.

[9] S. Ashraf and T. Ahmed, "Machine learning shrewd approach for an imbalanced dataset conversion samples," *Journal of Engineering and Technology (JET)*, vol. 11, no. 1, 2020.

[10] S. Balsamo, A. Marin, and E. Vicario, Eds., *New Frontiers in Quantitative Methods in Informatics: 7th Workshop, InfQ 2017, Venice, Italy, December 4, 2017, Revised Selected Papers*, Springer, New York, NY, 1st edition edition, 2018.

[11] J. Wang, Y. Gao, C. Zhou, R. Simon Sherratt, and L. Wang, "Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs," *Computers, Materials & Continua*, vol. 62, no. 2, pp. 695–711, 2020.

[12] Q. Zhang and M. Fok, "A two-phase coverage-enhancing algorithm for hybrid wireless sensor networks," *Sensors*, vol. 17, no. 12, p. 117, 2017.

[13] R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[14] G. Wang, G. Cao, and T. F. la Porta, "Movement-assisted sensor deployment," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 640–652, 2006.

[15] J. Wang, Y. Yang, T. Wang, R. S. Sherratt, and J. Zhang, "Big Data Service Architecture: A Survey," *Journal of Internet Technology*, vol. 21, no. 2, 2020.

[16] H. Mahboubi and A. G. Aghdam, "Distributed deployment algorithms for coverage improvement in a network of wireless mobile sensors: relocation by virtual force," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 736–748, 2017.

[17] S. Das, A. Biswas, S. Dasgupta, and A. Abraham, "Bacterial foraging optimization algorithm: theoretical foundations, analysis, and applications," in *Foundations of Computational Intelligence Volume 3: Global Optimization*, A. Abraham, A.-E. Hassanien, P. Siarry, and A. Engelbrecht, Eds., pp. 23–55, Springer, Berlin, Heidelberg, 2009.

[18] H. Stringer, *Behavior of variable-length genetic algorithms under random selection*, University of Central Florida, 2007.

[19] L. Sun, X. Song, and T. Chen, "An improved convergence particle swarm optimization algorithm with random sampling of control parameters," *Journal of Control Science and Engineering*, vol. 2019, Article ID 7478498, 11 pages, 2019.

[20] M. Huang, A. Liu, M. Zhao, and T. Wang, "Multi working sets alternate covering scheme for continuous partial coverage in WSNs," *Peer-to-Peer Networking and Applications*, vol. 12, no. 3, pp. 553–567, 2019.

[21] S. Ashraf, Z. A. Arfeen, M. A. Khan, and T. Ahmed, "SLM-OJ: surrogate learning mechanism during outbreak juncture," *International Journal for Modern Trends in Science and Technology*, vol. 6, no. 5, pp. 162–167, 2020.

[22] J. Wang, X. Gu, W. Liu, A. K. Sangaiah, and H.-J. Kim, "An empower Hamilton loop based data collection algorithm with mobile agent for WSNs," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 2019.

[23] S. Goyal and M. S. Patterh, "Flower pollination algorithm based localization of wireless sensor network," in *2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS)*, Chandigarh, India, Dec. 2015.

[24] J. Wang, Y. Gao, W. Liu, and W. Wu and Se-Jung Lim, "An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks," *Computers, Materials & Continua*, vol. 58, no. 3, pp. 711–725, 2019.

[25] S. Ashraf, D. Muhammad, and Z. Aslam, "Analyzing challenging aspects of IPv6 over IPv4," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 6, no. 1, pp. 54–67, 2020.

[26] "How to calculate Euclidean distance," May 2020, https://sciencing.com/how-to-calculate-euclidean-distance-12751761.html.

[27] Yourim Yoon and Yong-Hyuk Kim, "An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1473–1483, 2013.

[28] S. Ashraf, D. Muhammad, M. Shuaeeb, and Z. Aslam, "Development of shrewd cosmetology model through fuzzy logic," *Journal of Research in Engineering and Applied Sciences*, vol. 5, no. 3, pp. 93–99, 2020.

[29] S. Ashraf, A. Raza, Z. Aslam, H. Naeem, and T. Ahmed, "Underwater resurrection routing synergy using astucious energy pods," *Journal of Robotics and Control (JRC)*, vol. 1, no. 5, 2020.

[30] J. Zhang, Y. Lei, C. Chen, and F. Lin, "Directional probability perceived nodes deployment based on particle swarm optimization," *International Journal of Distributed Sensor Networks*, vol. 12, no. 4, Article ID 2046392, 2016.

[31] A. Shahzad and A. Tauqeer, "Dual-nature biometric recognition epitome," *Trends in Computer Science and Information Technology*, vol. 5, no. 1, pp. 8–14, 2020.

[32] S. Ashraf, T. Ahmed, S. Saleem, and Z. Aslam, "Diverging mysterious in green supply chain management," *Oriental journal of computer science and technology*, vol. 13, no. 1, pp. 22–28, 2020.

[33] J. E. Franklin and R. J. Urick, "A binary detection model for at-sea sonar prediction," *The Journal of the Acoustical Society of America*, vol. 66, no. S1, p. S15, 1979.

[34] S. Ashraf, S. Saleem, A. H. Chohan, Z. Aslam, and A. Raza, "Challenging strategic trends in green supply chain management," *Journal of Research in Engineering and Applied Sciences*, vol. 5, no. 2, pp. 71–74, 2020.

[35] S. Ashraf, A. Ahmad, A. Yahya, T. Ahmed, and 3 Dow University of Health Sciences Karachi Pakistan, "Underwater routing protocols: analysis of link selection challenges," *AIMS Electronics and Electrical Engineering*, vol. 4, no. 3, pp. 234–248, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8857272, 20 pages
https://doi.org/10.1155/2020/8857272

Research Article

A Lightweight Nature Heterogeneous Generalized Signcryption (HGSC) Scheme for Named Data Networking-Enabled Internet of Things

Manazara Rehman,[1] Hizbullah Khattak,[1] Ahmed Saeed Alzahrani,[2] Insaf Ullah,[3] Muhammad Adnan,[4] Syed Sajid Ullah,[1] Noor Ul Amin,[1] Saddam Hussain,[1] and Shah Jahan Khattak[5]

[1]*IT Department, Hazara University, Mansehra, 21120 KP, Pakistan*
[2]*Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia*
[3]*Department of Computing, HIET, Hamdard University, Islamabad Campus, Islamabad 44000, Pakistan*
[4]*Division of Computer and Information Sciences, Higher Colleges of Technology, 17155, Al Ain, UAE*
[5]*Pakistan Engineering Council (PEC), Attatruk Avenue (East) Sector G-5/2, Islamabad 44000, Pakistan*

Correspondence should be addressed to Insaf Ullah; insafktk@gmail.com

Received 27 April 2020; Revised 3 July 2020; Accepted 30 July 2020; Published 18 November 2020

Academic Editor: Fawad Zaman

Internet of Things (IoT) is the collection of different types of smart objects like mobile phones, sensors, cars, smart cities, smart buildings, and healthcare, which can provide a quality life to humans around the globe. These smart objects sense and produce a huge amount of data for distribution. The current hostcentric networking paradigm is not that scalable to provide a suitable solution to the idea of IoT. For scalable connectivity and efficient distribution, Named Data Networking (NDN) has been envisioned as a promising solution for future internet architecture. On the other hand, the significant issues regarding the adaptation of NDN with IoT possess security concerns such as authentication, confidentiality, integrity, and forward secrecy. As IoT is a heterogeneous environment, it demands a different type of security, according to the environmental situation such as public key infrastructure (PKI), identity-based cryptosystem (IBC), and certificateless cryptosystem (CLC). This paper presents a new concept of CLC to IBC heterogeneous generalized signcryption for the first time to fulfil the prime security requirements of NDN-based IoT. The proposed scheme provides the security properties according to situational needs without disturbing the structural policy of NDN. Considering the resource-constrained nature of IoT, we used a lightweight type of elliptic curve called the hyperelliptic curve cryptosystem which offers the same level of security as that of bilinear pairing and an elliptic curve cryptosystem using a minimum key size. Further, we compare the proposed scheme with recently proposed identity-based as well as certificateless generalized signcryption schemes, and the results give satisfactory outputs in terms of computational and communication resources. Furthermore, we simulate the proposed scheme with Automated Validation of Internet Security Protocols and Applications (AVISPA), and the results show that our scheme is valid and safe. Additionally, we provide a practical scenario of the proposed on NDN with an IoT-based smart city.

1. Introduction

Nowadays, Internet of Things (IoT) is deployed in almost every environmental domain, such as smart grid, smart homes, smart cities, smart building, healthcare, and smart agriculture, by connecting and controlling a large number of objects [1]. The increasing numbers of smart applications and their heterogeneity raise some challenges regarding their

connectivity, communication, scalability, mobility, and amount of generating data [2]. To address these challenges, Named Data Networking (NDN) has been projected as a future internet architecture [3]. In general, NDN deals with two packets: the interest packet and the data packet. The communication of the NDN is based on the altercation on interest packets that carry the request. Further, the NDN node maintains three types of data structures that are the Content Store (CS), which stores the copy of the contents with itself in the CS for future use; Pending Interest Table (PIT), which enlists all the requests of the incoming interfaces in the PIT table; and Forwarding Information Base (FIB), which forwards the requests from one node to another based on routing protocols [4]. Security is instigated on each packet, so the authenticity can be achieved at a time inside the network [5]. Whenever a consumer sends an interest packet for some specific contents, the NDN router performs CS lookup; if the requested contents are available, then the router simply forwards the contents directly from its CS to the requested consumer [6]. If the requested contents are not available in the CS, then the NDN router checks its PIT table for that requested content; if the contents have been requested before, then the PIT table updates with an entry of that specific interface in the PIT table. If the contents are being requested for the first time, then the PIT table marks up an entry of that interface and forwards the request to the next router based on FIB as shown in Figure 1. The monumental features of NDN like in-network caching, scalability, name-based routing, and mobility are a suitable option for fulfilling the demands of IoT applications.

However, security is considered to be the fundamental need for NDN-based IoT devices. Additionally, the NDN-based IoT environment requires different types of security properties such as authentication, confidentiality, and integrity, which can be achieved from a digital signature, encryption, or signcryption according to the environmental situation. Moreover, IoT is a heterogeneous environment where the sender and receivers may come from different types of environments. Here, the concept of heterogeneous signcryption is a suitable option that makes use of two different types of cryptosystems in a single algorithm [7]. On the other hand, the IoT devices may demand a digital signature, encryption, or signcryption, separately or in combination. For this type of situation, the heterogeneous signcryption becomes effortless due to its nongeneralized nature such as providing signcryption only. Here, the concept of generalized signcryption may be able to provide a digital signature, encryption, or signcryption using a single algorithm [8]. Likewise, the generalized signcryption cannot fulfill the requirement of IoT devices due to its homogeneity.

Generally, the security and efficiency of the aforementioned schemes are based on computationally hard problems like RSA, bilinear pairing, and elliptic curve cryptosystem. The RSA provides a solution using a 1024-bit large key which is firmly based on large factorization [9–11]. However, due to the limited processing capabilities of IoT devices, the 1024-bit key is not an efficient solution. On the other hand, bilinear pairing suffers from the issue of high pairing operations and is 12.93 times worse than RSA [12]. Hence, to tackle the weaknesses of both RSA and bilinear pairing, a new type of cryptosystem was introduced [13] called the elliptic curve cryptosystem. Unlike RSA and bilinear pairing, the security difficulty of the elliptic curve cryptosystem is based on a small key size of 160 bits. However, the 160-bit key is still not appropriate for resource-limited IoT devices [14]. Hence, in [15], a new type of cryptosystem was introduced, called the hyperelliptic curve cryptosystem, which suits the resource-limited nature of IoT devices by using a small key of 80 bits [16, 17].

The above discussion motivates us to contribute a new concept of heterogeneous generalized signcryption for NDN-based IoT which will combine the idea of heterogeneous signcryption with generalized signcryption to fulfill the conditional demands of IoT. The features of this new concept are mentioned as follows:

(1) First, we introduced a new concept of CLC to IBC heterogeneous generalized signcryption

(2) We provide the proper syntax of our proposed scheme

(3) We also provide a proper algorithm for the proposed scheme on the basis of the hyperelliptic curve cryptosystem which is suited for the IoT environment

(4) We prove the security properties such as authentication, confidentiality, unforgeability, forward secrecy, and integrity of the proposed scheme

(5) We also compared our proposed scheme with recently published CLC and IBC generalized signcryption schemes, and the results give satisfactory outputs in terms of computational and communication resources

(6) We also validate the security of our scheme through AVISPA, and the results show that our proposed scheme is valid and safe

(7) We practically deployed our scheme on the NDN-based smart city

1.1. Paper Organization. The organization of the paper is shown in Figure 2.

2. Related Work

Here, we divided the related work into three parts such as identity-based generalized signcryption, certificateless generalized signcryption, and heterogeneous generalized signcryption.

2.1. Identity-Based Generalized Signcryption Schemes. Lal and Kushwah in 2008 [18] introduced the concept of an identity-based generalized signcryption (ID-BGS) scheme for the first time to solve the certificate management issues of PKI-based generalized signcryption. In 2010, the concept was used by Liang et al. [19] for key management issues in mobile ad hoc networks (MANET). The proposed scheme saves memory storage of users and minimizes computational and communication resources. Kushwah and Lal in 2011 [20]

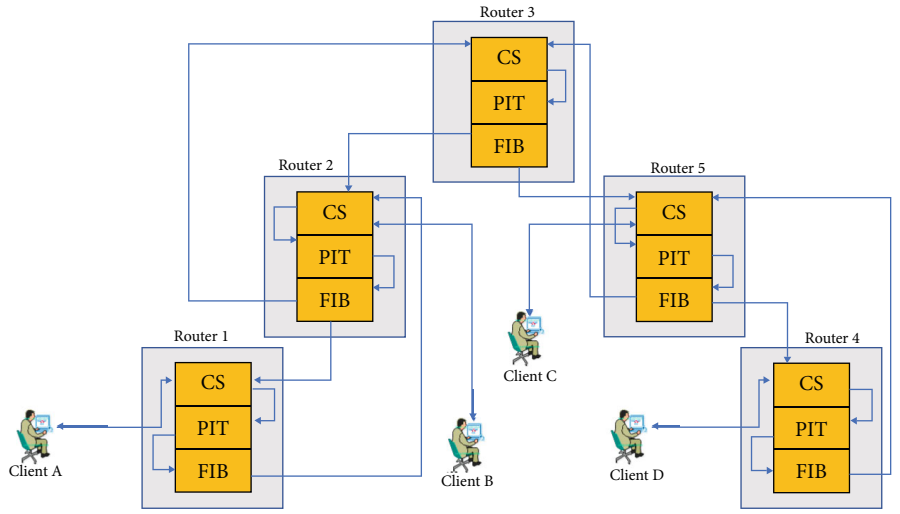


Figure 1: The basic architecture of NDN.

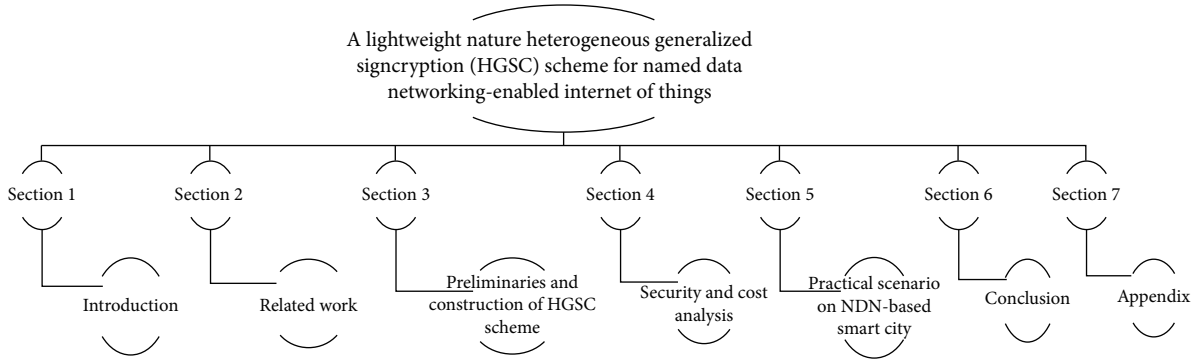


Figure 2: Paper organization.

proposed an efficient ID-BGS scheme for wireless sensor networks (WSN). The authors used bilinear pairing and proved the security of the proposed scheme under the random oracle model (ROM). Wei et al. [21] proposed an efficient ID-BGS for obtaining the confidentiality and authenticity of big data. Mishra and Singh in 2014 [22] surveyed the existing identity and certificateless generalized signcryption schemes. Based on security limitations in the existing schemes, the authors proposed two schemes to improve the limitations. Shen et al. in 2017 [23] improve the security of existing IBGS schemes which is suitable for low storage devices. Waheed et al. in 2019 [24] proved that the security of the Wei et al. [21] scheme is susceptible to attack and insecure. In the proposed cryptanalysis, the authors launched a security attack on the Wei et al. [21] scheme and found that the master secret key of the proposed scheme can be easily compromised.

However, the schemes [18–23] suffer from a heavy pairing operation due to the use of bilinear pairing. In [24], the authors did not provide any sort of solution to the proposed claims.

2.2. Certificateless Generalized Signcryption Schemes. Huifang et al. in 2010 [25] defined the notion of certificateless generalized signcryption (CGS) to solve the key escrow problem of IBGS. Later, Kushwah and Lal in 2012 [26] improved the security flaws of Huifang et al. [25] and proposed a new CGS scheme which is unforgeable against insider attacks. In 2014, Zhou et al. [27] proposed a provable CGS scheme for resource-constrained environment devices. The scheme provides security against malicious, but passive, key generation centre attacks. Zhang et al. in 2016 [28] proposed a CGS scheme for mobile health (M-Health). The scheme reduces the computation and communication costs by the use of the elliptic curve cryptosystem. Zhou et al. in 2017 [29] proposed a GSC scheme for security insurance in cloud storage. Zhang et al. in 2018 [30] proposed an efficient CGS scheme that is suitable for low power and low processor devices due to the use of the elliptic curve cryptosystem. Further, the scheme provides security against ciphertext attacks. In 2019, Zhou [31] improved the scheme of Zhang et al. [30] and proposed a new scheme for the mobile health system that can monitor the human body status in real time. Waheed et al. in 2019 [32] analyzed the proposed scheme of Zhou et al. [29] and proved that the scheme of Zhou et al. [29] is insecure against ciphertext indistinguishability under adaptive chosen-ciphertext attacks (IND-CCA2). Further, the

author proposed a new and improved scheme at the same cost which is secure against the aforementioned attacks. Karati et al. in 2019 [33] proposed a new CGS for resource-constrained IoT devices.

However, the schemes [25–33] suffer from heavy computation and communication costs due to the use of bilinear pairing and an elliptic curve cryptosystem.

2.3. Heterogeneous Signcryption Schemes. In 2011, Huang et al. [7] introduced the concept of heterogeneous signcryption (HS) which uses two different types of cryptosystem such as IBC at the sender side and CBC at the receiver side. The proposed scheme was suitable for the practical scenario of IoT where the sender and receiver belonged to different environments. Li et al. in 2016 [34] proposed a multireceiver heterogeneous signcryption scheme for wireless area network applications. The authors used CLC on the sender side and IBC on the receiver side (CLC-IBC). Raveendranath and Aneesh in 2016 [35] proposed a multireceiver HS scheme by using the elliptic curve cryptosystem to reduce the computation and communication costs of the existing HS scheme. Niu et al. in 2017 [36] proposed a CLC to IBC HS scheme by using bilinear pairings in the random oracle model. In the same year, Niu et al. [37] proposed a hybrid IBC to CLC scheme for multimessage and multireceiver. Li et al. in 2017 [38] proposed a PKI to IBC HS scheme for vehicle ad hoc networks. Niu et al. in 2017 [39] proposed a CLC to IBC HS scheme for the privacy-preserving multiparty aggregate scheme. Saeed et al. in 2017 [40] proposed a CLC to PKI online/offline HS scheme for IoT. Furthermore, the authors practically deployed the scheme on healthcare and the smart grid. Wang et al. in 2017 [41] proposed a PKI to IBC HS scheme for broadcast communication in ad hoc networks. Jin et al. in 2018 [42] proposed an IBC to PKI HS scheme for secure communication in the smart grid. Liu et al. in 2018 [43] proposed two HS schemes, such as PKI to CLC and CLC to PKI for secure communications between 5G network slicing. Liu and Ma in 2018 [44] proposed a cross domain of the PKI and IBC HS scheme for the medical information system. The authors use an elliptic curve cryptosystem to reduce computation and communication resources. Omala et al. in 2018 [45] proposed a CLC to IBC heterogeneous access control scheme for body area networks. Zhou et al. in 2019 [46] proposed a PKI to IBC HS scheme for vehicular ad hoc networks.

The aforementioned schemes [7], [34–46] suffer from heavy computation and communication costs due to the use of bilinear pairing and an elliptic curve cryptosystem. Moreover, these schemes are not suitable for NDN-based IoT due to its nongeneralized nature.

3. Preliminaries and Construction of HGSC Scheme

In this section, we will discuss the background of the hyperelliptic curve, threat model, and construction of our proposed HGSC scheme.

3.1. Hyperelliptic Curve (HEC). First, we will define the basic mathematics of hyperelliptic curves (HEC). Let $\mathscr{A}$ be a finite set and G be a genus of HEC with an order $G \geq 2$. Suppose (u), $f(u)\, \varepsilon\, d\,[u]$ and deg $(\mathrm{H}(u)) \leq G$, and $f(u)$ is a monic polynomial possessing deg $(f(u)) = 2d + 1$ [47]. Furthermore, HEC of genus $G \geq 2$ over d is a set of points (u), $d * d$ as in the mentioned equation:

$$\mathrm{HEC} : y^2 + (u)y = f(u). \tag{1}$$

Note: the point of HEC is not the same as elliptic curves [48]. It forms a divisor (D) that is the formal sum of finite integers such as $D = \sum \varkappa_i z_{\varkappa_i}$ where $\varkappa_i\, \varepsilon\, d$ and $z_{\varkappa_i}\, \varepsilon$ HEC. Additionally, HEC over the Jacobian group J_{HEC} has the brief order mentioned in the following equation:

$$(u_{\tau-1})^{2\mathrm{G}} \leq J_{\mathrm{HEC}}(\mathscr{A}) \leq (u_{\tau+1})^{2G}. \tag{2}$$

3.1.1. Hyperelliptic Curve Discrete Logarithm Problem. Assume D is a divisor, which is publicly known to everyone, and ℓ is a private number that is randomly chosen from $\mathscr{A}$ where finding ℓ from $d\,\ell = D$ is known to be an HEC discrete logarithm problem.

3.2. Syntax of Proposed Heterogeneous Generalized Signcryption Scheme. Here, we first explain different notations in Table 1, which can be used in the syntax and our proposed HGSC algorithms.

The syntax proposed scheme consists of 10 algorithms such as setup, generate secrete value, generate public key, generate partial private key, generate full private key, consumer private key generation, signcryption, unsigncryption, signature, and signature verification.

(1) *Setup*: the Key Generation Centre (KGC) executes this algorithm by taking the security parameter $\mathscr{k}$ to generate the master secret key $\mathscr{W}$, master public key $\mathscr{X}$, and public parameter set φ, then publishes φ and $\mathscr{X}$ openly in the network.

(2) *Generate secret value (GSVL)*: the producer takes (k, φ) and generates a secret value ∂.

(3) *Generate public key (GPK)*: in this algorithm, the producer then takes $(\mathscr{k}, \varphi, \partial)$ and generates a public key $\mathscr{B}_p$.

(4) *Generate partial private key (GPPK)*: in this algorithm, the KGC takes $(\mathscr{k}, \varphi, I\mathscr{D}_p, \mathscr{W}, \mathscr{B}_p)$ and generates a partial private key $(\mathscr{N}, \mathscr{J})$.

(5) *Generate full private key (GFPTK)*: in this algorithm, the producer takes $(\mathscr{k}, \varphi, I\mathscr{D}_p, \mathscr{N}, \mathscr{J}, \partial)$ as an input and generates his own full private key $(\mathscr{A}_p)$.

(6) *Consumer private key generation (CPKG)*: in this algorithm, the KGC takes $(ID_c, \mathscr{W})$ as an input by using IBS and a produces private key $(\mathscr{A}_c)$ and a public key $(\mathscr{B}_c)$ for the consumer.

Table 1: Abbreviations used in these algorithms.

Abbreviation	Definition
KGC	Key generation centre
$\mathscr{W}$	Master secret key
$\mathscr{k}$	Security parameter
$\mathscr{X}$	Master public key
φ	Public parameter set
$\mathscr{D}$	Divisor of the hyperelliptic curve
∂	Secret value
$\mathscr{B}_p$	Producer public key
$\mathbf{ID}_p$	Producer identity
$\mathscr{N}, \mathscr{J}$	Partial private key
$\mathscr{A}_p$	Producer full private key
$\mathbf{ID}_c$	Consumer identity
$\mathscr{L}$	Random number
$\mathscr{B}_c$	Consumer public key
$\mathscr{A}_c$	Consumer private key
$\mathscr{m}$	Message
$\mathscr{S}$	Producer digital signature
$\mathscr{H}_1$	Hash
$\mathscr{C}$	Content
Ψ	Singcrypted message/content
$\mathscr{C}$	Fresh nonce
Φ	Sign message/content

(7) *Signcryption*: in this algorithm, the producer takes $(ID_c, \mathscr{B}_c, \mathscr{A}_p, \mathscr{m}, \mathscr{X})$ and generates a signcrypted content/messageΨ and sends it to consumers.

(8) *Unsigncryption*: in this algorithm, the consumer unsigncrypts the Ψ by using $(ID_c, \mathscr{B}_c, \mathscr{A}_c, \mathscr{C}, \mu, \mathscr{S})$.

(9) *Signature*: in this algorithm, the producer takes $(\mathscr{A}_p, \mathscr{m})$ and generates a content/message sign Φ and sends it to the consumer.

(10) *Signature verification*: in this algorithm, the consumer verifies Φ by using $(ID_c, \mathscr{B}_c, \mathscr{A}_c, \mathscr{C}, \mu, \mathscr{S})$.

3.3. Threat Model. In our proposed HGSC scheme, we consider the Dolev-Yao (DY) [49] threat model. According to DY, the communication between two or more entities is not reliable and secure, as the attacker has full commands to reveal the contents of the ciphertext and inject a false signcryption/signature text to the network. The NDN-based IoT environment possesses different types of estimated security threats; it means that the adversaries can easily modify or delete the user's sensitive information. To maintain the security and authentication of NDN-based IoT devices, it is necessary to perform authentic and secure communication among entities in the NDN-based environment. The basic security requirements used in HGSC scheme are as follows: (1) Confidentiality: it means to keep the information secret from unauthorized users. The attackers can break the confidentiality of the HGSC scheme if he/she gets access to the encryption of decryption keys. The attacker here cannot access the original content in the message without having the encryption or decryption keys which are called confidentiality. (2) Unforgeability: it means that the signature could not be reproduced by any other party. Here, the attacker can generate a forged signature if he/she gets access to the digital signature generation secret key. If the attacker fails to do so, then it is called unforgeability. (3) Forward secrecy: forward secrecy means that if one of the session keys gets compromised by any malicious user, the data from the other session could not be affected. Here, the attacker cannot get access to the encryption or decryption keys even if the attacker got access to the sender private key. If the attacker is not able to access the encryption/decryption key of the user, it is called forward secrecy. (4) Antireplay attack: an antireplay attack means the attacker can resend a copy of an authenticated message again. Here, the attacker cannot reply to the existing message again if the sender and receiver use nonce and time stamping techniques for the freshness of a message.

3.4. Proposed Network Model. Here, we explain the workflow of our proposed HGSC scheme for NDN-enabled IoT. In our proposed scheme, we consider four entities such as the producer, consumer, NDN node, and Key Generation Centre (KGC) as shown in Figure 3. Here, we consider that the consumer belongs from IBC while the producer belongs from CLC. For registration of consumers and producers with KGC, the KGC announces the public parameter set and master public key.

3.4.1. Role of KGC. In the producer registration phase, the producer generates its public key from the public parameter set and sends it to the KGC. The KGC then generates a partial private key for the producer and sends it to the producer in reverse order using a secure network. After receiving the partial private key, the producer generates its full private key.

In the consumer registration phase, the consumer sends its identities to the KGC. The KGC after receiving the identities of the consumer generates private as well as public keys for the consumer and sends them back to the consumer using a secured network.

3.4.2. Role of Consumer. Suppose a consumer sends an interest for some content/message in the NDN-based IoT environment to any producer.

3.4.3. Role of Producer. After receiving the interest, the producer then signs/signcrypts the content using its private key and sends it back to the requested consumer. However, the NDN node will store the copy content/message in their CS according to the caching policies of NDN. After receiving the content/message, the consumer verifies the signature or unsigncrypts the respective content/message.

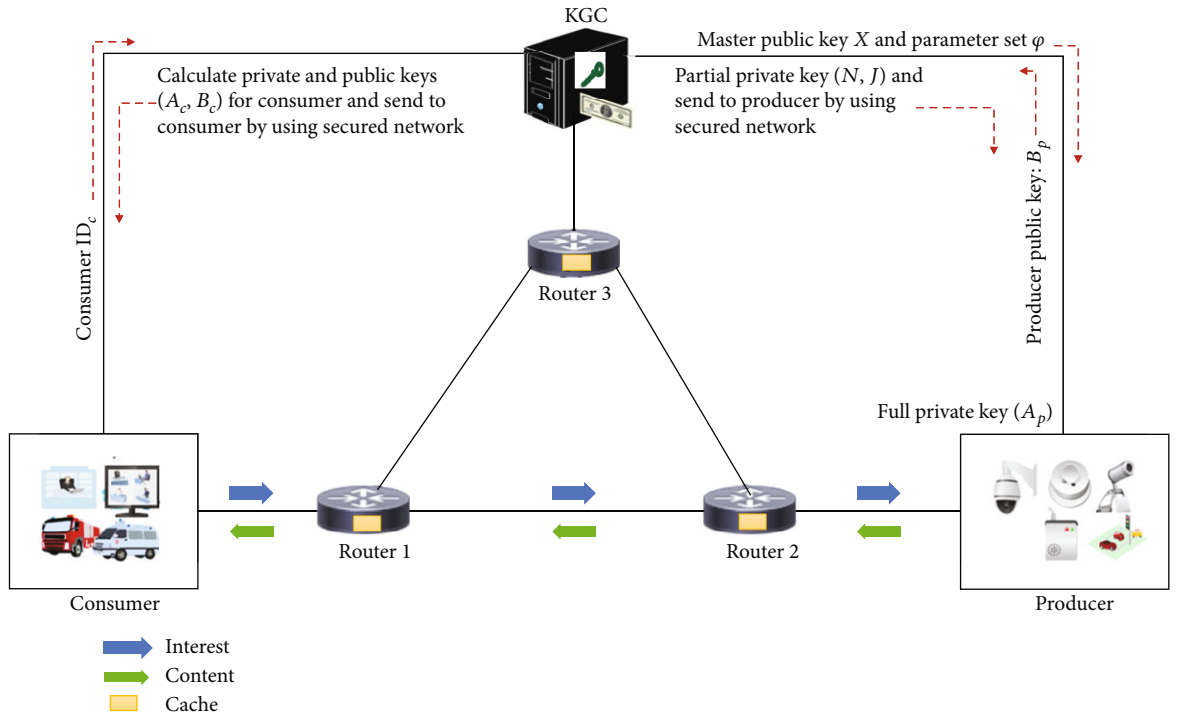


Figure 3: Proposed network model.

Setup:It is processed by KGC
Input: Security parameter $\mathscr{k}$
Output:Master secret key $\mathscr{W}$, master public key $\mathscr{X}$, and public parameter set $\varphi = \{\mathscr{X}, \mathscr{D}, \mathrm{G} \geq 2, \mathrm{HEC}, \mathscr{H}_1, \mathscr{H}_2, \mathscr{H}_3, \mathscr{H}_4\}$.
Process: KGC first produces public parameter set φ, then after randomly pick a master secret key $\mathscr{W} \in \{1, 2, 3, \cdots, z-1\}$ where $\varkappa = 2^{80}$, computes a master public key $\mathscr{X} = \mathscr{W}.\mathscr{D}$, where $\mathscr{D}$ is a divisor of the hyperelliptic curve.
Publishing KGC published φ and $\mathscr{X}$ openly in the network.

Algorithm 1: Setup.

3.5. Proposed Heterogeneous Generalized Signcryption Algorithms. The proposed HGSC consists of the following eight steps.

3.5.1. Algorithm 1. In this step, the KGC generates a master public key, master secret key, and public parameter as shown in Algorithm 1.

3.5.2. Algorithm 2. In this step, the producer generates secrete value as shown in Algorithm 2.

3.5.3. Algorithm 3. In this step, the KGC generates a partial private key for the producer as shown in Algorithm 3.

3.5.4. Algorithm 4. In this step, the producer generates its full private key as shown in Algorithm 4.

3.5.5. Algorithm 5. In this step, the KGC public as well as the private key for the consumer are shown in Algorithm 5.

3.5.6. Algorithm 6. In this step, the producer sign/signcrypts the requested contents as shown in Algorithm 6.

3.5.7. Algorithm 7. In this step, the consumer verifies the sign contents or unsigncrypts the signcrypted contents as shown in Algorithm 7.

4. Security and Cost Analyses

In this section, we briefly discuss the informal analysis and computation and communication cost analyses of our proposed scheme.

4.1. Informal Analysis. This section describes the contribution in upholding the security properties of confidentiality, unforgeability, forward secrecy, and antireplay attack.

4.1.1. Confidentiality. Confidentiality means to keep the contents secret; the attacker (ζ) cannot calculate the plaintext from signcrypted ciphertext. Let the ζ want to break the confidentiality of our proposed scheme and generate the plaintext from signcrypted ciphertext $\Psi = (\mathscr{C}, \mu, \mathscr{S})$. For this purpose, the ζ needs to calculate $\mathscr{C}$ from $\Psi = (\mathscr{C}, \mu, \mathscr{S})$, and to do so, ζ needs δ, β, and α from $K = \mathscr{H}_3(\delta, \alpha, \beta, ID_c$,

Generate secret value (GSVL):It is run by producer
Input: (k, φ)
Output:Secret value ∂
Process: Producer randomly picks a secret value $\partial \in \{1, 2, 3, \cdots, z-1\}$

Algorithm 2: Secret value generation.

Generate partial private key (GPPK):It is executed by KGC
Input: $(\mathbb{k}, \varphi, I\mathbb{D}_p, \mathbb{W}, \mathbb{B}_p)$
Output:Partial private key $(\mathbb{N}, \mathbb{J})$
Process:KGC randomly picks a number $\mathbb{Q} \in \{1, 2, 3, \cdots, z-1\}$, compute $\mathbb{N} = \mathbb{Q} \cdot \mathbb{D}$, compute $\mathbb{J} = \mathbb{Q} + \mathbb{W} \cdot \mathbb{H}_1(ID_p, \mathbb{Q}, \mathbb{B}_p) \mod z$, and send $(\mathbb{N}, \mathbb{J})$ to the producer by using secured network.

Algorithm 3: Partial private key generation.

Generate full private key (GFPTK):It is executed by the producer
Input: $(\mathbb{k}, \varphi, I\mathbb{D}_p, \mathbb{N}, \mathbb{J}, \partial)$
Output:Full private key $(\mathbb{A}_p)$
Process: Producer computes $\mathbb{A}_p = (\mathbb{J}, \partial)$.

Algorithm 4: Partial private key generation.

Consumer private key generation (CPKG):It is executed by KGC and note that KGC acts like a private key generator in an identity-based cryptosystem
Input: $(ID_c, \mathbb{W})$
Output:Private key $(\mathbb{A}_c)$, public key $(\mathbb{B}_c)$
Process: KGC randomly picks a number $\mathbb{L} \in \{1, 2, 3, \cdots, z-1\}$, computes $\mathbb{B}_c = \mathbb{L} \cdot \mathbb{D}$, computes $\mathbb{A}_c = \mathbb{L} + \mathbb{W} \cdot \mathbb{H}_1(I\mathbb{D}_c, \mathbb{L}, \mathbb{B}_c) \mod \mathbb{x}$, and sends $(\mathbb{A}_c, \mathbb{B}_c)$ to the consumer by using a secured network.

Algorithm 5: Consumer's key generation.

Heterogeneous generalized signcryption (HGSN):It is executed by producer
Input: $(ID_c, \mathbb{B}_c, \mathbb{A}_p, m, \mathbb{X})$
Output:Signcryption Ψ
Process: Producer randomly picks a number $\mathbb{R} \in \{1, 2, 3, \cdots, z-1\}$,
If $(\mathbb{C} = m)$
{
(1) Select a fresh nonce $\mathbb{T}$
(2) Compute $\mu = \mathbb{H}_2(\mathbb{T}, m)$
(3) Compute $\mathbb{S} = \mathbb{R} + \mu(\mathbb{J} + \partial)$ and go to step 11
}
Else
{
(4) Compute $\delta = \mathbb{R} \cdot \mathbb{D}$
(5) Compute $\Upsilon = \mathbb{H}_3(ID_c, \mathbb{B}_c, \mathbb{X})$
(6) Compute $\alpha = \mathbb{B}_c + \mathbb{X} \cdot Y$
(7) compute $\beta = \alpha \cdot \mathbb{D}$
(8) compute $K = \mathbb{H}_4(\delta, \alpha, \beta, ID_c, \mathbb{B}_c)$
(9) Repeat step 1 and compute $\mathbb{C} = \mathbb{E}_K(\mathbb{T}, m)$
(10) Repeat steps 2 and 3 and go to step 11
}
(11) Send $\Psi = (\mathbb{C}, \mu, \mathbb{S}, \delta)$ to the consumer by using an open network

Algorithm 6: Signcryption and signature generation.

Heterogeneous generalized unsigncryption (HGUSN)It is executed by consumer
Input: $(ID_c, \mathscr{B}_c, \mathscr{A}_c, \mathscr{C}, \mu, \mathscr{S}, \delta)$
Output and verifications:$(\mathscr{T}, m)$ and $\mu' \stackrel{?}{=} \mu$
Process: (1) Consumer computes $\beta = \delta \cdot \mathscr{A}_c$
(2) Calculates $K = \mathscr{H}_4(\delta, \alpha, \beta, ID_c, \mathscr{B}_c)$
(3) Perform decryption $(\mathscr{T}, m) = D_K(\mathscr{C})$
(4) Compute $\mu' = \mathscr{H}_2(\mathscr{T}, m)'$
(5) Compare $\mu' \stackrel{?}{=} \mu$; if it holds, then accept; otherwise, reject

Algorithm 7. Unsigncryption and signature verification.

Table 2: Comparative analysis in terms of major operations with CGS schemes.

Schemes	Signcryption	Unsigncryption	Total
Zhang et al. [28]	4 SEPM	5 SEPM	9 SEPM
Zhou et al. [29]	7 SEPM	8 SEPM	15 SEPM
Zhang et al. [30]	5 SEPM	4 SEPM	9 SEPM
Zhou [31]	5 SEPM	7 SEPM	12 SEPM
Waheed et al. [32]	1 SBP + 5 SPBPM	3 SBP + 1 SPBPM	4 SBP + 6 SPBPM
Karati et al. [33]	3 SPBPM + 3 SEXP	2 SPBPM + 2 SBP + 5 SEXP	5 SPBPM + 2 SBP + 8 SEXP
Proposed scheme	2 SHEDM	2 SHEDM	4 SHEDM

$\mathscr{B}_c$). Here, $\delta = \mathscr{R} \cdot \mathscr{D}$, $\beta = \alpha \cdot \mathscr{D}$, and $\alpha = \mathscr{B}_c + \mathscr{X} \cdot Y$ where $\mathscr{R}$ and α are discrete logarithm problems over the hyperelliptic curve cryptosystem which is not possible to calculate. Thus, our proposed scheme provides the property of confidentiality.

4.1.2. Unforgeability. Unforgeability means that no one can sign the content, except the valid provider. To forge the signature, ζ needs to calculate $\mathscr{R}$, μ, and $\mathscr{J} + \partial$. Here, $\mathscr{R}$ is a private number, and for calculating $\mu = \mathscr{H}_2(\mathscr{T}, m)$, ζ needs to calculate a private number $\mathscr{T}$ from $\mu = \mathscr{H}_2(\mathscr{T}, m)$. Further, ζ needs to $\mathscr{J} + \partial$ where $\mathscr{J}$ is a fresh nonce and ∂ is a private number, so to forge the signature $\mathscr{S}$, ζ needs to calculate 3 private numbers $\mathscr{R}$, $\mathscr{T}$, and ∂ with a fresh nonce $\mathscr{J}$ which is not possible to calculate. So, our proposed scheme provides the property of unforgeability.

4.1.3. Forward Secrecy. Forward secrecy means if the private key of the signer is compromised, still it could not affect the respective contents, because the content is encrypted via a session secret key. Here, in our scheme, to break forward secrecy, ζ needs to calculate $K = \mathscr{H}_3(\delta, \alpha, \beta, ID_c, \mathscr{B}_c)$ which requires δ where $\delta = R \cdot D$. So, for this purpose, ζ needs to calculate R, which is a private number, and δ is a discrete logarithm problem over the hyperelliptic curve, which is infeasible for ζ to break.

4.1.4. Antireplay Attack. In our proposed scheme, before communication, the provider generates a $\mathscr{T}$ and stores it in his memory. Then after, it sends the encrypted text as $\mathscr{C} = \mathscr{E}_K(\mathscr{T}, m)$ to the consumer. After receiving the FNs, the consumer, by using secret key K, performs the decryption process on the received ciphertext. Once the $\mathscr{T}$ is recovered, the consumer verifies the freshness, and if it is fresh, then the ciphertext is new. However, ζ cannot replay the old messages because he/she needs fresh FNs for every new session.

4.2. Cost Analysis. In this section, we compare the proposed scheme with existing certificateless generalized signcryption (CGS) and identity-based generalized signcryption (ID-BGS) schemes in terms of computation and communication costs.

4.2.1. Computation Cost. Here, we compare our proposed scheme with existing CGS and ID-BGS schemes in terms of expansive mathematical operations such as single pairing-based point multiplication (SPBPM), single bilinear pairing (SBP), single exponential (SEXP), single elliptic curve point multiplication (SEPM), and hyperelliptic curve point multiplication (SHEDM). Moreover, operations like addition, division, subtraction, encryption, decryption, and hash are neglected, due to its minimal consumption time during the computation.

Furthermore, we compare our scheme with the existing CGS and ID-BGS schemes in milliseconds (ms) by using the above major operation, according to the experiments performed in [50] with the following hardware and software specifications:

(i) Intel Core i7-4510U CPU

(ii) 2 GHz processor

(iii) 8 GB RAM

(iv) Windows 7, 64 bits

(v) Multiprecision integer and rational arithmetic C library

According to [50], an SPBPM will take 4.32 ms, a single SBP will take 14.90 ms, SEXP will take 1.25 ms, and SEPM will take 0.97 ms. Based on the experiments performed in [51, 52], we consider that a SHEDM will take 0.48 ms. On the bases of the above expansive mathematical operations, we conduct the computation cost comparison of our proposed scheme with existing CGS schemes which are Zhang et al. [28], Zhou et al. [29], Zhang et al. [30], Zhou [31], Waheed et al. [32], and Karati et al. [33] as shown in Tables 2 and 3. Further, the computation cost comparison of our proposed scheme with existing ID-BGS schemes which are Wei et al. [21] and Shen et al. [23] is shown in Tables 4 and 5. Moreover, a clear computation reduction is shown in Figures 4 and 5.

(1) Computation Cost Reduction of Our Scheme from CGS Schemes. The following formula will be used to calculate cost reduction

$$\left(\frac{\text{existing scheme} - \text{our scheme}}{\text{existing scheme}}\right) * 100. \tag{3}$$

(i) Computation cost reduction from Zhang et al. [28]:

$$\left(\frac{9\,\text{SEPM} - 4\,\text{SHEDM}}{9\,\text{SEPM}}\right) * 100 = \left(\frac{8.73 - 1.92}{8.73}\right) * 100 = 78.09\% \tag{4}$$

(ii) Computation cost reduction from Zhou et al. [29]:

$$\left(\frac{15\,\text{SEPM} - 4\,\text{SHEDM}}{15\,\text{SEPM}}\right) * 100 = \left(\frac{14.55 - 1.92}{14.55}\right) * 100 = 86.80\% \tag{5}$$

(iii) Computation cost reduction from Zhang et al. [30]:

$$\left(\frac{9\,\text{SEPM} - 4\,\text{SHEDM}}{9\,\text{SEPM}}\right) * 100 = \left(\frac{8.73 - 1.92}{8.73}\right) * 100 = 78.09\% \tag{6}$$

(iv) Computation cost reduction from Zhou [31]:

$$\left(\frac{12\,\text{SEPM} - 4\,\text{SHEDM}}{12\,\text{SEPM}}\right) * 100 = \left(\frac{11.64 - 1.92}{11.64}\right) * 100 = 83.50\% \tag{7}$$

Table 3: Computation cost comparison (CGS) in ms.

Schemes	Signcryption	Unsigncryption	Total
Zhang et al. [28]	3.88	4.85	8.73
Zhou et al. [29]	6.79	7.76	14.55
Zhang et al. [30]	4.85	3.88	8.73
Zhou [31]	4.85	6.79	11.64
Waheed et al. [32]	36.5	49.02	85.52
Karati et al. [33]	16.71	44.69	61.4
Proposed scheme	0.96	0.96	1.92

Table 4: Comparative analysis in terms of major operations with ID-BGS.

Schemes	Signcryption	Unsigncryption	Total
Wei et al. [21]	6 SEXP	2 SEXP + 5 SBP	8 SEXP + 5 SBP
Shen et al. [23]	5 SEXP + 1 SBP	3 SBP	5 SEXP + 4 SBP
Proposed	2 SHEDM	2 SHEDM	4 SHEDM

Table 5: Computation cost comparison (ID-BGS) in ms.

Schemes	Signcryption	Unsigncryption	Total
Wei et al. [21]	7.5	77	84.5
Shen et al. [23]	14.90	29.8	44.7
Proposed	0.96	0.96	1.92

(v) Computation reduction from Waheed et al. [32]:

$$\left(\frac{4\,\text{SBP} + 6\,\text{SPBPM} - 4\,\text{SHEDM}}{4\,\text{SBP} + 6\,\text{SPBPM}}\right) * 100 = \left(\frac{85.52 - 1.92}{85.52}\right) * 100 = 97.75\% \tag{8}$$

(vi) Computation cost reduction from Karati et al. [33]:

$$\left(\frac{5\,\text{SPBPM} + 2\,\text{SBP} + 8\,\text{SEXP} - 4\,\text{SHEDM}}{5\,\text{SPBPM} + 2\,\text{SBP} + 8\,\text{SEXP}}\right) * 100 = \left(\frac{61.4 - 1.92}{61.4}\right) * 100 = 96.87\% \tag{9}$$

(2) Computation Cost Reduction of Our Scheme from ID-BGS Schemes.

(i) Our Computation Cost Reduction from Wei et al. [21]:

$$\left(\frac{8\,\text{SEXP} + 5\,\text{SBP} - 4\,\text{SHEDM}}{8\,\text{SEXP} + 5\,\text{SBP}}\right) * 100 = \left(\frac{84.5 - 1.92}{84.5}\right) * 100 = 97.84\% \tag{10}$$

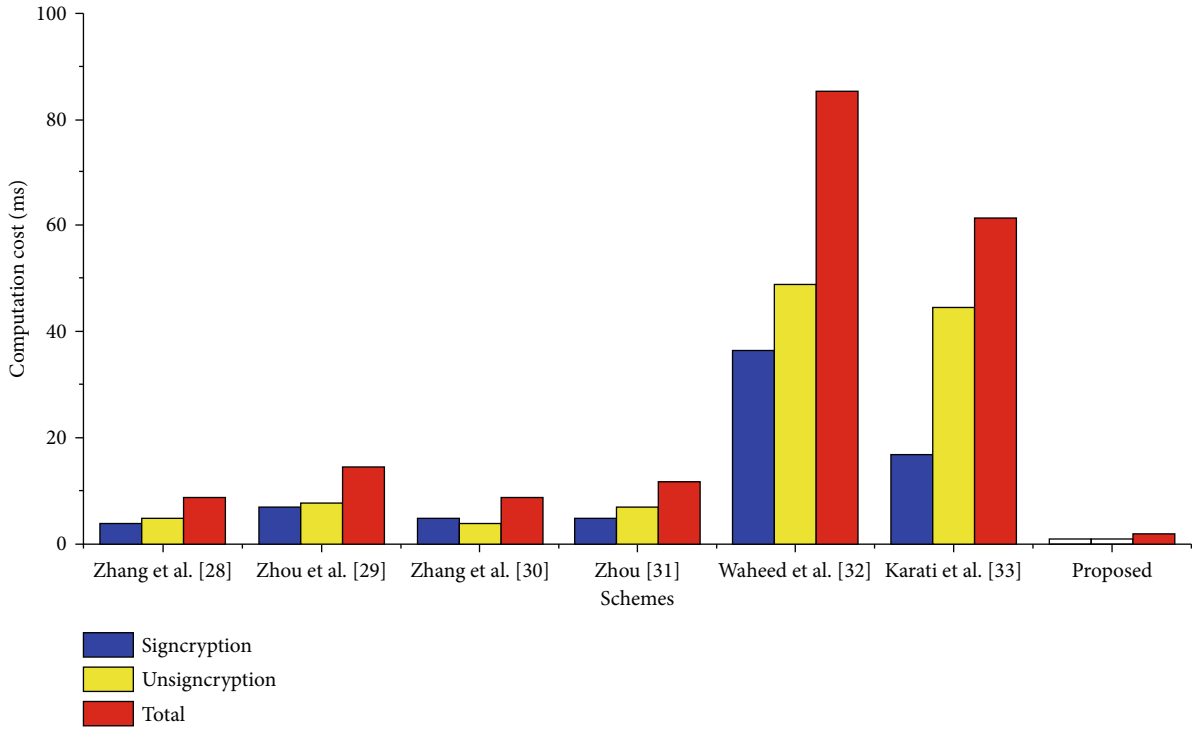


Figure 4: Computational cost reduction from CGS schemes.

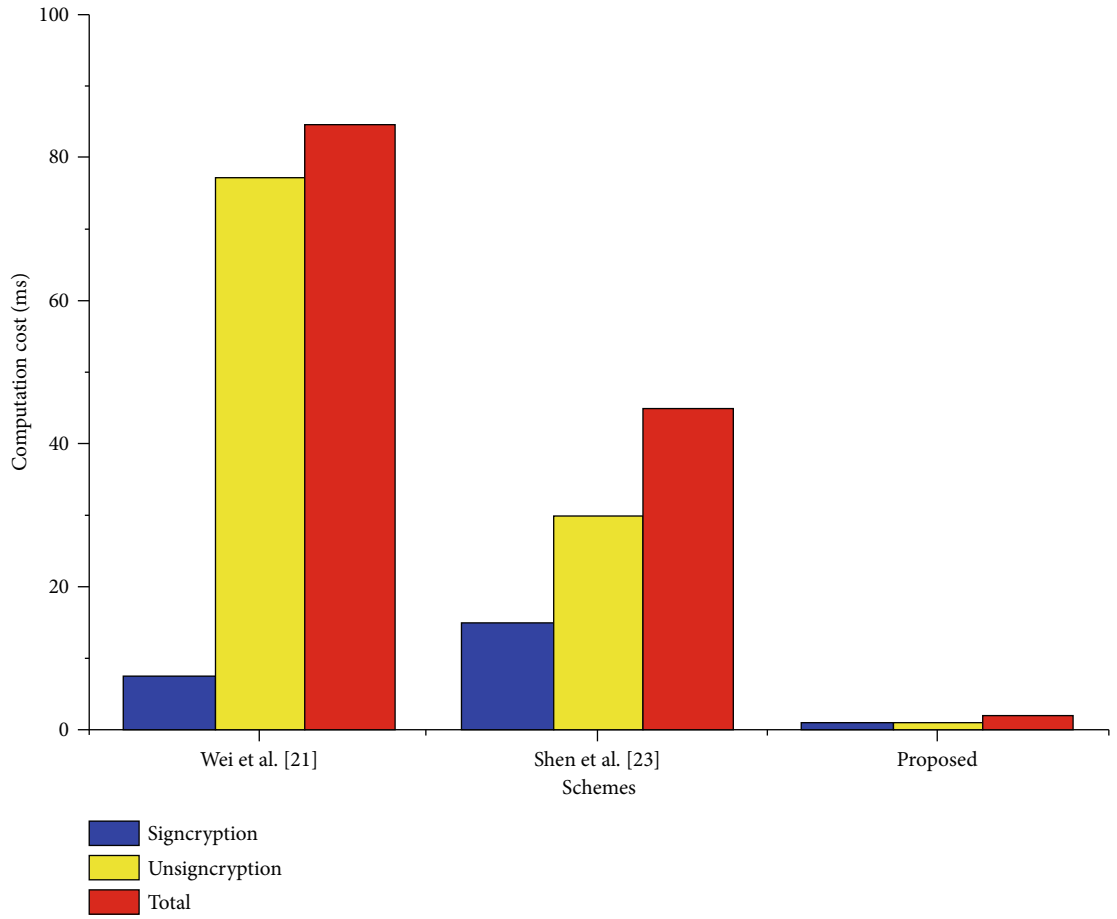


Figure 5: Computational cost reduction from ID-BGS schemes.

Table 6: Communication cost comparison with CGS schemes.

Schemes	Communication cost	Ciphertext size
Zhang et al. [28]	$1 \mid m \mid +3 \mid Q \mid = 1\|100\| + 3\|160\| = 100 + 540$	640 bits
Zhou et al. [29]	$1 \mid m \mid +3 \mid Q \mid = 1\|100\| + 3\|160\| = 100 + 540$	640 bits
Zhang et al. [30]	$1 \mid m \mid +2 \mid Q \mid = 1\|100\| + 2\|160\| = 100 + 360$	460 bits
Zhou [31]	$1 \mid m \mid +3 \mid Q \mid = 1\|100\| + 3\|160\| = 100 + 540$	640 bits
Waheed et al. [32]	$1 \mid m \mid +3 \mid G \mid = 1\|100\| + 3\|1024\| = 100 + 3072$	3172 bits
Karati et al. [33]	$1 \mid m \mid +4 \mid G \mid = 1\|100\| + 4\|1024\| = 100 + 4096$	4196 bits
Proposed	$1 \mid m \mid +3 \mid N \mid = 1\|100\| + 3\|80\| = 100 + 240$	340 bits

(ii) Our computation cost reduction from Shen et al. [23]:

$$\left(\frac{5\,\text{SEXP} + 4\,\text{SBP} - 4\,\text{SHEDM}}{5\,\text{SEXP} + 4\,\text{SBP}}\right) * 100 = \left(\frac{44.7 - 1.92}{44.7}\right) * 100 = 95.70\% \tag{11}$$

4.2.2. Communication Cost. In this section, we compare our proposed scheme with existing CGS and ID-BGS schemes in terms of bits. For this purpose, we suppose elliptic curve $|Q| = 160$ bits, bilinear pairing $|G| = 1024$ bits, hyperelliptic curve $|N| = 80$ bits, and message $|M| = 100$ bits. According to our suppositions, for CGS schemes, the communication cost of the Zhang et al. [28] scheme is $1 \mid m \mid +3|Q|$, of the Zhou et al. [29] scheme is $1 \mid m \mid + 3|Q|$, of the Zhang et al. [30] scheme is $1 \mid m \mid +2|Q|$, of the Zhou [31] scheme is $1 \mid m \mid +3|Q|$, of the Waheed et al. [32] scheme is $1 \mid m \mid +3|G|$, and of the Karati et al. [33] scheme is $1 \mid m \mid +4|G|$, and the communication cost of our proposed scheme is $1 \mid m \mid +3|N|$. Furthermore, Table 6 shows the efficiency of our scheme from Zhang et al. [28], Zhou et al. [29], Zhang et al. [30], Zhou [31], Waheed et al. [32], and Karati et al. [33]. Moreover, a clear communicational cost reduction is shown in Figure 6.

Furthermore, for the ID-BGS schemes, the communication cost of the Wei et al. [21] scheme is $1 \mid m \mid +4 \mid Q \mid$ and of the Shen et al. [23] scheme is $1 \mid m \mid +7 \mid Q \mid$. Furthermore, Table 7 shows the efficiency of our scheme from Wei et al. [21] and Shen et al. [23]. Additionally, a clear communicational cost reduction is shown in Figure 7.

(1) Communication Cost Reduction of Our Scheme from CGS Schemes. The following formula can be used to calculate the cost reduction.

$$\left(\frac{\text{existing scheme} - \text{our scheme}}{\text{existing scheme}}\right) * 100. \tag{12}$$

(i) Our communication cost reduction from Zhang et al. [28]:

$$\left(\frac{1 \mid m\mid+3 \mid Q\mid-1 \mid m\mid+3 \mid N \mid}{1 \mid m\mid+3 \mid Q \mid}\right) * 100 = \left(\frac{1|100| + 3|160| - 1|100| + 3|80|}{1|100| + 3|160|}\right) * 100 = \left(\frac{640\,\text{bits} - 340\,\text{bits}}{640\,\text{bits}}\right) * 100 = 46.87\% \tag{13}$$

(ii) Our communication cost reduction from Zhou et al. [29]:

$$\left(\frac{1 \mid m\mid+3 \mid \text{Q}\mid-1 \mid m\mid+3 \mid N \mid}{1 \mid m\mid+3 \mid Q \mid}\right) * 100 = \left(\frac{1|1\text{OO}| + 3|160| - 1|100| + 3|80|}{1|100| + 3|160|}\right) * 100 = \left(\frac{640\,\text{bits} - 340\,\text{bits}}{640\,\text{bits}}\right) * 100 = 46.87\% \tag{14}$$

(i) Our communication cost reduction from Zhang et al. [30]:

$$\left(\frac{1 \mid m\mid+2 \mid Q\mid-1 \mid m\mid+3 \mid N \mid}{1 \mid m\mid+2 \mid Q \mid}\right) * 100 = \left(\frac{1|1\text{OO}| + 2|160| - 1|1\text{OO}| + 3|80|}{1|1\text{OO}| + 3|160|}\right) * 100 = \left(\frac{460\,\text{bits} - 340\,\text{bits}}{460\,\text{bits}}\right) * 100 = 21.73\% \tag{15}$$

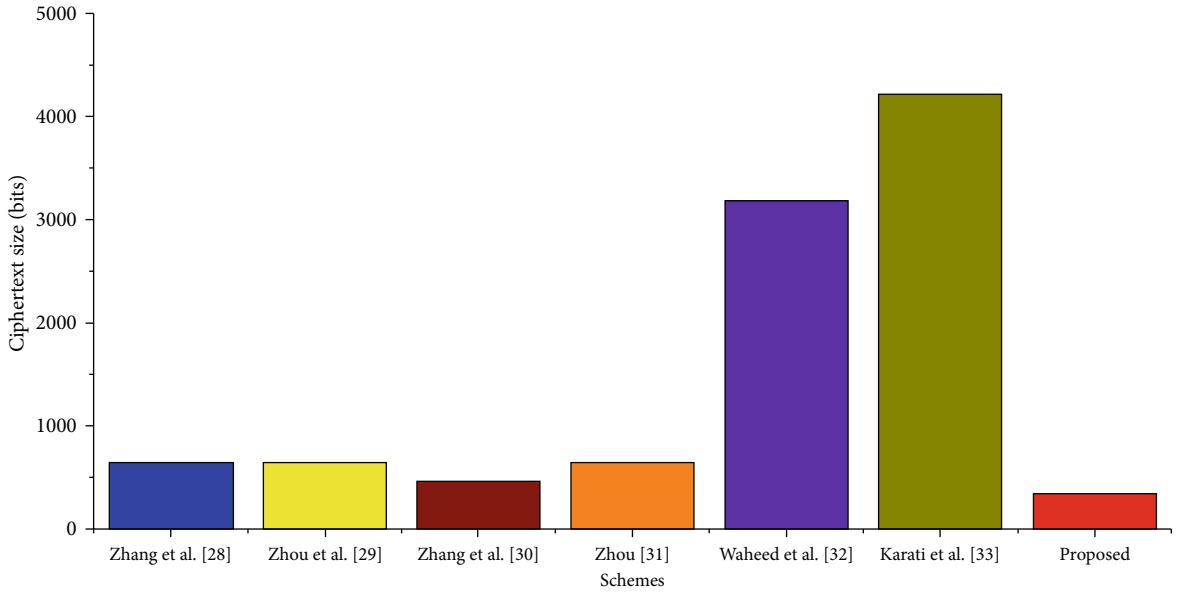


Figure 6: Communication cost reduction from CGS schemes.

Table 7: Communication cost comparison with CGS schemes.

Schemes	Communication cost	Ciphertext size
Wei et al. [21]	$1 \mid m \mid +4 \mid Q \mid = 1\|100\| + 4\|160\| = 100 + 640$	740 bits
Shen et al. [23]	$1 \mid m \mid +7 \mid Q \mid = 1\|100\| + 7\|160\| = 100 + 1120$	1220 bits
Proposed	$1 \mid m \mid +3 \mid N \mid = 1\|100\| + 3\|80\| = 100 + 240$	340 bits

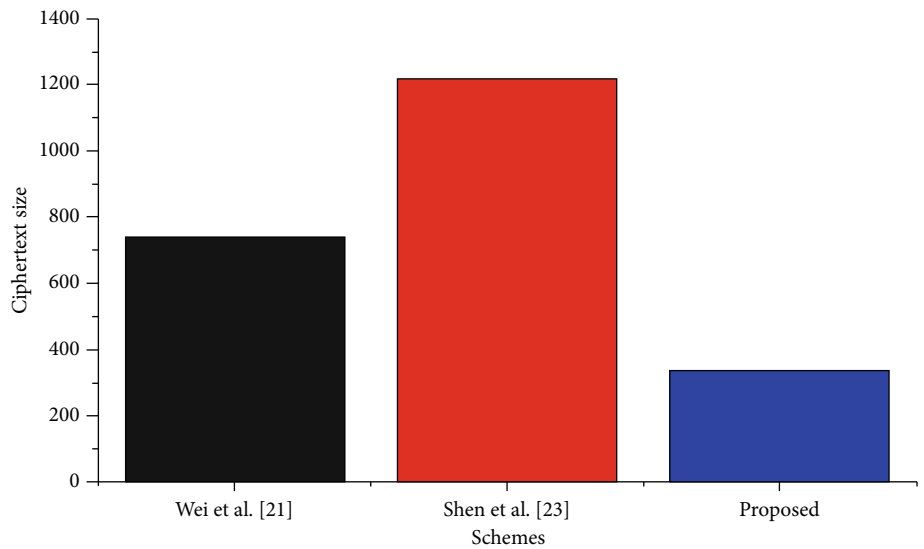


Figure 7: Communication cost reduction from ID-BGS schemes.

(ii) Our communication cost reduction from Zhou [31]:

$$\left(\frac{1\,|\,m|+3\,|\,Q|-1\,|\,\mathrm{m}|+3\,|\,\mathrm{N}\,|}{1\,|\,\mathrm{m}|+3\,|\,\mathrm{Q}\,|}\right)*100 = \left(\frac{1|1\mathrm{OO}|+3|160|-1|1\mathrm{OO}|+3|80|}{1|1\mathrm{OO}|+3|160|}\right)*100 = \left(\frac{640\text{ bits}-340\text{ bits}}{640\text{ bits}}\right)*100 = 46.87\% \tag{16}$$

(iii) Our communication cost reduction from Waheed et al. [32]:

$$\left(\frac{1\,|\,\mathrm{m}|+3\,|\,\mathrm{G}|-1\,|\,m|+3\,|\,N\,|}{1\,|\,m|+3\,|\,G\,|}\right)*100 = \left(\frac{1|1\mathrm{OO}|+3|1024|-1|1\mathrm{OO}|+3|80|}{1|1\mathrm{OO}|+3|1024|}\right)*100 = \left(\frac{3172\text{ bits}-340\text{ bits}}{3172\text{ bits}}\right)*100 = 89.28\% \tag{17}$$

(iv) Our communication cost reduction from Karati et al. [33]:

$$\left(\frac{1\,|\,m|+4\,|\,G|-1\,|\,m|+3\,|\,N\,|}{1\,|\,m|+4\,|\,G\,|}\right)*100 = \left(\frac{1|1\mathrm{OO}|+4|1024|-1|1\mathrm{OO}|+3|80|}{1|1\mathrm{OO}|+4|1024|}\right)*100 = \left(\frac{4196\text{ bits}-340\text{ bits}}{4196\text{ bits}}\right)*100 = 91.89\% \tag{18}$$

(2) Communication Cost Reduction of Our Scheme from ID-BGS Schemes.

(i) Our communication cost reduction from Wei et al. [21]:

$$\left(\frac{1\,|\,m|+4\,|\,Q|-1\,|\,m|+3\,|\,N\,|}{1\,|\,m|+4\,|\,Q\,|}\right)*100 = \left(\frac{1|100|+4|160|-1|1\mathrm{OO}|+3|80|}{1|100|+4|160|}\right)*100 = \left(\frac{740\text{ bits}-340\text{ bits}}{740\text{ bits}}\right)*100 = 54.05\% \tag{19}$$

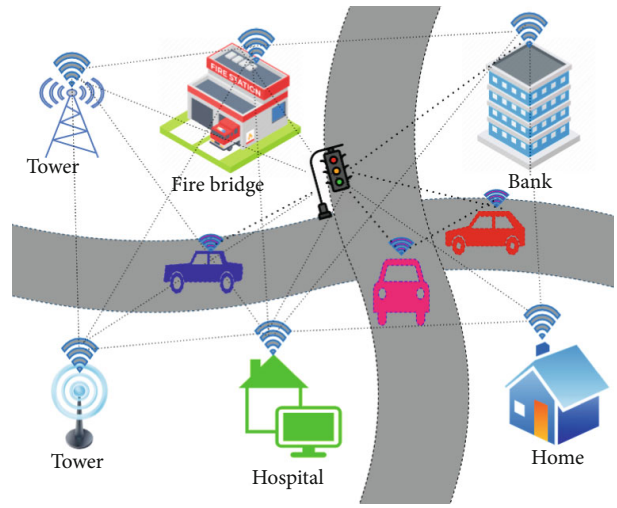


Figure 8: Deployment in smart city.

(ii) Our communication cost reduction from Shen et al. [23]:

$$\left(\frac{1\,|\,m|+7\,|\,Q|-1\,|\,m|+3\,|\,N\,|}{1\,|\,m|+7\,|\,Q\,|}\right)*100 = \left(\frac{1|1\mathrm{OO}|+7|160|-1|1\mathrm{OO}|+3|80|}{1|1\mathrm{OO}|+7|160|}\right)*100 = \left(\frac{1220\text{ bits}-340\text{ bits}}{1220\text{ bits}}\right)*100 = 72.13\% \tag{20}$$

5. Practical Scenario on NDN-Based Smart City

Assume an NDN-based smart city, where the number of sensors deployed for monitoring environmental conditions is shown in Figure 8. The sensors can monitor some emergency parameters such as fire, leakage of water, and vehicle accident, which require authentication as well as confidentiality. Furthermore, these sensors can sense some normal parameters (e.g., temperature, humidity, and energy consumption) which require authentication only.

These sensed parameters are forwarded through NDN routers using the following transmission modes.

(1) *Pull-based mode*: in this mode, a consumer sends an interest in some content/message. The sensor nodes provide the requested contents according to given interest.

(2) *Push-based mode*: in this mode, the sensor nodes intermittently forward content/message without receiving any interests of the consumer. This mode better suits the secure transfer of emergency contents/messages to a specific destination in run time.

Our deployment consists of entities such as KGC (authorization provider), content/message producer (sensors and

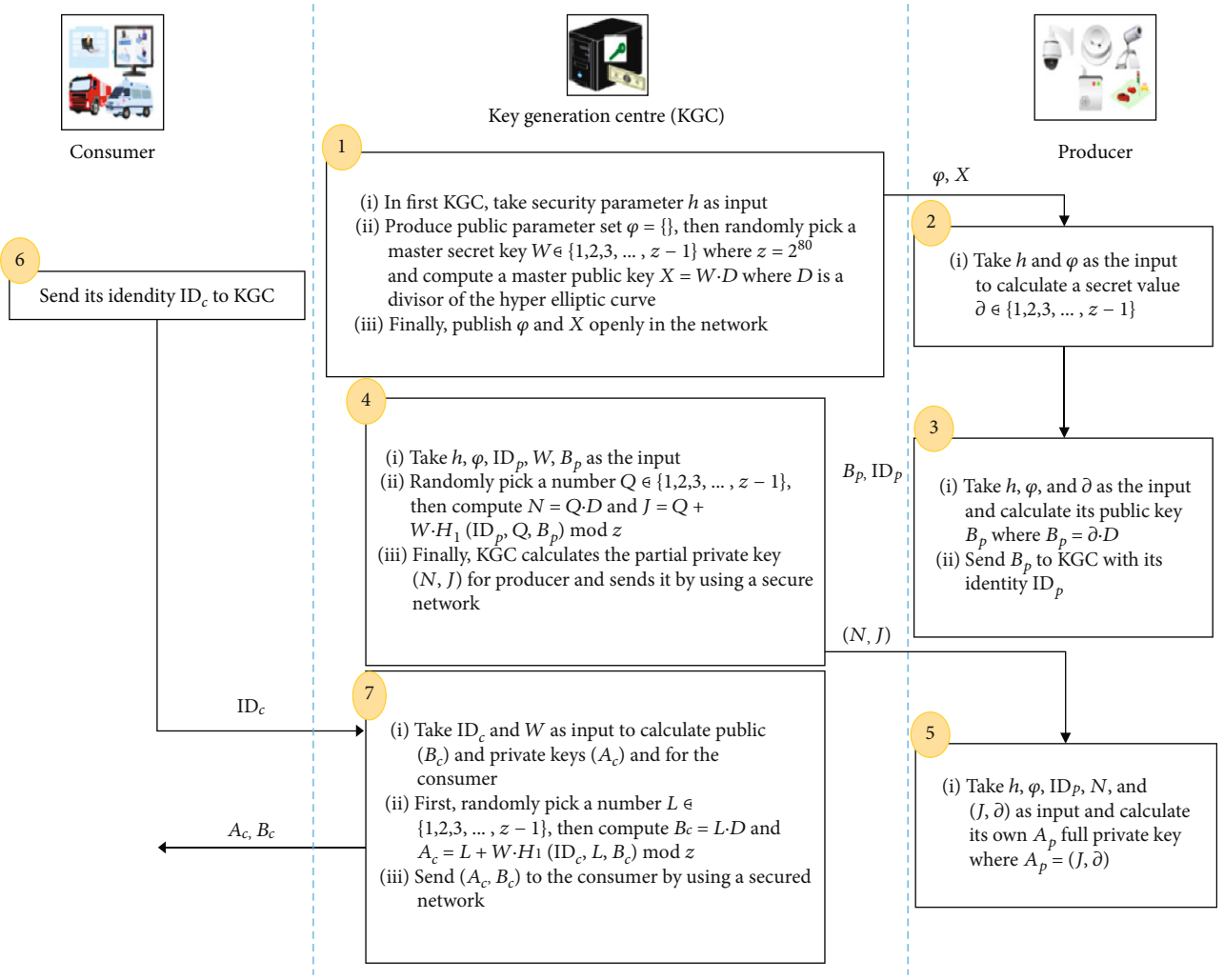


Figure 9: Registration and key generation process.

NDN router), and consumer (mobile user, fire centre, hospital, etc.).

The overall process is discussed below.

5.1. Registration and Key Generation Phase. In Figure 9, we explain the registration and key generation of consumers and providers. In step 1, the KGC takes security parameters $\mathscr{k}$ as input and produces public parameter set φ for generating master secret key $\mathscr{W} \in \{1, 2, 3, \cdots, z-1\}$ and master public key $\mathscr{X} = \mathscr{W} \cdot \mathscr{D}$. Then, publish φ and $\mathscr{X}$ in the entire network. In step 2, the producer takes (k, φ) as an input and generates a secret value $\partial \in \{1, 2, 3, \cdots, \varkappa - 1\}$.

In step 3, the producer then takes the parameters $(\mathscr{k}, \varphi, \partial)$ as input and computes its public key $\mathscr{B}_p = \partial \cdot \mathscr{D}$. After computing $\mathscr{B}_p$, the producer sends it alongside with his identity ID_p to the KGC. In step 4, after receiving the $\mathscr{B}_p$ and ID_p, the KGC takes $(\mathscr{k}, \varphi, ID_p, \mathscr{W}, \mathscr{B}_p)$ as input and randomly picks a number from $\mathscr{Q} \in \{1, 2, 3, \cdots, z-1\}$, computes $\mathscr{N} = \mathscr{Q} \cdot \mathscr{D}$ and $\mathscr{J} = \mathscr{Q} + \mathscr{W} \cdot \mathscr{H}_1(ID_p, \mathscr{Q}, \mathscr{B}_p)$, and generates a partial private key $(\mathscr{N}, \mathscr{J})$ for the producer. The KGC then sends $(\mathscr{N}, \mathscr{J})$ to the producer using a secure network. In step 5, upon receiving $(\mathscr{N}, \mathscr{J})$, the producer takes $(\mathscr{k}, \varphi, ID_p, \mathscr{N}, \mathscr{J}, \partial)$ as an input and computes its own full private key $(\mathscr{A}_p)$.

In step 6, the consumer sends the identity ID_c to KGC for registration. In step 7, upon receiving the ID_c, the KGC takes $(ID_c, \mathscr{W})$ as input and randomly picks a number from $\mathscr{L} \in \{1, 2, 3, \cdots, \varkappa - 1\}$ to calculate the public key $(\mathscr{B}_c)$ and private key $(\mathscr{A}_c)$ for the consumer. The KGC then sends (B_c, A_c) to the consumer using a secure channel.

5.2. Communication Phase. In Figure 10, we explain the secure communication of the consumer and provider after a successful registration and key generation phase. If the consumer wants the signed/signcrypted contents from the producer or the producer wants to deliver signed/signcrypted contents to the consumer securely, first, for the signcrypted content, the producer takes content (m) and$(ID_c, \mathscr{B}_c, \mathscr{A}_p, \mathscr{X})$ with a randomly picked number from $\mathscr{R} \in \{1, 2, 3, \cdots, z-1\}$; computes the secret value δ, hash of $(ID_c, \mathscr{B}_c, \mathscr{X})$, a fresh nonce $\mathscr{T}$, encrypted contents $\mathscr{C} = \mathscr{E}_K(\mathscr{T}, m)$, and a hash of the encrypted contents $\mu = \mathscr{H}_2(\mathscr{T}, m)$; and applies signature $\mathscr{S} = \mathscr{R} + \mu(\mathscr{J} + \partial)$ on it. Finally, generate the signcrypted contents $\Psi = (\mathscr{C}, \mu, \mathscr{S}, \delta)$ and send it to

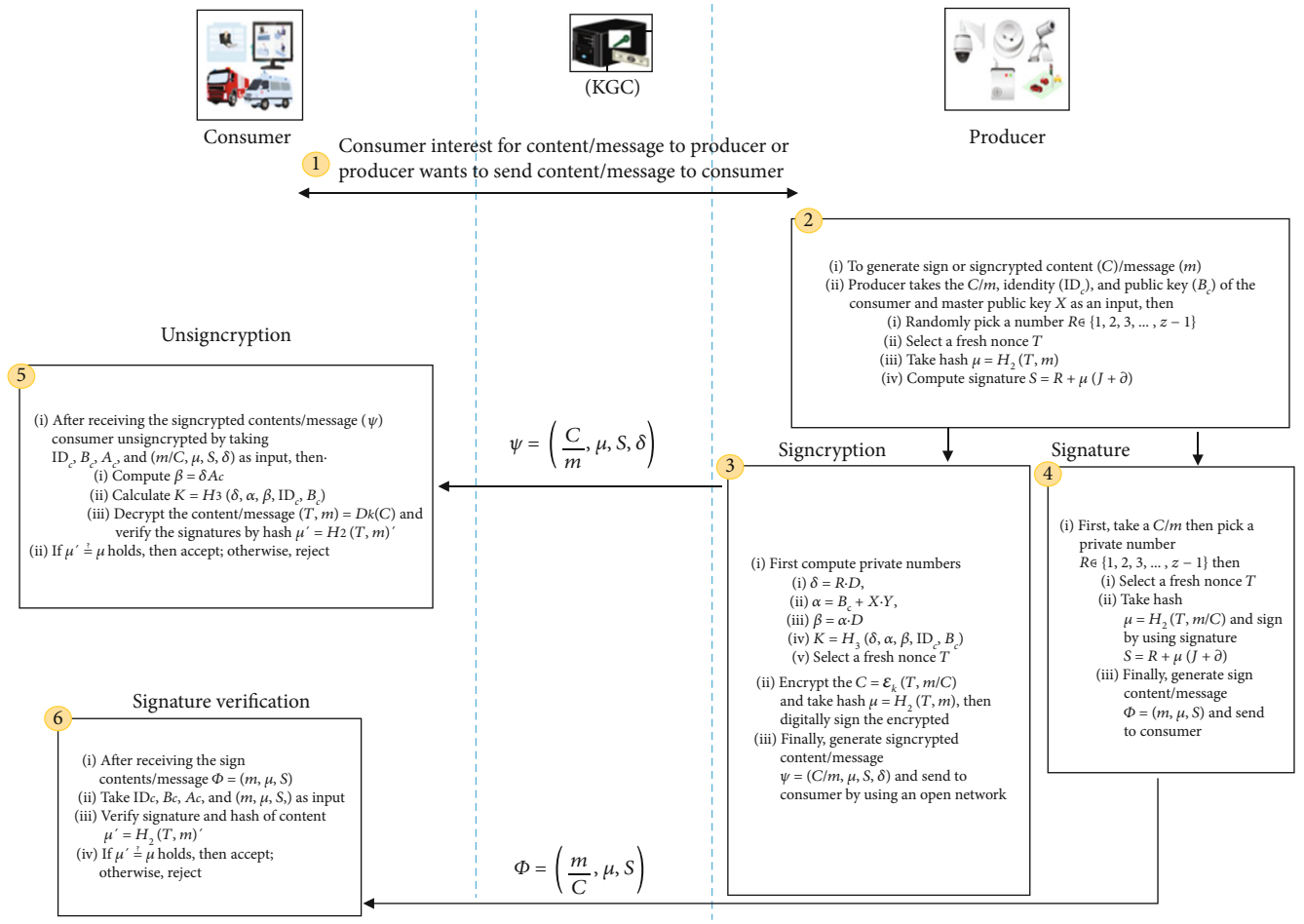


Figure 10: Communication process of the proposed scheme.

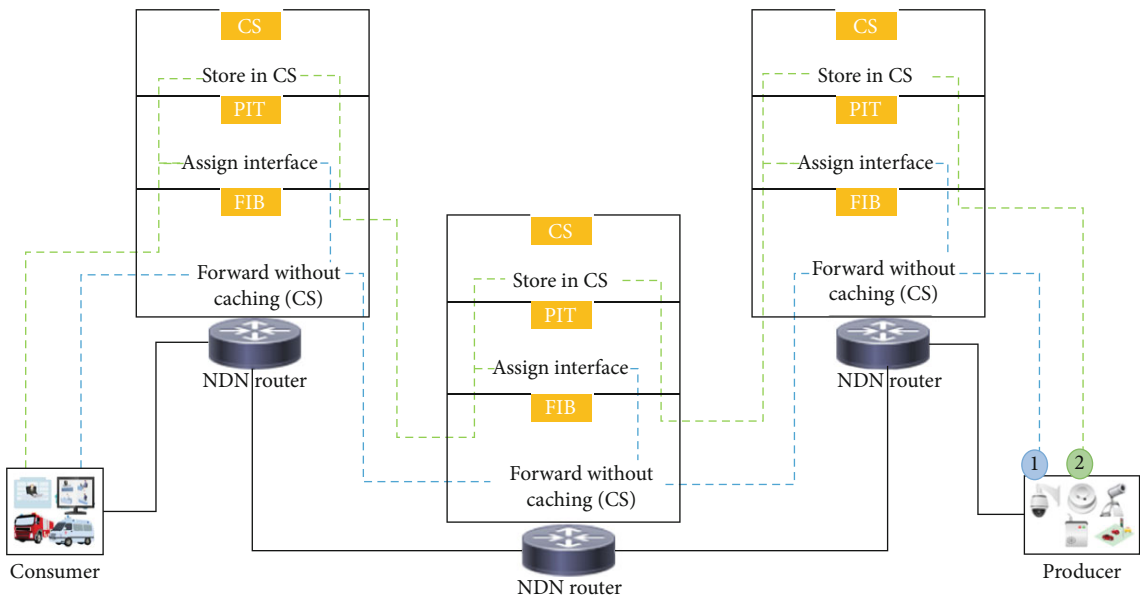


Figure 11: Workflow process of the proposed scheme in NDN architecture.

the consumer. For signed contents, the producer takes content (m) with randomly picked numbers from $\mathscr{R} \in \{1, 2, 3, \cdots, z-1\}$ and selects a fresh nonce $\mathscr{T}$, takes hash of $\mu = \mathscr{H}_2(\mathscr{T}, m)$, and applies signature $\mathscr{S} = \mathscr{R} + \mu(\mathscr{J} + \partial)$. Finally, it generates signed contents $\Phi = (m, \mu, \mathscr{S}, \delta)$ and sends it to the consumer.

After receiving the signcrypted contents Ψ, the consumer unsigncrypts the contents by taking $(ID_c, \mathscr{B}_c, \mathscr{A}_c, \mathscr{C}, \mu, \mathscr{S},)$ as an input and computing $\beta = \delta \cdot \mathscr{A}_c$, calculates the hash of signature $K = \mathscr{H}_3(\delta, \alpha, \beta, ID_c, \mathscr{B}_c)$, decrypts the content $(\mathscr{T}, m) = D_K(\mathscr{C})$, and computes the hash of the content $\mu' = \mathscr{H}_2(\mathscr{T}, m)'$; if $\mu' \stackrel{?}{=} \mu$ holds, then the contents are accepted; otherwise, they are rejected. In the case of signed contents Φ, the consumer takes $(ID_c, \mathscr{B}_c, \mathscr{A}_c, \mathscr{C}, \mu, \mathscr{S}, \delta)$ as input and calculates hash $\mu' = \mathscr{H}_2(\mathscr{T}, m)'$; if $\mu' \stackrel{?}{=} \mu$ holds, then the contents are accepted; otherwise, they are rejected.

5.3. The Workflow in NDN Architecture. NDN provides in-network caching, which means that the router of NDN will store and forward every message. Here, we divide the overall scenario into two types such as emergency situation and routine-based situation. In case of an emergency situation (fire, leakage of water, vehicle accident, etc.) that requires signcryption (confidentiality and authentication) for successful delivery to the intended destination in run time, the signcryption algorithm will execute and the NDN routers must not store these messages in the CS as shown in step 1 (Figure 11). The storage of emergency messages in CS does not facilitate any consumer later with the expense of latency.

In the routine-based situation, some parameters like, e.g., temperature, humidity, energy consumption, and video streaming, require authentication only and facilitate a number of consumers at a time. For this type of situation, the signature algorithm will execute and the NDN routers will store the copy of these contents/messages in its CS for future use as shown in step 2.

6. Conclusion

In this paper, we introduce the concept of lightweight in a natural heterogeneous generalized signcryption for the NDN-based Internet of Things (IoT). The proposed scheme provides the security properties of unforgeability, confidentiality, forward secrecy, and antireplay attack. We did the computation and communication cost comparisons with existing schemes, and the results give a satisfactory output due to the use of the hyperelliptic curve. So, our scheme reduced the computation cost of certificateless generalized signcryption (CGS) schemes from 78.09 to 97.23% and the communication from 21.73 to 91.89%. Furthermore, our scheme reduced the computation cost of identity-based generalized signcryption (ID-BGS) schemes from 95.70 to 97.84% and the communication cost from 54.05 to 72.13%. In addition, we practically deployed our scheme in the NDN-based smart city. Additionally, the scheme is validated through a security verification tool called AVISPA. The simulation results show that our scheme is valid and safe under the back-end protocols (OFMC, ATSE) of AVISPA.

Appendix

In this section, we discuss the simulation and validation of our proposed scheme in AVISPA. The simulation tool, code, and results are shown in the subsection below.

A. Automated Validation of Internet Security Protocols and Applications (AVISPA)

Automated Validation of Internet Security Protocols and Applications (AVISPA) is a security simulation tool used to check the validity of cryptographic schemes [39]. The AVISPA tools work under two states such as "safe" if the scheme resists against security threats and "unsafe" if the scheme cannot resist against security threats. AVISPA uses a role-oriented language called a high-level protocol specification language (HLPSL) for a specification of a cryptographic scheme. For checking the security, the user needs to convert the pseudocode of the proposed algorithm into the HLPSL. Then, the HLPSL2IF translator translates it to the intermediate format (IF). HLPSL2IF then verifies the security of the proposed scheme under four back-end tools called on-the-fly model checker (OFMC), CL-based attack searcher (CL-AtSe), SAT-based model checker (SATMC), and tree-automata-based protocol analyzer (TA4SP). According to the requirement of the scheme, each backed tool has its own functionality as further discussed in [40, 41], as shown in Figure 12.

B. Simulation Code

Here, we divide the simulation code according to the entities that participate in our scheme such as the producer and consumer. Note: for simulation of the proposed algorithm, the pseudocode of the proposed algorithm needs to be changed for the HLPSL library. Moreover, the notation used in the proposed algorithm is different as compared to the notation used in HLPSL. Further, the simulation code is shown in Pseudocodes B.1 and B.2.

C. Simulation Results

This section contains the simulation results of the proposed scheme according to the back-end protocols of the AVISPA tool such as OFMC and ATSE.

C.1. OFMC. The results of the proposed scheme after applying the OFMC protocol show that our scheme is safe against malicious attacks as shown in Figure 13.

C.2. ATSE. The results of the proposed scheme after applying the ATSE protocol show that our scheme is safe against malicious attacks as shown in Figure 14.

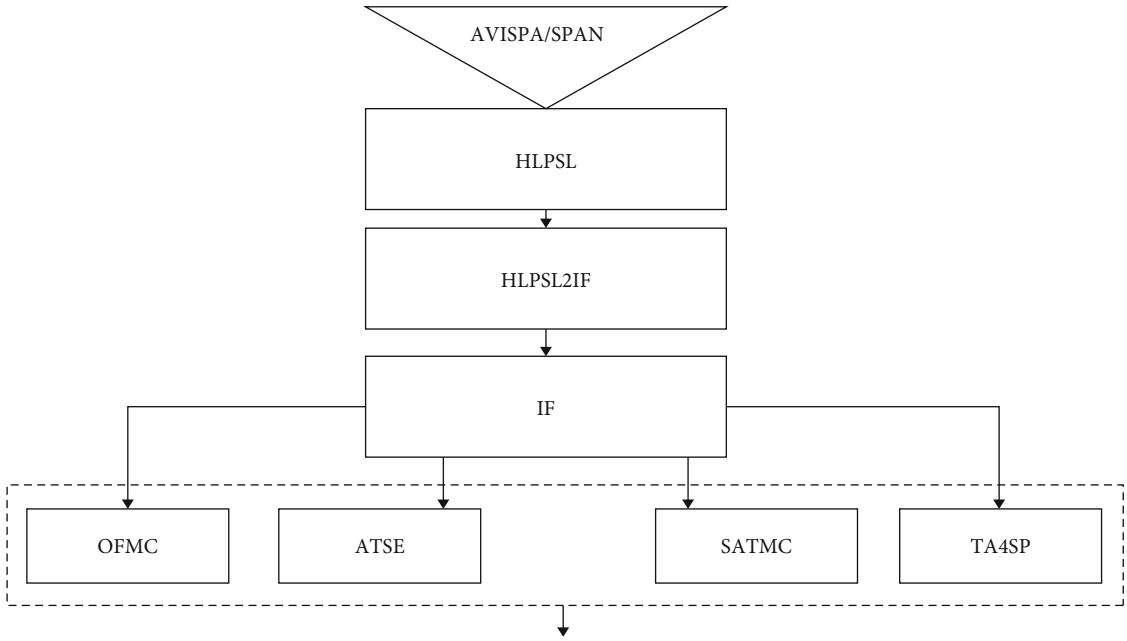


Figure 12: AVISPA workflow.

```
role role_Producer(Producer:agent, Consumer:agent, Bp:public_key, Bc:public_key, SND, RCV:channel(dy))
played_by Producer
def=
  local
    State:nat, T:text, Plus:hash_func, R:text, U:text, Mmm:text, Encryptionnn:hash_func, Kk:symmetric_key
  init
    State := 0
  transition
    1. State=0 /\ RCV(start) = |> State':=1 /\ SND(Producer.Consumer)
    2. State=1 /\ RCV(Consumer.{T'}_Bc) = |> State':=2 /\ U':=new() /\ R':=new() /\ Kk':=new() /\ Mmm':=new() /\ SND(Produ-
cer.{Encryptionnn(Mmm')}_Kk'.{Plus(R'.U')}_inv(Bp))
end role
```

Pseudocode B.1: HLPSL code for producer role.

```
role role_Consumer(Producer:agent, Consumer:agent, Bp:public_key, Bc:public_key, SND, RCV:channel(dy))
played_by Consumer
def=
  local
  State:nat, T:text, Plus:hash_func, R:text, U:text, Mmm:text, Encryptionnn:hash_func, Kk:symmetric_key
  init
    State := 0
  transition
    1. State=0 /\ RCV(Producer.Consumer) = |> State':=1 /\ T':=new() /\ SND(Consumer.{T'}_Bc)
    6. State=1 /\ RCV(Producer.{Encryptionnn(Mmm')}_Kk'.{Plus(R'.U')}_inv(Bp)) = |> State':=2
end role
```

Pseudocode B.2: HLPSL code for consumer role.

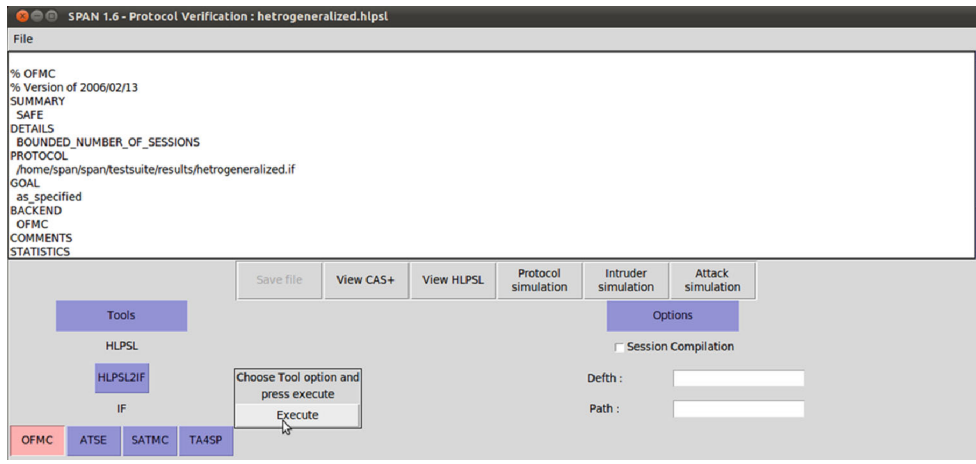


Figure 13: OFMC protocol result of proposed scheme.

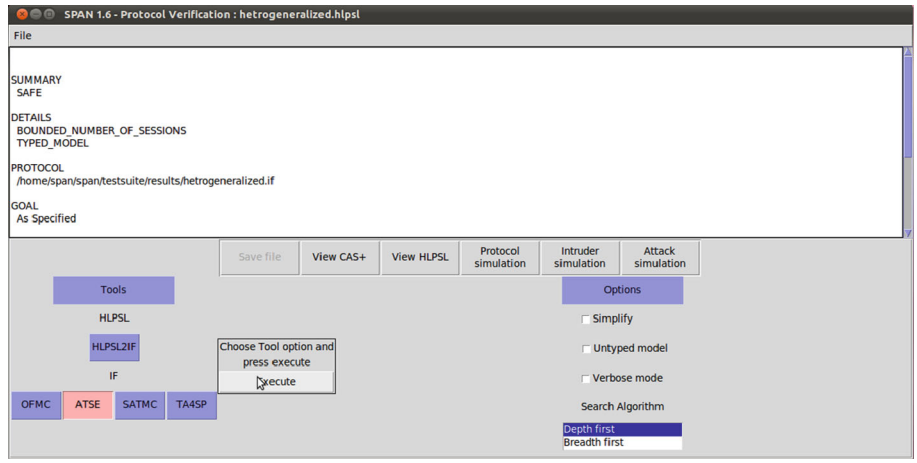


Figure 14: ATSE protocol result of proposed scheme.

Data Availability

All data generated or analyzed during this study are included in this published article.

Conflicts of Interest

The authors declare no conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] D. Mars, S. Mettali Gammar, A. Lahmadi, and L. Azouz Saidane, "Using information centric networking in internet of things: a survey," *Wireless Personal Communications*, vol. 105, no. 1, pp. 87–103, 2019.

[2] A. Khanna and S. Kaur, "Evolution of internet of things (IoT) and its significant impact in the field of precision agriculture," *Computers and electronics in agriculture*, vol. 157, pp. 218–231, 2019.

[3] V. Jacobson, D. K. Smetters, J. D. Thornton, and M. F. Plass, "Networking named content," in *CoNEXT '09: Proceedings of the 5th international conference on Emerging networking experiments and technologies*, pp. 1–12, Rome, Italy, 2009.

[4] C. Fang, F. Yu, T. Huang, J. Liu, and Y. Liu, "A survey of energy-efficient caching in information-centric networking," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 122–129, 2014.

[5] M. Amadeo, G. Ruggeri, C. Campolo, and A. Molinaro, "IoT services allocation at the edge via named data networking: from optimal bounds to practical design," *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 661–674, 2019.

[6] C. Fang, F. R. Yu, T. Huang, J. Liu, and Y. Liu, "A survey of green information-centric networking: research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1455–1472, 2015.

[7] Q. Huang, D. S. Wong, and G. Yang, "Heterogeneous signcryption with key privacy," *The Computer Journal*, vol. 54, no. 4, pp. 525–536, 2011.

[8] Y. Han, X. Yang, P. Wei, Y. Wang, and Y. Hu, "ECGSC: elliptic curve based generalized signcryption," in *Ubiquitous Intelligence and Computing*, pp. 956–965, Springer Berlin Heidelberg, 2006.

[9] M. Suárez-Albela, P. Fraga-Lamas, and T. Fernández-Caramés, "A practical evaluation on RSA and ECC-based cipher suites for IoT high-security energy-efficient fog and mist computing devices," *Sensors*, vol. 18, p. 3868, 2018.

[10] M. Yu, J. Zhang, J. Wang et al., "Internet of things security and privacy-preserving method through nodes differentiation, concrete cluster centers, multi-signature, and blockchain," *International Journal of Distributed Sensor Networks*, vol. 14, 2018.

[11] A. Braeken, "PUF based authentication protocol for IoT," *Symmetry*, vol. 10, no. 8, 2018.

[12] I. Ullah, N. Ul Amin, M. Zareei et al., "A lightweight and provable secured certificateless signcryption approach for crowdsourced IIoT applications," *Symmetry*, vol. 11, p. 1386, 2019.

[13] S. Kumari, M. Karuppiah, A. K. Das, X. Li, F. Wu, and N. Kumar, "A secure authentication scheme based on elliptic curve cryptography for IoT and cloud servers," *The Journal of Supercomputing*, vol. 74, no. 12, pp. 6428–6453, 2018.

[14] Z. Ullah, A. Zeb, I. Ullah et al., "Certificateless proxy reencryption scheme (CPRES) based on hyperelliptic curve for access control in content-centric network (CCN)," *Mobile Information Systems*, vol. 2020, Article ID 4138516, p. 13, 2020.

[15] C. Tamizhselvan and V. Vijayalakshmi, "An energy efficient secure distributed naming service for IoT," *International Journal of Advanced Studies of Scientific Research*, vol. 3, no. 8, 2019.

[16] V. S. Naresh, R. Sivaranjani, and N. V. E. S. Murthy, "Provable secure lightweight hyper elliptic curve-based communication system for wireless sensor network," *International Journal of Communication Systems*, vol. 31, no. 15, article e3763, 2018.

[17] A. Rahman, I. Ullah, M. Naeem, R. Anwar, H. S. Khaǰak, and A. Ullah, "Lightweight multi-message and multi-receiver heterogeneous hybrid signcryption scheme based on hyper elliptic curve," *International Journal of Advanced Computer Science and Applications*, vol. 9, p. 5, 2018.

[18] S. Lal and P. Kushwah, "ID based generalized signcryption," Cryptology ePrint Archive, Report, 2008, http://eprint.iacr.org.

[19] W. Liang, Z. Chuan-Rong, and L.-Q. Zheng, "A key management scheme based generalized Signcryption in mobile ad hoc network," in *2010 International Conference on Communications and Intelligence Information Security*, Nanning, China, 2010.

[20] P. Kushwah and S. Lal, "An efficient identity based generalized signcryption scheme," *Theoretical Computer Science*, vol. 412, no. 45, pp. 6382–6389, 2011.

[21] G. Wei, J. Shao, Y. Xiang, P. Zhu, and R. Lu, "Obtain confidentiality or/and authenticity in big data by ID-based generalized signcryption," *Information Sciences*, vol. 318, pp. 111–122, 2015.

[22] D. Mishra and S. Singh, "A survey on ID based and certificateless generalized signcryption scheme," *International Journal of Innovative Research in Advanced Engineering*, vol. 2, no. 11, 2014.

[23] X. Shen, Y. Ming, and J. Feng, "Identity based generalized signcryption scheme in the standard model," *Entropy*, vol. 19, no. 3, p. 121, 2017.

[24] A. Waheed, A. I. Umar, N. Din, N. U. Amin, S. Abdullah, and P. Kumam, "Cryptanalysis of an authentication scheme using an identity based generalized signcryption," *Mathematics*, vol. 7, no. 9, p. 782, 2019.

[25] J. Huifang, H. Wenbao, and Z. Long, "Certificateless generalized signcryption," Cryptology ePrint Archive, Report, 2010, http://eprint.iacr.org.

[26] P. Kushwah and S. Lal, "Provable secure certificateless generalized signcryption scheme," *Technology & Applications*, vol. 3, pp. 925–939, 2012.

[27] C. Zhou, W. Zhou, and X. Dong, "Provable certificateless generalized signcryption scheme," *Designs, Codes and Cryptography*, vol. 71, no. 2, pp. 331–346, 2014.

[28] A. Zhang, L. Wang, X. Ye, and X. Lin, "Light-weight and robust security-aware D2D-assist data transmission protocol for mobile-health systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 3, pp. 662–675, 2016.

[29] M. A. Khan, I. Ullah, S. Nisar et al., "An efficient and provably secure certificateless key-encapsulated signcryption scheme for flying ad-hoc network," *IEEE Access*, vol. 8, pp. 36807–36828, 2020.

[30] B. Zhang, Z. Jia, and C. Zhao, "An efficient certificateless generalized signcryption scheme," *Security and Communication Networks*, vol. 2018, Article ID 3578942, 2018.

[31] C. Zhou, "An improved lightweight certificateless generalized signcryption scheme for mobile-health system," *International Journal of Distributed Sensor Networks*, vol. 15, no. 1, 2019.

[32] A. Waheed, J. Iqbal, N. Din, S. Ul, A. Iqbal, and N. Ul, "Improved cryptanalysis of provable certificateless generalized signcryption," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 4, 2019.

[33] A. Karati, C. I. Fan, and R. H. Hsu, "Provably secure and generalized Signcryption with public verifiability for secure data transmission between resource-constrained IoT devices," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10431–10440, 2019.

[34] Y. Li, C. Wang, Y. Zhang, and S. Niu, "Privacy-preserving multi-receiver signcryption scheme for heterogeneous systems," *Security and Communication Networks*, vol. 9, no. 17, 4584 pages, 2016.

[35] S. Raveendranath and A. Aneesh, "Efficient multi-receiver heterogenous signcryption," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 1693–1697, Chennai, India, 2016.

[36] S. Niu, Z. Li, M. Tian, C. Wang, and X. Jia, "An efficient heterogeneous signcryption scheme from certificateless to identity-based cryptosystem," *MATEC Web of Conferences*, vol. 139, article 00037, 2017.

[37] S. Niu, L. Niu, X. Yang, C. Wang, and X. Jia, "Heterogeneous hybrid signcryption for multi-message and multi-receiver," *PloS One*, vol. 12, no. 9, article e0184407, 2017.

[38] Y. Li, Y. Qi, and L. Lu, "Secure and efficient V2V communications for heterogeneous vehicle ad hoc networks," in *2017 International Conference on Networking and Network Applications (NaNA)*, pp. 93–99, Kathmandu, Nepal, 2017.

[39] S. Niu, Z. Li, and C. Wang, "Privacy-preserving multi-party aggregate signcryption for heterogeneous systems," in *International Conference on Cloud Computing and Security*, pp. 216–229, Nanjing, China, 2017.

[40] M. E. Saeed, Q. Liu, G. Tian, B. Gao, and F. Li, "HOOSC: heterogeneous online/offline signcryption for the internet of things," *Wireless Networks*, vol. 24, no. 8, pp. 3141–3160, 2018.

[41] C. Wang, C. Liu, Y. Li, H. Qiao, and L. Chen, "Multi-message and multi-receiver heterogeneous signcryption scheme for ad-hoc networks," *Information Security Journal: A Global Perspective*, vol. 26, no. 3, pp. 136–152, 2017.

[42] C. Jin, G. Chen, C. Yu, J. Shan, J. Zhao, and Y. Jin, "An efficient heterogeneous signcryption for smart grid," *PloS One*, vol. 13, no. 12, article e0208311, 2018.

[43] J. Liu, L. Zhang, R. Sun, X. Du, and M. Guizani, "Mutual heterogeneous signcryption schemes for 5G network slicings," *IEEE Access*, vol. 6, pp. 7854–7863, 2018.

[44] X. Liu and W. Ma, "CDAKA: a provably-secure heterogeneous cross-domain authenticated key agreement protocol with symptoms-matching in TMIS," *Journal of Medical Systems*, vol. 42, no. 8, 2018.

[45] A. A. Omala, A. S. Mbandu, K. D. Mutiria, C. Jin, and F. Li, "Provably secure heterogeneous access control scheme for wireless body area network," *Journal of Medical Systems*, vol. 42, no. 6, 2018.

[46] F. Zhou, Y. Li, and Y. Ding, "Practical V2I secure communication schemes for heterogeneous VANETs," *Applied Sciences*, vol. 9, no. 15, 2019.

[47] I. Ullah, N. U. Amin, M. Naeem, S. J. KhaJak, and H. Ali, "A novel provable secured signcryption scheme????: A hyperelliptic curve-based approach," *Mathematics*, vol. 7, no. 8, p. 686, 2019.

[48] S. Ullah, X.-Y. Li, and L. Zhang, "A Review of signcryption schemes based on hyper elliptic curve," in *2017 3rd International Conference on Big Data Computing and Communications (BIGCOM)*, pp. 10-11, Chengdu, China, 2017.

[49] D. Dolev and A. C. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–208, 1983.

[50] C. Zhou, Z. Zhao, W. Zhou, and Y. Mei, "Certificateless key-insulated generalized signcryption scheme without bilinear pairings," *Security and Communication Networks*, vol. 2017, Article ID 8405879, 2017.

[51] S. S. Ullah, I. Ullah, H. Khattak et al., "A lightweight identity-based signature scheme for mitigation of content poisoning attack in named data networking with internet of things," *IEEE Access*, vol. 8, pp. 98910–98928, 2020.

[52] S. Hussain, I. Ullah, H. Khattak et al., "A lightweight and formally secure certificate based Signcryption with proxy re-encryption (CBSRE) for internet of things enabled smart grid," *IEEE Access*, vol. 8, pp. 93230–93248, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8819545, 9 pages
https://doi.org/10.1155/2020/8819545

Research Article

Intrusion Detection into Cloud-Fog-Based IoT Networks Using Game Theory

Poria Pirozmand,[1] **Mohsen Angoraj Ghafary**,[2] **Safieh Siadat**,[2] **and Jiankang Ren**[3]

[1]*School of Computer and Software, Dalian Neusoft University of Information, Dalian 116023, China*
[2]*Department of Computer Engineering and Information Technology, Payame Noor University (PNU), P.O. Box 19395-4697 Tehran, Iran*
[3]*School of Computer Science and Technology, Dalian University of Technology, China*

Correspondence should be addressed to Safieh Siadat; safieh.siadat@gmail.com

Received 11 April 2020; Revised 25 September 2020; Accepted 27 October 2020; Published 16 November 2020

Academic Editor: Fawad Zaman

The Internet of Things is an emerging technology that integrates the Internet and physical smart objects. This technology currently is used in many areas of human life, including education, agriculture, medicine, military and industrial processes, and trade. Integrating real-world objects with the Internet can pose security threats to many of our day-to-day activities. Intrusion detection systems (IDS) can be used in this technology as one of the security methods. In intrusion detection systems, early and correct detection (with high accuracy) of intrusions is considered very important. In this research, game theory is used to develop the performance of intrusion detection systems. In the proposed method, the attacker infiltration mode and the behavior of the intrusion detection system as a two-player and nonparticipatory dynamic game are completely analyzed and Nash equilibrium solution is used to create specific subgames. During the simulation performed using MATLAB software, various parameters were examined using the definitions of game theory and Nash equilibrium to extract the parameters that had the most accurate detection results. The results obtained from the simulation of the proposed method showed that the use of intrusion detection systems in the Internet of Things based on cloud-fog can be very effective in identifying attacks with the least amount of errors in this network.

1. Introduction

Advances in various technologies like sensors, wireless communications, hidden computing, automatic detection and tracking, extensive Internet access, and distributed services enhance the potential for the integration of intelligent things in our daily lives through the Internet. The convergence of the Internet and intelligent things that can communicate and interact with each other is defined as the Internet of Things (IoT). [1].

However, integrating real-world smart objects with the Internet may pose security threats in many of our daily activities, too [2].

Given the wide range of standards and communication stacks, limited computing power, and the large number of interconnected devices, common security measures against threats cannot effectively operate in Internet of Things (IoT) systems. Accordingly, it is essential to develop certain security solutions by means of mathematical methods and statistical points for the IoT, to make it possible for the users of organizations to carefully analyze and detect all the weaknesses of the system in this way [3].

Due to widespread communication standards and stacks, limited computing power, and a high number of interconnected devices, common security measures against threats cannot be effective in IoT systems. For this reason, it is necessary to develop specific security solutions for the IoT, to allow users of organizations to identify all the weaknesses of the system [3].

Some of the ongoing projects to improve the security of the IoT include methods providing confidentiality of data and authentication, access control within the IoT network,

privacy, and trust between users and things, as well as the implementation of security and privacy policies [Sicari, et al., 2010]. Nevertheless, even with these methods, IoT networks are vulnerable to multiple attacks designed to disrupt and destroy these networks. Thus, one of the required defense methods is to design methods detecting attackers. Intrusion detection systems are for this purpose.

Security concepts are being considered with the rapid growth of IoT technology applications. Concerns are raised about intrusion, privacy, and people's inability to control their personal lives. If people's daily activities are monitored and they produce information outputs, political, economic, and social activities will be affected. The benefits of IoT technology will diminish in case of security breaches, attacks, or malfunctions [4].

Given the security challenges in the virtual world and the emerging technology of the IoT and due to the challenges of infiltrating these systems, it is significant to provide an optimal way to detect intrusion and maintain security in these systems.

Therefore, to deal with intruders and attackers on computer systems and networks, several methods have been developed called intrusion detection methods, responsible for monitoring the events occurring in a computer system or network. In the current study, the following sections are considered to achieve the objectives and provide an efficient mathematical model in intrusion detection systems. The research background is presented in the second section, and the statement of the problem is given in the third section. Modeling and definition of game parameters, information, and the used data are stated in the fourth section. The fifth and sixth sections present the results using the findings obtained, while analyzing, evaluating, and implementing; ultimately, the effective suggestions are presented in the seventh section.

2. Related Works

Over recent years, various papers and methods based on game theory in the field of computer network security have been published to model, analyze, and optimize the performance and efficiency of intrusion detection systems in IoT-related technologies like ad hoc mobile networks ([5]; Mishra et al., 2014), wireless sensor networks (Buton et al., 2016; [6]), cloud computing [7], and physical cyber systems [8].

The report by Moudi et al. [7] provided a variety of intrusions affecting accessibility, confidentiality, and integration in cloud computing. The authors of this reference have divided the intrusion detection system technology used in cloud into three categories: host-based, network-based, and hyper-based systems (virtual machine monitor). Moreover, they have discussed the pros and cons of each protocol and identified challenges to make cloud computing a reliable platform for providing IoT services.

The results of a study by Midi et al. [9] reveal that an intrusion detection system is able to monitor and control multiple communication protocols, a combination of signature rules, and anomaly detection processes.

Buton et al. [10] performed an extensive study of intrusion detection systems in wireless sensor networks and made a comparative analysis between the intrusion detection systems provided for wireless sensor networks given the network architecture and detection methods.

Granjal et al. [11] presented a comprehensive security analysis of several Internet protocols. More specifically, they checked IEEE802.15.4 security issues on low-power wireless regional networks (6LoWPAN), IPv6 routing protocols for low-power and lossy networks (RPL), Datagram Transport Layer Security (DTLS), and constrained application protocols (CoAP).

Goa et al. (2016) addressed a two-step hybrid approach first examining the initial diagnosis of whether or not the data is invasive using the K-means cluster and then, at the second stage, finally diagnosing the closest neighbor using the K algorithm.

Kumar and Dota [5] have examined the intrusion detection methods provided for mobile ad hoc networks through focusing on their detection algorithms. They have introduced a tree classification for intrusion detection methods based on the nature of the processing method used in the detection method.

Walgren et al. (2017) have provided an intrusion detection system for LOWPAN-RPL6 networks able to detect Sinkhole, Sybil, and Selective attacks using a hybrid approach connecting various parameters.

Atli and Jung [12] have developed an intrusion detection system based on the characteristics of the supervisor as well as the use of the leading neural network. Their paper gives a brief overview on ISCX-IDS 2012 and CIC Android. To perform the phase, SVM feature selection has been used with incremental learning; the rankings selected 20 features with the highest ranking out of 43 features in the data set, and then using the neural network, the final diagnosis was 94% to 98.7% accurate.

Shen et al. [13] have provided an optimal framework for demonstrating the potential and practical application of malware repression to protect the privacy of smart things on IoT networks through an intrusion detection system with theoretical calculation of the Bayesian game.

Pagitus et al. (2019) have investigated the security of the IoT, its challenges, threats, and its solutions. After reviewing and assessing the potential threats and determining security measures and requirements in the field of IoT, they have performed a quantitative and qualitative risk analysis examining security threats at each layer.

In their study titled "A Game Theoretic Approach to Decision and Analysis in Network Intrusion Detection," Susan and Rayford (2019) have developed a model for IDS distributed with a network of sensors, in addition to suggesting two plans independent from the flexible platform based on game theory techniques. In the presented plan, through implementing participatory game theory, Shapley values have been especially used for analysis and configuration; Nash equilibrium solutions have been obtained by means of analysis method and analyzed for the defined game security.

In their review paper, Hajiheidari et al. (2019) have comprehensively investigated the IDSS in IoT networks. The

research systematically investigates IDSS with a precise classification, considering the common features of IoT tools, analyzes the advantages and disadvantages of these mechanisms and guidelines, and finally presents future trends.

In their study entitled "Deep Learning Approaches for Anomaly- based Intrusion Detection Systems," Arwa et al. (2020) discussed on the efficiency and effectiveness of the proposed methods through analyzing the solutions and experimental studies and by employing the role of deep learning in detecting the intrusion. Deep-learning-based guidelines and identifiers are recommended by identifying the challenges of past research.

Wenjua et al. (2019) have designed a participatory blockchain signature-based intrusion detection model that can be used as a general framework for signature-based IDS for security sharing and reliable database building.

Research efforts on intrusion detection devices for the IoT have started and accelerated. Considering the provided research backgrounds, it is worth noting that the proposed solutions have not investigated the strengths and weaknesses of each method of diagnosis and strategy in depth. Most authors have focused on a few types of IoT attacks and technologies. Ultimately, very simple accreditation strategies have provided the basis for reproducing other proposed approaches.

3. Problem Definition

In fact, intrusion detection is the process of identifying intruders and attackers into information systems. Known as infiltration, these measures are taken aiming at unauthorized access to computer systems. Intruders may be internal or external users. Internal intruders are in fact network users with varying degrees of access trying to increase the level of access and privileges to exploit unauthorized privileges. External intruders are actually users outside the target network trying to gain unauthorized access to system information.

The intrusion detection system includes sensors, an analytical engine, and a reporting system. The sensors are located in different locations or hosts of the network. Their function is to collect network or host data such as traffic statistics, packet headers, and service requests, besides operating system calls, placed in different locations according to network architecture. Sensors send the collected data to the analytical engine, which is responsible for investigating the collected data and detecting the ongoing infiltration with various signature-based, anomaly-based, feature-based, and combination-based approaches. When the analytical engine detects an intrusion, it will equip the reporting system with infiltration information, including intruder detection, intrusion location, and intrusion time and type, and the system will generate an alert for the network manager [Shen & Huang, 2019].

Classified into three strategies: centralized, distributed, and hybrid, in IoT networks, the intrusion detection systems may be placed in different strategies, in one or more specific hosts, or in any physical thing.

In centralized mode, intrusion detection system's agents are deployed in a centralized component, for example, a border router or a dedicated host. However, due to the need for intrusion detection system's agents to collect many data from smart things, this mode establishes a connection between smart things and the border router. In distributed locating strategy mode, intrusion detection systems are placed on each physical thing, which can obviously decline the above connection while increasing the capacity to consume limited resources of smart things. Nevertheless, unlike the two mentioned modes, infiltration detection system's hybrid agents are deployed in nodes or monitoring nodes, for instance, the guard nodes to take the advantage of centralized and distributed strategies and prevent their weaknesses. This strategy may reduce the requirements for communication between smart things and the boundary router and meet more processing capacity [Shen & Huang, 2019].

Figure 1 shows the independent layers, hardware, and software of the agent, as well as how to deploy and influence intrusion detection systems, indicating that intrusion detection systems in cloud fog-based IoT can be located on a border router in one or more dedicated hosts, or in any physical thing [13].

Today, various measures have been taken to establish security, communications, and information exchange in cyberspace, including data encryption, secure protocol design, and the use of firewalls, tracking systems, and intrusion detection prevention systems. In some network security methods like intrusion tracking systems or firewalls, a decision-making process based on certain data is required to set a specific security policy on the network. Various mathematical tools have been used so far to perform such processes in network security systems and optimize them, such as statistical methods of hypothesis testing, decision theory, pattern identification method, machine learning, graph theory, and control theory.

However, since in many security incidents on the network, the attacker is a human being or a smart program, a method is needed that can decide how a smart attacker can make decisions in order to appropriately change the strategy of his attackers in proportion to the precautionary and model countermeasures. Accordingly, in recent decades, some efforts have been made to apply the game theory to network security.

Since game theory was originally created to model and optimize decision-making in situations where a number of smart factors compete or interact with each other, it is a good tool to be used in many issues related to network security. This theory has been so far used in issues like the optimal allocation of resources, the safe design of network topology, and the optimal configuration of intrusion tracking systems, as well as firewalls.

Given the large volume of data faced by an intrusion detection system, the application of a powerful tool able to enable an intrusion detection system to achieve the desired result by exploring the vast amount of network data is inevitable. The use of game-theory-based systems is one of the powerful tools. Game theory has gained great success in solving the optimization of resources and costs in the economic

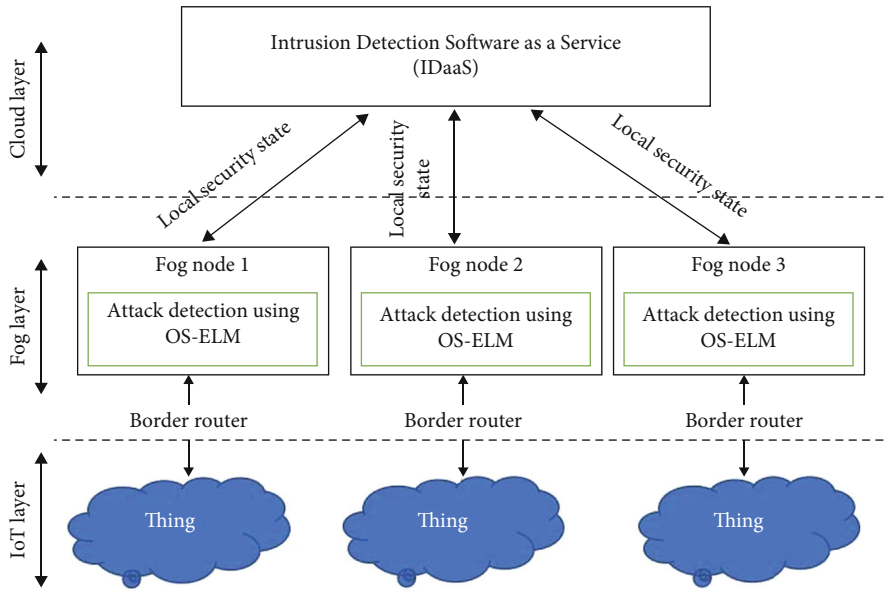


Figure 1: Infrastructures of intrusion detection system for cloud fog-based IoT networks.

field. Accordingly, in recent years, it has been considered by researchers in other fields, too [14].

Game theory is based on the behavior of each player, and it can be based on cooperation or noncooperation in a participatory game [14].

In recent years, the provision of mathematical inferences for wireless networks has become very popular by means of game theory methods. Since game theory is a natural and flexible tool for studying the intelligent and decision-making users, the interaction and cooperation of automated users in wireless networks may be examined with this tool [Pavlidou & Pavlidov, 2010]. Hence, if the issue of security and intrusion detection is investigated from the perspective of game theory, common points between this issue and the models may be gained in this theory.

Detection tools and placement strategy are among important specifications of intrusion detection systems. The studied and analyzed papers point to a general consensus indicating that the game theory and finding the best solution through Nash equilibrium are the most important tools to detect attacks against intrusion detection systems in IoT. However, although the game models proposed to detect IoT intrusion attacks have many similarities, they fundamentally differ from each other in the scope of attack detection. Despite lots of potential attacks against IoT networks, the proposed game model for the intrusion detection system is capable of detecting more attacks simultaneously.

In the proposed model, a mathematical pattern is presented to detect more classes of attacks and correct detection rate and to minimize incorrect detection rate using game theory.

Considering research gap in other studies, in the proposed model, we put emphasis on the lowest amount of error and it can be observed that by considering the dissemination rate parameter and the possibility of the next infraction for a smart object, which is an effective indicator on a smart object behavior, the error and time problem is significantly taken into account and resolved. This way, the smart sensor series detect the attacking smart object *faster and more accurately* and avoid malware dissemination in the IoT network layers.

In the present study, we aimed to model the interactions between attackers and the intrusion detection system as a dynamic two-player game. In game theory, nonparticipatory game is a game in which players may not exchange or negotiate with each other and reach an agreement or form a coalition in any way.

The selection and use of nonparticipatory game are due to the nature of the interactions between the intrusion detection system and the IoT network subsystems. These interactions are indeed a dynamic game with complete information, in which the intrusion detection system is uncertain about the type of player's performance.

4. Information and Data

The main elements in game theory include players, actions, profits, and information, all of which are known as the rules of the game.

The objective in modeling using game theory is to design a situation based on the rules of the game in order to determine what will happen in a specific situation. Game theory is based on the behavior of each player, and players strive to increase their profits in the game and make decisions called strategies [Behounek, 2016]. Accordingly, game theory may be defined as the science of modeling and investigating decision-making systems.

In the current study, dynamic game modeling is defined based on time, completely and strategically according to the information, and the following two conditions have been observed and considered in the proposed model:

(1) Players are fully aware of all the parameters and rules of the game

(2) At least one of the players is unaware of the strategy of the other player; hence, the first player first makes his move, then the second player chooses his move when he is aware of the selected move (operator) of the first player

Defining the players and determining their preferences through the profit function are two of the key elements in describing the game. In the proposed game model, the player is a potential attacker and the other player, the defender of the intrusion detection system.

(i) Players: $\mathcal{N} = \{\text{possible attacker, intrusion detection system}\}$

(ii) First player strategy: $\mathcal{S}_1 = \{\text{attack, no attack}\}$

(iii) Second player strategy: $\mathcal{S}_2 = \{\text{alert by detection, no alert}\}$

Given the provided definitions, we consider the intrusion detection system with the network of sensors $S = \{S_1, S_2, \cdots, S_p\}$, where the sensors are defined as an operating software, reporting the possible attacks in the large subsystem of IoT using a variety of signature-based, anomaly-based, feature-based, and hybrid-based approaches. Alerts reported by the intrusion detection system may be displayed as a set of subsystems, including computer programs or network components, as well as the independent processes distributed across multiple hosts as $A = \{a_1, a_2, \cdots, a_M\}$ which are the target of an attacker. We define the set $T = \{t_1, t_2, \cdots, t_K\}$ as a set of recorded recognizable threats that each member of the set represents a possible intrusion. The properties of one of the T elements can be described by assigning it to one or more classes of the function between $\{F_1, F_2, \cdots\}$ that each class of the function F represents a common property of its members.

In order to be able to detect more than one intrusion by the sensors, by mapping from the S set to the $T \cup \{0\}$ set, the sensor output vector $d = \{d_1, d_2, \cdots, d_L\}$ is defined, so that $L \geq P$. The element i, the output vector associated with the $s_j \epsilon S$ sensor, in the form of $d_1(s_j)$, is equal to one, if the sensor has detected the possible intrusion of $T_k \epsilon T$; otherwise, $d_1(s_j) = 0$.

Given the above argument and since each smart sensor may report a maximum of one of any possible intrusions, we will have

$$d_i(s_k) \neq d_j(s_k), \quad \perp i, j, k > 0 s_k \epsilon S, \tag{1}$$

$$\text{Unless, } d_1(s_k) \neq d_j(s_k). \tag{2}$$

Now, using the definitions and hypotheses of the game, the matrix of the M system is defined by describing the relationship between the output vector of the sensor j and the subsystem i as the matrix (3):

$$M_{i,j} = \begin{cases} 1, & \text{if the sensor } j \text{ alerts for intrusion } i, \\ 0, & \text{if the sensor } j \text{ does not alert for intrusion } i. \end{cases} \tag{3}$$

In Figure 2, the parameters t_1, t_2, and t_3 are as threat targets of subsystems 1, 2, and 3 by the attacker; nt_1 and nt_2 identify the operator of not attacking by the attacker; a_1, a_2, and a_3 warnings show the intrusion detection system alerts for relevant subsystems; and na_1 and na_2 indicate an alert from the intrusion detection system.

The tree modeled in Figure 2, representing an example of the proposed game with two information sets and three subsystems, may be studied by a reversible method. In the first set of information, where the threat t_1 defined by the attacker targets the first subsystem, or does nothing (nt_1), the whole applications of the intrusion detection system are an alert report for the first subsystem with a_1 identifier or not sending an alert with a_2 identifier. Consequently, using the game tree, Figure 2 and definitions may be employed to show the matrix of 2×2 games and how the strategies work in Table 1.

Always $\alpha, \beta \geq 0$.

The parameters Q_{IDS} and Q_{Attack} defined in Table 1 represent the values of the profit function of each player and similar rows and columns like the matrix, performance, strategy spaces of the players, the intrusion detection, and attack system. The $-\alpha_h$ value is the gain of the intrusion detection system for the target detection alert report. On the other hand, α_f and α_m indicate the costs of the detection system for false alarms and attack loss. The cost of β_h shows the penalty for the attacker, and $-\beta_s$ shows the gain of an undetected intrusion.

As a result, strategies of the player's intrusion detection system depend on the relative values of α_f and α_m and false alerts and the cost of losing an attack and threat. If $\alpha_f > \alpha_m$, then the intrusion detection system will not have an alert ($n\alpha$ identifier), and in the other case, if $\alpha_f > \alpha_m$, then the intrusion detection system will always specify an alert (α identifier).

5. Finding the Best Response and Analyzing the Nash of the Game

The study of Nash equilibrium existence in a game has two advantages. First, if we have a game with Nash equilibrium assumptions, we can hope that the attempt to find balance will be successful. The second and the more important is that the existence of equilibrium indicates that the game is compatible with the mode-space solution. Moreover, the equilibrium existence for a family of games allows us to study their properties without finding them explicitly or being faced with the risk of studying an empty collection.

The presence of Nash equilibrium in the Q_{IDS} matrix is investigated. We develop the results by considering strategies similar to those players defined in the form of probability distributions on the space of certain strategies. It is supposed that P_1 and $1 - P_1$ are the probabilities of the t_1 and nt_1 strategies of the attacking player and that q_1 and $1 - q_1$ are the

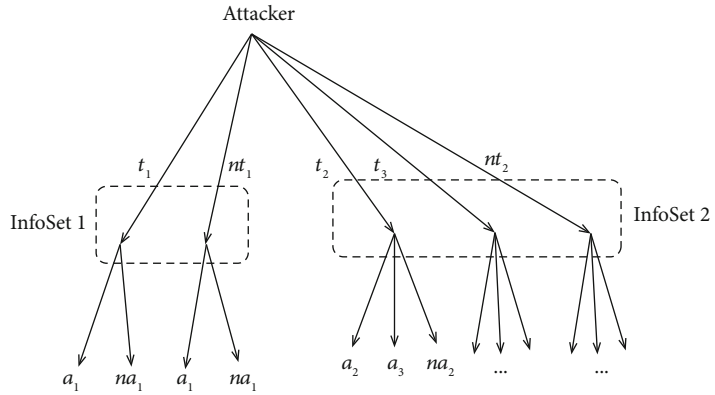


Figure 2: The extended form of the game with 2 information sets and 3 subsystems.

Table 1: Parametric description of similar strategies of the first database.

	t_1	β_h	$-\beta_s$		t_1	$-a_h$	α_m
Q_{Attack}	nt_1	0	0	Q_{IDS}	nt_1	a_f	0
		a_1	na_1			a_1	na_1

probabilities of the strategies a_1 and na_1 of the intrusion detection system. Pair (P^*, q^*) proposes a noncooperative Nash equilibrium solution for 2×2 matrix game operator $(Q_{\text{Attack}}, Q_{\text{IDS}})$ provided that the inequalities (4) and (5) hold true given the fundamental theorem of Nash equilibrium.

$$p_1^*(\beta_h q_1^* - \beta_s(1 - q_1^*)) \le p_1(\beta_h q_1^* - \beta_s(1 - q_1^*)), \quad (4)$$

$$p_1^* a_m + q_1^*\left[a_f - (a_f + a_h + a_m)p_1^*\right] \le p_1^* a_m + q_i\left[a_f - (a_f + a_h + a_m)p^*\right], \quad (5)$$

where $0 \le p_1, q_1 \le 1$. The only solution for the set of inequalities presented as the parameters of the best response is to form a unique Nash equilibrium of the game obtained through

$$p_1^* = \frac{a_f}{a_f + a_h + a_m}, \quad (6)$$

$$p_1^* = \frac{B_s}{B_h + B_s}. \quad (7)$$

In addition, the equilibrium costs of the attacker Q^*_{Attack} and the intrusion detection system Q^*_{IDS} for the designed subsystem matrix of Table 1 are obtained from

$$Q^*_{\text{Attack}} = [p_1^*(1 - p_1^*)]Q_{\text{Attack}}[q_1^*(1 - q_1^*)]^T, \quad (8)$$

$$Q^*_{\text{IDS}} = [p_1^*(1 - p_1^*)]Q_{\text{IDS}}[q_1^*(1 - q_1^*)]^T. \quad (9)$$

Given the Nash equilibrium equations (6) and (7) and the best response parameters of the (8) and (9) equations, the likelihood that the attacker will attack and target the first subsystem at the Nash equilibrium point is reduced by a decrease in a_f since the lower the cost of not reporting an alert to the intrusion detection system, the more likely it is to set an alert and trap the attacker. Then, of course, increasing a_n and a_m plays a key role for the attacker, and $-\beta_s$ the likelihood that the intrusion detection system will detect an alert is affected by the attacker's gain from successful intrusion.

The parametric analysis for the second set of information is examined by establishing a relationship between costs in subsystems two and three and in the form of a 2×2 matrix in Table 2.

In Table 2 α_d and $-\beta_d$ are the deception costs for the intrusion detection system and attack. It can be assumed that $\alpha_d > \alpha_m$ and $\beta_d > -\beta_s$ since the lack of alert of the intrusion detection system is much more costly than the lack of attack, and the attacker disrupts the security mechanisms by deceiving the intrusion detection system. Let us assume that $\bar{p}_1$, $\bar{p}_2$, and $1 - \bar{p}_1 - \bar{p}_2$ are the probabilities of t_2, t_3, and nt_2 strategies of the attacker, and assume that $\bar{q}_1$, $\bar{q}_2$, and $1 - \bar{q}_1 - \bar{q}_2$ are the probabilities of the a_1, a_2, and na_2 strategies. The intrusion detection systems' operating strategy is presented with relative values such as

$$\bar{p}_1^* = \bar{p}_2^* = \frac{a_f}{2a_f + 2a_m + a_{h-}a_d}, \quad (10)$$

$$\bar{q}_1^* = \bar{q}_2^* = \frac{\beta_f}{2\beta_s + \beta_h - \beta_d}, \quad (11)$$

if $\beta_d < \beta_{h\beta_d<\beta_h}$ and $a_d < 2a_f + 2a_m + a_{ha_d<2a_f+2a_m+a_h}$. Finally, the Nash equilibrium strategy of the intrusion detection system may be presented in the form of

$$\begin{cases} \dfrac{a_1 \text{ with probability } q_1^*}{na_1 \text{ with probability } q_2^*}, \text{ InfoSet 1,} \\ a_2 \text{ with probability } \bar{q}_1^* \\ a_3 \text{ with probability } \bar{q}_2^* \quad \text{, InfoSet 2.} \\ na_1 \text{ with probability } 1 - \bar{q}_1^* - \bar{q}_2^* \end{cases} \quad (12)$$

Table 2: Parametric description of similar strategies of the second information set.

t_2	β_h	β_d	β_s
Q_{Attack}			
t_3	$-\beta_d$	$-\beta_h$	$-\beta_s$
nt_2	0	0	0
	a_2	a_3	na_2
t_2	$-\alpha_h$	α_d	α_m
Q_{IDS}			
t_3	α_d	$-\alpha_h$	α_m
nt_2	α_f	α_f	0
	a_2	a_3	na_2

Always $\alpha, \beta \geq 0$.

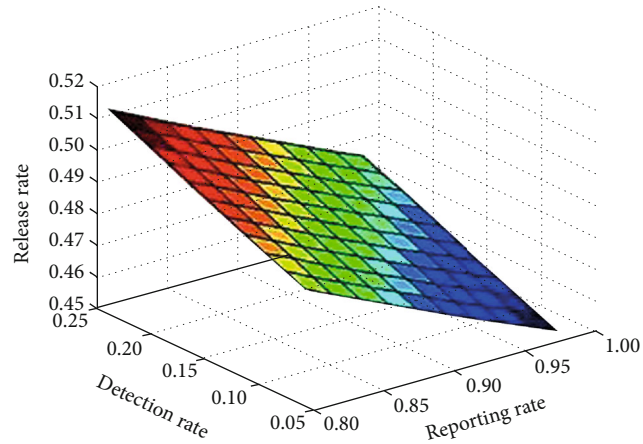


Figure 3: Release rate of an aggressive smart object based on the parameters of detection rate and reporting rate.

6. Evaluation and Validation of the Proposed Game Model

Today's intrusion detection system's architecture is a passive information-processing model.

Nevertheless, with the abundance and complexity of security attacks, intrusion detection systems cannot distinguish between the real intentions and target of the attackers. To correctly identify and detect the target of an attack, intrusion detection systems must be able to process the attack information in the text. Through establishing a network of sensors in the system and by a theoretical analysis of the game's sensor output data, the attacker's behavior, intention, and target may be modeled. In addition, due to the flexibility of the proposed game model, not only attacks targeting the specific parts of the network but also single targets such as processes distributed across multiple physical subsystems may be detected. Besides modeling the attacker's behavior and intention, the game's theoretical framework may be employed in order to analyze and model the response process of the intrusion detection system through calculating the relationship between security succession and statistical points. The response and reaction of the intrusion detection system vary from a simple alert setting to a high-cost reconfiguration of the system, including shutting down relatively less important services in the system.

In this section, the theoretical framework of the proposed game is first validated by performing numerical experiments in MATLAB software environment and augmentation, and to investigate and explain the Nash equilibrium of the numerical samples, in mixed and behavioral strategies, the attacker's vector with application of t_1, t_2, t_3, nt_1nt_2 and the intrusion detection system's vector with application of $[a_1, a_2, a_3, na_1na_2]$ were related; and the Nash equilibrium was calculated according to equations (6), (7), (10), and (11). Then, by entering the proposed game model into the IoT using cloud-fog-based IDSaaS, a potential application is presented.

As the intrusion detection system and the potential attacker interact and play in several different strategies in the proposed game model, the game results are observed and recorded at each stage. We present and calculate some statistical points from these results.

The optimal smart thing rate criteria as an attacker have been considered by choosing release and the possibility of subsequent infection. The reason for choosing this criterion may determine the effective parameters on the behavior of a smart thing in the network, as well as the principles of timely judgment about whether the attacker's smart thing is infected or not.

The parameters of various game strategies have been specifically evaluated in software experimentation, although if the values of these parameters are logically changed, similar trends towards statistical points can be reached. Thus, given the parameters of different strategies, it is believed that the following numerical results are helpful for showing the characteristics of the proposed game model and they can be easily reproduced for more specific situations.

The parameters used to evaluate the proposed method in this research are time, correct detection rate, reporting rate, and emission rate of the infected smart object. The faster intelligent sensors can detect and report an attacking smart object, the faster the propagation rate converges over time t , which prevents malware (attacker) from spreading across layers of the Internet of Things.

Obviously, a higher detection rate and a higher reporting rate (alert) allow IDSaaS to more easily trap an attacking smart object, which in turn, as shown in Figure 3, causes the malware in attacking smart object makes less effort to propagate, which reduces the propagation rate.

In addition, lower reporting rates mean that attacker detection rates are reduced and the privacy of IoT networks cannot be adequately protected for research purposes, so it can be concluded that an attacker is a smart object. Release at a higher rate means that the intrusion detection system is less likely to detect that attacker. As expected, the actual implementation trends in Figure 3 confirm the analysis presented.

However, different factors have different impacts on the players in the proposed game model, affecting the rate of different detection strategies and the release rate.

Table 3: Comparison of the proposed model with the other three models.

	Security threats coverage	Emphasized detection method	Type of game model	Complexity of protocol and architecture	Scalability
The proposed model	Noncooperative game	Combined	Common attacks in IoT networks and wireless sensors	Low	Yes
Susan and Rayford (2019)	Noncooperative game	Signature-based	Threats in mobile ad hoc networks	Moderate	Yes
Wenjua et al. (2019)	Cooperative game	Signature-based	Dissemination of malware, to protect privacy in IoT networks	High	Yes
Shen et al. [13]	Cooperative game	Combined	Dissemination of malware, to protect privacy in IoT networks	Moderate	Yes

Table 3 includes a comparison of the proposed model with the three models in other articles.

7. Conclusions

In the present study, a strategic, dynamic, and complete game model has been defined to detect the intrusion of attacks in IoT networks in the distributed intrusion detection system. An analytical research of the game in the form of 2×2 matrix subgames and finding the best response parameters in Nash equilibrium bring valuable insights for the attacker and the behavior of intrusion detection. Furthermore, the simple assumptions proposed to achieve analytical results may be easily expanded to achieve more realistic scenarios, and smart intrusion detection system, defined as a software agent, reports attacks on the large subsystems of IoT using a variety of signature-based, anomaly-based, feature-based, and hybrid approaches.

Thus, it can be stated that given the equilibrium solutions and costs of each subgame in the presented matrices, the intrusion and attack detection systems specify the performance of their strategies. Furthermore, compared to a related work, the distinguishing feature and the used innovation are the presentation of a game model to detect attacks on the IoT between sensor's nodes and the platform server used to detect more attacks, correct detection rates, and minimize wrong detection rate.

Consequently, it is important to note that other common security measures, as well as the implementation of privacy, cannot be directly applied to IoT technologies. Therefore, the development of specific security solutions such as intrusion detection systems is essential to allow users and organizations to identify and repair all weaknesses and attacks in their system. Further, this method has been used in smart systems efficiently in the future in real-time applications.

Data Availability

Data are available on request through contacting safieh .siadat@gmail.com.

Conflicts of Interest

The author declare that they have no conflicts of interest.

References

[1] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[2] E. Borgia, "The Internet of things vision: key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.

[3] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of things: the road ahead," *Computer Networks*, vol. 76, pp. 146–164, 2015.

[4] D. Jin, "Application of IOT in electronic commerce," *Journal of Digital Content Technology and its Application*, vol. 6, 2012.

[5] S. Kumar and K. Dutta, "Intrusion detection in mobile ad hoc networks: techniques, systems, and future challenges," *Security and Communication Networks*, vol. 9, no. 14, pp. 2484–2556, 2016.

[6] A. Abduvaliyev, A. S. K. Pathan, Jianying Zhou, R. Roman, and W.-C. Wong, "On the vital areas of intrusion detection systems in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1223–1237, 2013.

[7] C. Modi, D. Patel, B. Borisaniya, H. Patel, A. Patel, and M. Rajarajan, "A survey of intrusion detection techniques in cloud," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 42–57, 2013.

[8] R. Mitchell and I.-R. Chen, "A survey of intrusion detection techniques for cyberphysical systems," *ACM Computing Surveys*, vol. 46, no. 4, pp. 1–29, 2014.

[9] D. Midi, A. Rullo, A. Mudgerikar, E. Bertino, and Kalis, "Kalis — A system for knowledge-driven adaptable intrusion detection for the Internet of things," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pp. 656–666, Atlanta, GA, USA, June 2017.

[10] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 266–282, 2014.

[11] J. Granjal, E. Monteiro, and J. Sa Silva, "Security for the internet of things: a survey of existing protocols and open research issues," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1294–1312, 2015.

[12] B. G. Atli and A. Jung, "Online feature ranking for intrusion detection systems," 2018, http://arxiv.org/abs/1803.00530.

[13] S. Shen, L. Huang, H. Zhou, S. Yu, E. Fan, and Q. Cao, "Multistage signaling game-based optimal detection strategies for suppressing malware diffusion in fog-cloud-based IoT networks," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1043–1054, 2018.

[14] T. Ramesh and S. Shaleni Priya, "A review on game theory based congestion control in wireless sensor network," *Journal of Network Communications and Emerging Technologies*, vol. 8, no. 4, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8860646, 29 pages
https://doi.org/10.1155/2020/8860646

Research Article

IMOC: Optimization Technique for Drone-Assisted VANET (DAV) Based on Moth Flame Optimization

Rehan Tariq, Zeshan Iqbal, and Farhan Aadil

Department of Computer Science, University of Engineering and Technology Taxila, Pakistan

Correspondence should be addressed to Rehan Tariq; rehan.tariq@students.uettaxila.edu.pk

Received 1 June 2020; Revised 3 September 2020; Accepted 29 September 2020; Published 7 November 2020

Academic Editor: Sungchang Lee

Technology advancement in the field of vehicular ad hoc networks (VANETs) improves smart transportation along with its many other applications. Routing in VANETs is difficult as compared to mobile ad hoc networks (MANETs); topological constraints such as high mobility, node density, and frequent path failure make the VANET routing more challenging. To scale complex routing problems, where static and dynamic routings do not work well, AI-based clustering techniques are introduced. Evolutionary algorithm-based clustering techniques are used to solve such routing problems; moth flame optimization is one of them. In this work, an intelligent moth flame optimization-based clustering (IMOC) for a drone-assisted vehicular network is proposed. This technique is used to provide maximum coverage for the vehicular node with minimum cluster heads (CHs) required for routing. Delivering optimal route by providing end-to-end connectivity with minimum overhead is the core issue addressed in this article. Node density, grid size, and transmission ranges are the performance metrics used for comparative analysis. These parameters were varied during simulations for each algorithm, and the results were recorded. A comparison was done with state-of-the-art clustering algorithms for routing such as Ant Colony Optimization (ACO), Comprehensive Learning Particle Swarm Optimization (CLPSO), and Gray Wolf Optimization (GWO). Experimental outcomes for IMOC consistently outperformed the state-of-the-art techniques for each scenario. A framework is also proposed with the support of a commercial Unmanned Aerial Vehicle (UAV) to improve routing by minimizing path creation overhead in VANETs. UAV support for clustering improved end-to-end connectivity by keeping the routing cost constant for intercluster communication in the same grid.

1. Introduction

Vehicular ad hoc networks (VANETs) are different from mobile ad hoc networks (MANETs); therefore, clustering algorithms designed for MANETs cannot be applied to VANETs. In traditional VANETs, infrastructure, like road-side units (RSUs), is used to provide network services to vehicular nodes, selecting the optimal paths and transmitting data. This infrastructure provides road safety information, road congestion, alternative routes, along with weather conditions to drivers. In urban areas where RSU support is available, VANETs work efficiently, but in those areas where infrastructure is not available, VANETs do not perform well [1]. On the other hand, scalability is one of the challenges in VANETs. Clustering is used to solve the scalability issue, but in the high-speed environment on highways where the vehicle speed is relatively much faster than in urban areas, the clustering does not work well, resulting in degraded network performance due to the higher rate of reclustering [2]. Existing VANET routing and clustering algorithms are computationally expansive, so we need to build a heterogeneous routing algorithm (for flying ad hoc network- (FANET-) assisted VANET) with low routing overhead, efficient utilization of computational resources, and high overall network throughput [3]. The addition of UAVs in existing VANETs is a challenging task because they have very distinct features as compared with ground nodes/vehicle. Another challenge is the efficient utilization of flight time of UAVs because UAVs carry limited energy resources [4]. In VANET, partial infrastructure support is available through RSUs; replacing the RSUs with UAVs to form a fully ad hoc network is another challenge to be addressed.

The current traffic system has many problems like road congestion, accident risks, mobility, node energy, node

Table 1: VANET routing challenge.

Topology-based routing	Geography-based routing
Performance at stake in rural areas	Performance on stake in urban areas
Transmission can be delayed	Transmission of data for longer distances
Higher routing overhead	Incorrect GPS coordinates for a node
Higher packet drop ration	Inherent loops can occur
Routes are broken more frequently	Network partitioning more frequently

Table 2: UAV classification.

UAV type	Weight (kg)	Altitude (m)	Hovering time (hrs)	Range (km)
Micro	<5	250	1	<10
Mini	150	150-300	<2	<10
Close range	150	3000	2-4	10-30
Short range	200	3000	3-6	30-70
Medium range	1250	5000	6-10	70-200
Medium-range endurance	1250	8000	10-18	>500
Low-altitude deep penetration	350	50-9000	0.5-1	>250
Low-altitude-long-endurance	<30	3000	>24	>500
Medium-altitude-long-endurance	1500	14000	24-48	>700

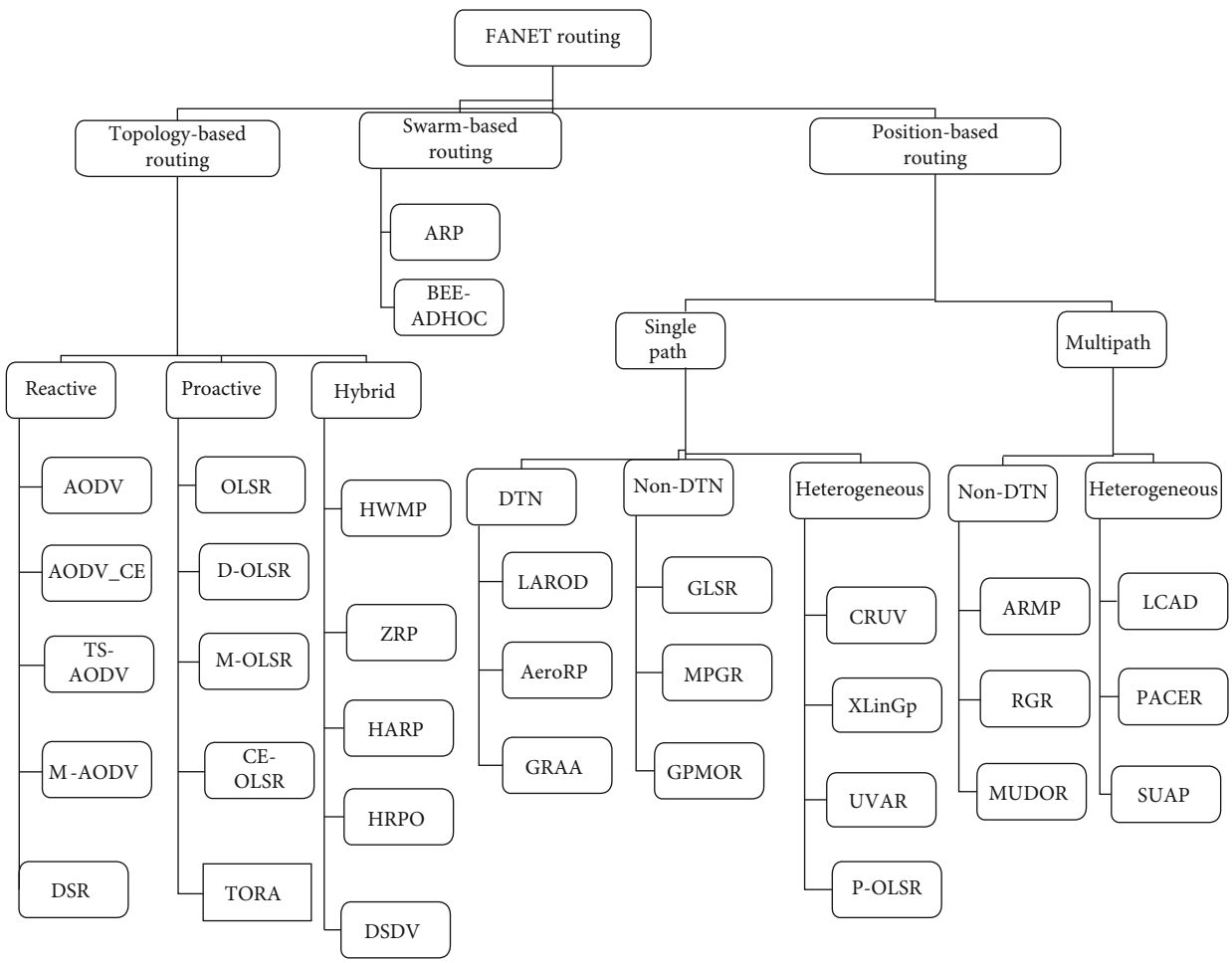


Figure 1: FANET routing classification.

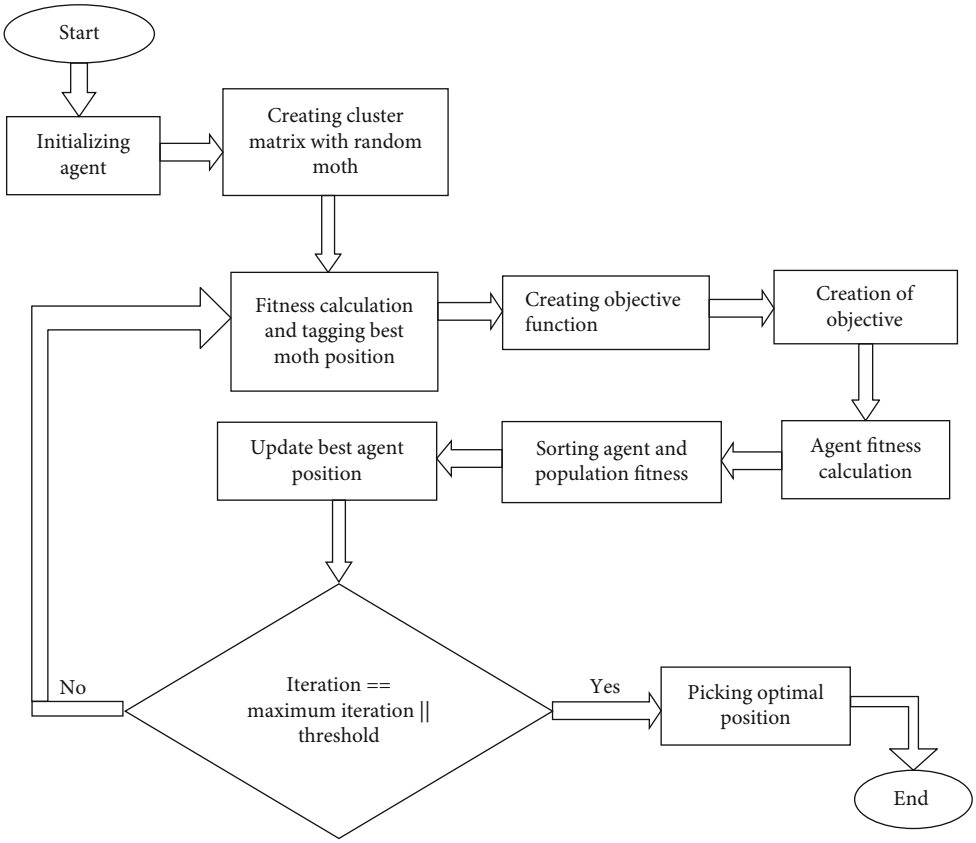


Figure 2: Flow chart of proposed IMOC algorithm.

physical condition, and received signal strength of node [5]. If nodes are participating in path construction and path maintenance phase, then the route will be considered as reliable [6]. The reliable route improves packet delivery ratio, reliability, and packet delays and achieves low overhead during transmission of data. Route reliability is essential and robust for application such as disaster management and audio and video conferencing. If the route is lost, then the packet takes a lot of time to reach a destination with higher travelling cost. So, to solve these issues, FANET assistance will provide a better solution to solve irregularities in traditional VANETs.

Genetic algorithms/programming, evolutionary strategies, and learning classifier systems are some types evolutionary algorithms [7, 8]. Evolutionary algorithms offer a decent solution for the problems that cannot be solved with other techniques. In situations where we must find a solution for unsolvable problems, evolutionary techniques are widely accepted. EA might be computationally expensive, but finding a near-optimal solution for unsolvable problems is acceptable. In FANETs and VANETs for a continuous node clustering problem, the choice of evolutionary algorithms is effective [9].

The natural evolution model of biological evolution is the base for evolutionary algorithms [10]. An environment will be generated in which possible solutions will be evolved to find a solution for the problem. For problem factors with regard to constructed surroundings, it is possible to get the best possible solution through evolution. To solve the scalability issue, nodes are grouped and they share the same geographical coordinates [11] [12]. To provide solutions for network scalability, clustering is one of the methods [13]. The clustering solution ensures the effective utilization of resources with load balancing in each cluster. A moth flame optimizer is one of the finest clustering techniques to provide an optimal number of clusters. Moths are the insects like butterflies. About 16000 species of moths are identified to date. Like other insects, moth larvae convert into cocoons in adulthood. The moths navigate at night-time and follow moonlight. The traverse orientation method is used for traveling by moths. During traveling, moths follow moonlight by keeping a fixed angle toward the moon. Their going after moonlight with a fixed angle keeps them in a straight line. Humans adopted the same method for traveling in a straight line at night [14, 15]. For example, at night, if a man wants to walk toward the west, the moon position must be on the northern side of the sky. By keeping the moon on the right side, a man can easily travel in straight line. With regard to the efficacy of transverse alignment, often, moths are tricked by nonnatural light and are inclined to fly spirally towards nonnatural light. If the light source is far away, then the same behavior for transverse orientation performs well.

Once artificial light comes across a straight path that is being followed, moths try to keep the angle toward the

```
1. START
2. Define grid size
3. In the 2d grid random deployment of vehicular nodes
4. Broadcasted position of each node in the search space
5. Node IDs as vertex, mesh topology is formed
6. Assignment of values to edges in the mesh network by distance calculation of each node
7. Create search space of m × n order, initialized moth position
8.        When i = 1, loop from i the total maximum number of search agents. FOR j to the total number of dimensions where j starts
from 1
9.        Position of each moth updated by moth position (i, j) = upperbound minus lowerbound both starts from I divided by node
position in the grid plus lower-bound
10.        End loop
11.     If simulation stalled or ended (20 iterations)
12.          FOR moth i to flame size and i starts from 1
13.            Calculation of fitness of moth_position (MP) as moth fitness = fitness_function (); fitness moth_position (MP)
calculated
14.               WHILE node list NOT empty for clustering nodes
15.               Allocation of the best solution to moth with cluster fitness of each moth less than the finest result
16.               END_WHILE
17.          END_FOR
18.     Sorted_fitness (), sorting fitness values of all moths
19.     Population_sorted () w.r.t sorted_fitness () population is sorted
20.     Among the updated fame position, the best obtained till now
21.     Best flame_score = sorted_fitness (1);
22.     Best flame_position = population_sorted (1, :)
23.     Update moth position-based on the corresponding flame
24.          FOR i from one to maximum search_agents
25.               FOR j from one to total dimensions
26.               Compute distance for ith-moth for jth-flame; equivalent to absolute (population_sorted (i, j) − MP (i, j));
27.               Moth location update
28.               END_FOR
29.          END_FOR
30.     IF convergence_curve = convergence_curve iteration no 131.          stall iteration++;
32.        ELSE
33.        stall_iteration =0;
34.     END_IF
35.     Iteration++;
36.     END loop
37.     Best solution from search is equal to total number of
38.  END
```

Algorithm 1: Intelligent moth flame clustering optimization for VANETs.

artificial light source. Deadly paths for moths occur when the artificial light source is too close to moths, because moths must converge toward the light. This convergence property of moths can be exploited mathematically as a moth flame optimizer (MFO) algorithm [16]. In this research, we proposed an intelligent moth flame clustering optimization for VANET (IMOC) to optimize the clustering problem in VANETs with air assistance of FANETs.

1.1. Vehicular Ad Hoc Network. The moving vehicles are equipped with advanced communication capabilities to form a wireless network referred to as VANETs. VANETs offer intelligent transportation services including road conditions, traffic density, alternative routes, vehicle conditions, nearby rest areas, and weather updates to drivers. Intelligent transportation integrated information systems, communication sensors, advanced mathematical methods, and high technologies to traditional transportation infrastructure. Traffimatics is the term used in intelligent transportation system (ITS) where moving vehicles act as network nodes for transceiving and routing packets in a network [17]. To ensure a safe and secure route for vehicles is the main application of ITS. Information including unseen traffic, road conditions, weather information, traffic density, and infotainment is broadcast to make the trip safer for drivers and passengers. To provide short wireless networks between vehicles, radio devices and onboard units (OBUs) are installed on the vehicle. These devices are used to provide communication between OBU and RSU to form VANETs [18].

To get an accurate geographical position of moving nodes, vehicles are equipped with a global positioning system (GPS) and a differential global positioning system (D-GPS). RSUs serve as a cellular base transceiver system and act as a backbone to provide communication between vehicles in

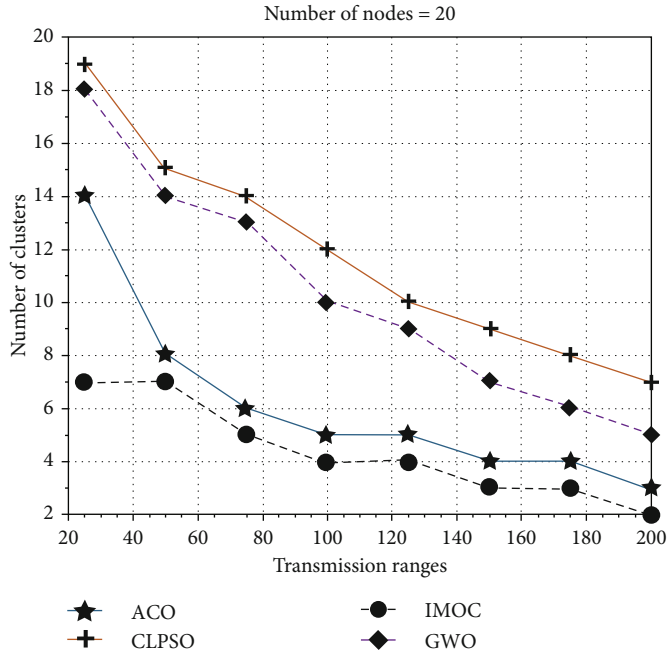
Table 3: Parameters for simulation.

Parameters	IMOC	CLPSO	GWO	ACO
Total population size	100	100	100	100
Maximum number iteration	150	150	150	150
Total runs	10	10	10	10
Weight for inertia	0.90	0.694	0.694	—
Rate of evaporation	—	—	—	0.5
C1	2	2	2	2
C2	2	2	2	2
Grid size for simulation	500 m^2, 1000 m^2, 1500 m^2, 2000 m^2	500 m^2, 1000 m^2, 1500 m^2, 2000 m^2	500 m^2, 1000 m^2, 1500 m^2, 2000 m^2	500 m^2, 1000 m^2, 1500 m^2, 2000 m^2
Number of vehicles	10 to 100	10 to 100	10 to 100	10 to 100
Interval between vehicles	+20	+20	+20	+20
Transmission ranges	25 m to 200 m	25 m to 200 m	25 m to 200 m	25 m to 200 m
Vehicle position	Fixed	Fixed	Fixed	Fixed
Minimum_distance between vehicles	1.5 m	1.5 m	1.5 m	1.5 m
Maximum_distance between vehicles	5 m	5 m	5 m	5 m
W1 (1st objective function's weight)	0.5	0.5	0.5	0.5
W2 (2nd objective function's weight)	0.5	0.5	0.5	0.5

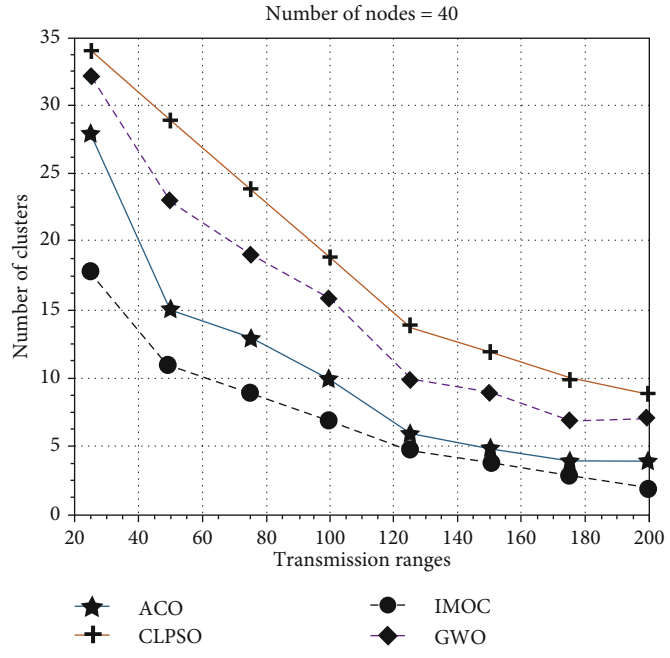
VANETs [19]. The mobility model is not random as vehicles follow the road trajectory, but speed is relatively high as compared with MANETs. The energy is not a critical issue because transceivers utilize engine power to establish communication in VANETs. The number of RSUs in VANETs depends on the communication protocol. The communication in VANETs might be intervehicle, vehicle to RSU, and routing-based communication [20]. The information needs to be broadcast efficiently in VANETs for effective information interchange during communication between nodes. To provide such capability, there is a need to have efficient routing protocols. The proactive (table-driven) and reactive (on-demand) are two main classifications for routing protocols.

In proactive routing protocols, the routing information is available every time in its packet header. Optimized link-state routing (OLSR) and fishy state routing (FSR) are types of proactive protocols. FSR minimizes the overhead because it does not broadcast; it only exchanges topological change with its neighboring nodes [21]. OLSR uses multipoint relays to the optimized broadcasting process of the control message to keep the routing table updated. Hello and topology control message are used to discover and disseminate link-state information. Nodes share topological change to their neighboring subset as nodes have limited repetitions for broadcasting [22]. In reactive protocols, the route is not stored permanently which helps in minimizing communication overhead and routes are established on-demand. In the route construction process, the control message is broadcasted through flooding to look up participating nodes for communication. Ad hoc on-demand distance vector (AODV), dynamic source routing (DSR), and temporally ordered routing algorithm (TORA) are examples of reactive routing. AODV apply route discovery by hop count and sequence number. Destination sequenced number is checked in the route construction process based on the route request/route reply messages [23]. DSR routing information is attached to the data packet header from the source. Route recovery or maintenance is the limitation of DSR [24]. The temporally ordered routing protocol (TORA) is based on a three-level route construction, route maintenance, and erasing route by using query (QRY), update (UPD), and clear (CLR) messages. Topological change does not have an effect on routing information until a complete path from source to destination has been lost [25]. To provide fast message data delivery along curved roads overlays the node selection based on optimal position and exponent partition range [26].

Geographic or position-based routing protocols used GPS to pick exact coordinates of nodes and used their current location for routing data [27]. GPSR, geographic source routing (GSR), and greedy perimeter coordinator routing (GPCR) are some of the examples of position-based routing protocols. A greedy perimeter stateless routing protocol works on greedy approval for transmission of data between the sender and receiver. The locations of the transmitting node and the destination are used to find other nodes to construct a route [28]. In GSR, the shortest path is calculated between the sender and receiver based on their locations [29]. GPCR coordinating nodes are given preference with noncoordinating nodes. Communication is established between the sender and the receiver on their geographical locations and road conditions [30]. Anchor-based street traffic-aware routing (A-STAR) is the best route established

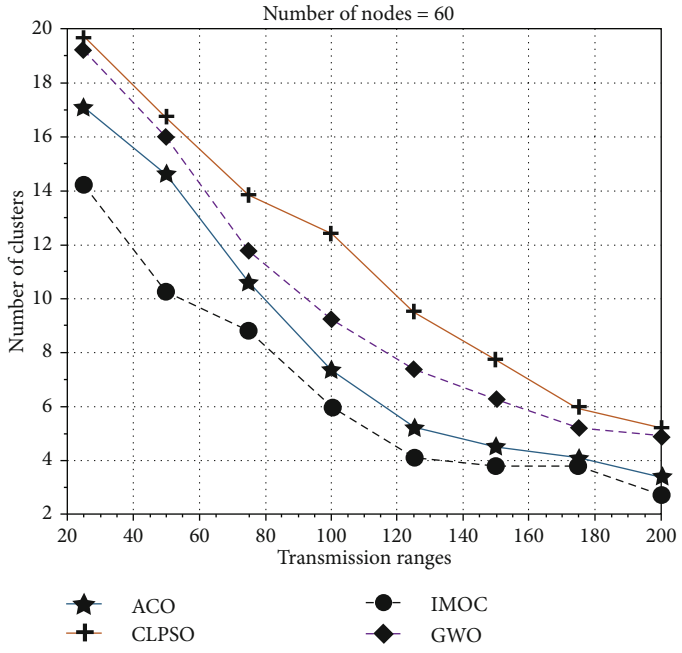


(a)

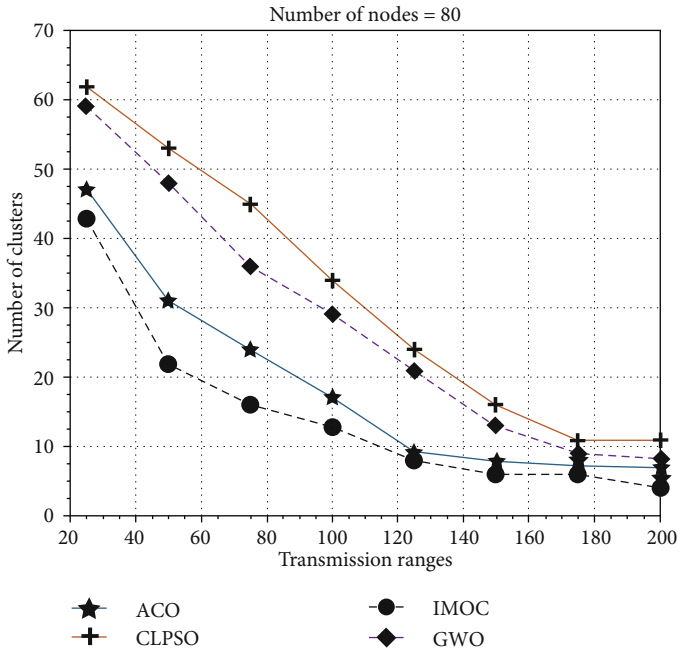


(b)

Figure 3: Continued.

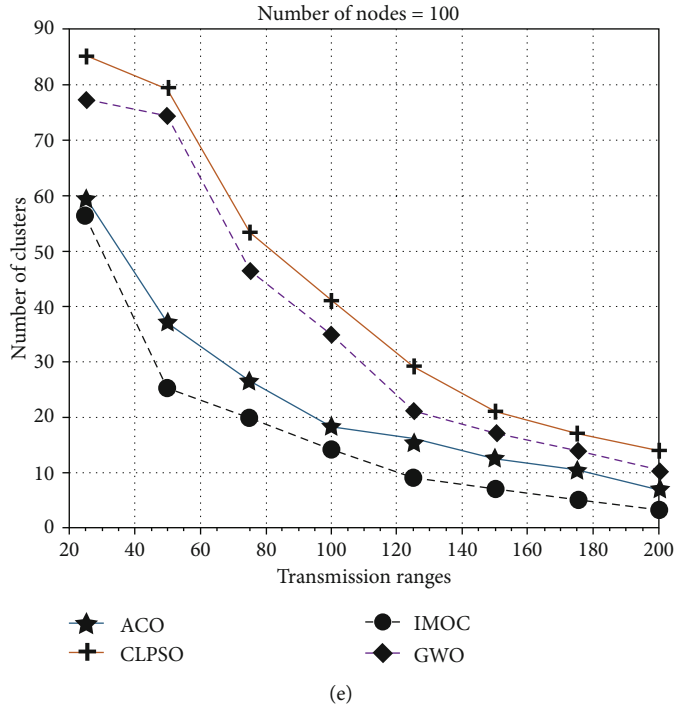


(c)



(d)

Figure 3: Continued.



(e)

Figure 3: Grid_size 500 m × 500 m, nodes 10 to 100.

on information gathered from nodes including their location and trust. The anchor path is computed with Dijkstra's least weight path [31]. Table 1 presents the routing challenges for topology-based and geographical-based routing protocols in VANET.

1.2. Flying Ad Hoc Network. Availability of low-cost Wi-Fi radio interfaces, GPS, micro-embedded systems, high-resolution cameras, and sensors raised a path for developing intelligent flying vehicles or UAVs [32]. These UAVs created a relatively new era of networks known as FANETs. To integrate drones with an existing vehicular network to improve overall network performance is known as a drone-assisted vehicular network (DAVN) [33]. The UAV-assisted applications have their unique features, competitive advantages, and characteristics [34]. The fundamental operation in Internet of Things (IoT) application is data aggregation. It can be seen in distributed internet-based industrial computing and control systems [35]. The FANET applications can be found everywhere from civilian to military use [36]. Such applications are traffic monitoring, disaster monitoring, providing coordination between rescue teams, crop monitoring, fire monitoring where human access is difficult, infotainment, autonomous tracking, and border surveillance [37, 38]. Two types of applications for FANET can be classified on the deployment of an aerial node in topology: one is a single aerial node application and the second is multiaerial node applications. In the single aerial node application, only one aerial node (AN) is deployed in the middle of base stations localized on the ground; the AN serves as a router between multiple base stations, whereas in multi-UAV application, a team of ANs works together to provide services [39].

Table 2 shows the classification for UAV type, coverage range, weight, climb rate, and endurance time in the air. FANETs are considered as a subclass of MANETs; UAV routing becomes more complex as AN characteristics vary from other ad hoc networks. The characteristics, including mobility rate, number of ANs, transmission range, weather conditions, and residual energy, need to be critically addressed for designing a routing scheme. Under these limitations, higher communication failures can result in high dynamic movements of ANs. In FANETs, effective routing will support to keep services and applications stable and available all the time. The FANET is an additional support to enhance the effectiveness of existing technologies such as VANETs and MANETs [40]. The ANs can be placed and dispatched in multiple scenarios to improve VANET and MANET applications to provide end-to-end connectivity between ground nodes.

Figure 1 presents the classification for FANET routing protocols. The classification for topological-based protocols can be divided as proactive, reactive, and hybrid. The position-based routing protocols are classified as single path schemes and multipath schemes. The swarm-based protocols are listed for FANET routing. The parameters for FANET-routing protocols are node density, link information between ANs, residual energy, coverage area, and mobility pattern.

The concept of FANET-assisted VANETs is the focus of researchers these days. Various approaches were proposed where drones/UAVs were used to assist VANETs. One of the initial approaches is a multi-UAV-aided network [41]. This approach proposed two-layer networking, i.e., aerial networking and ground networking where the former is responsible for air-to-air communication and the latter is a

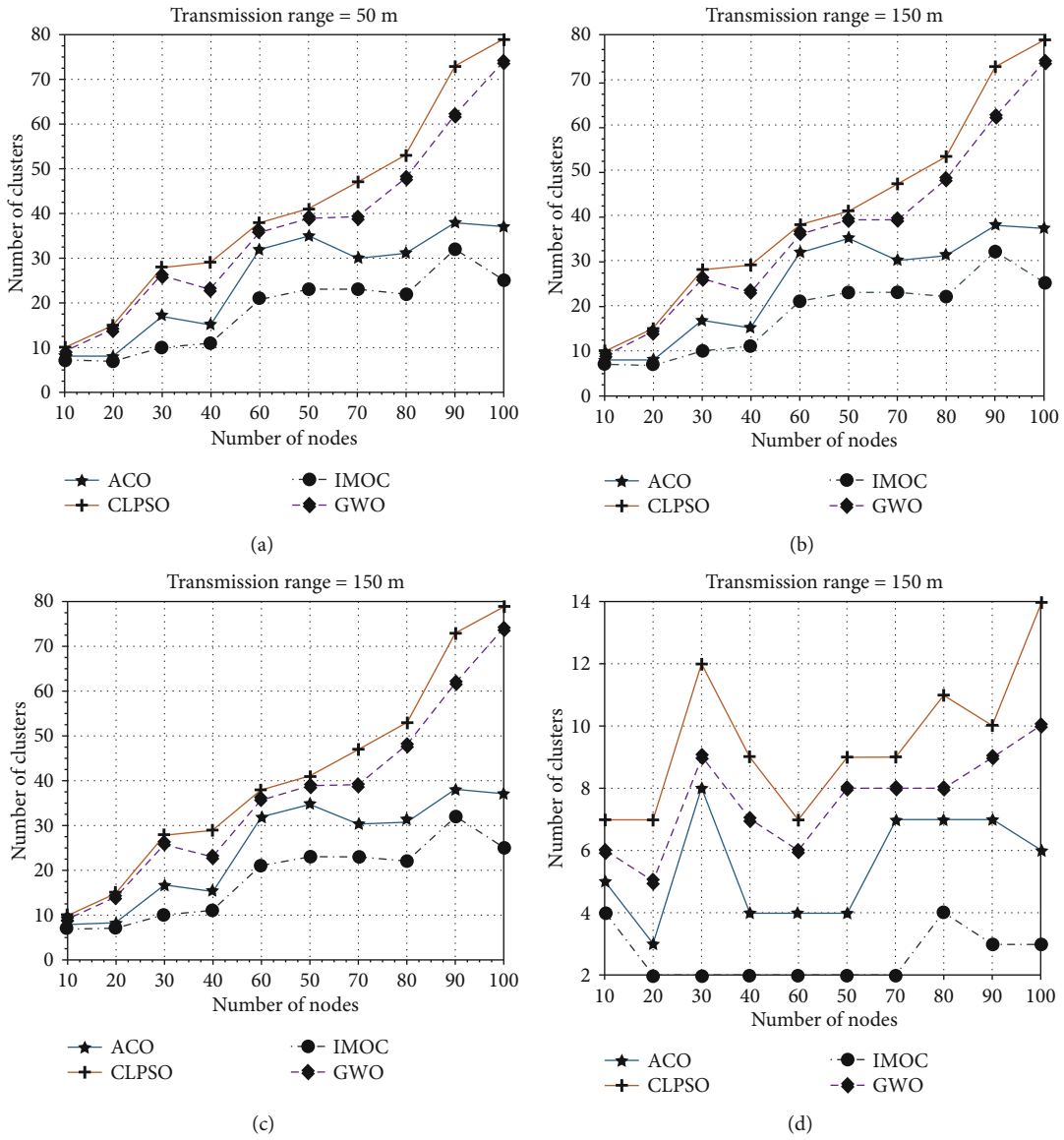
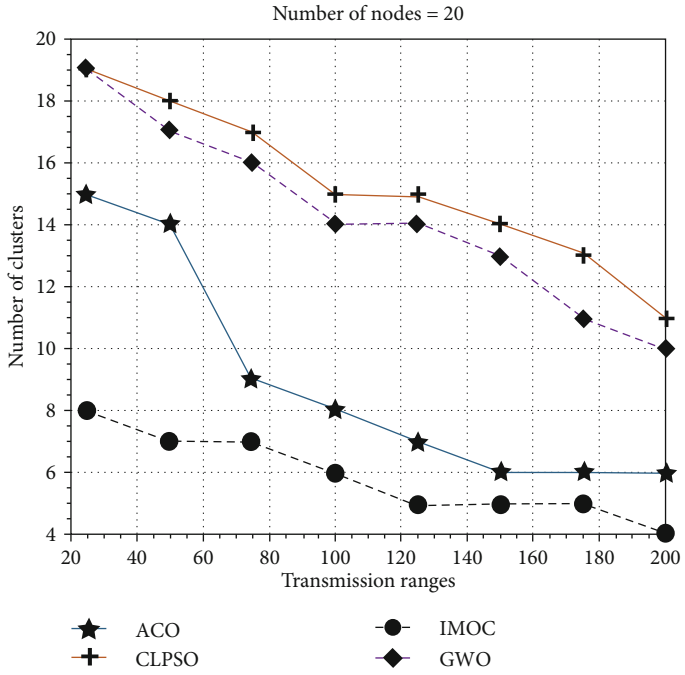


Figure 4: Grid_size 500 m × 500 m, transmission_ranges 50 to 200.

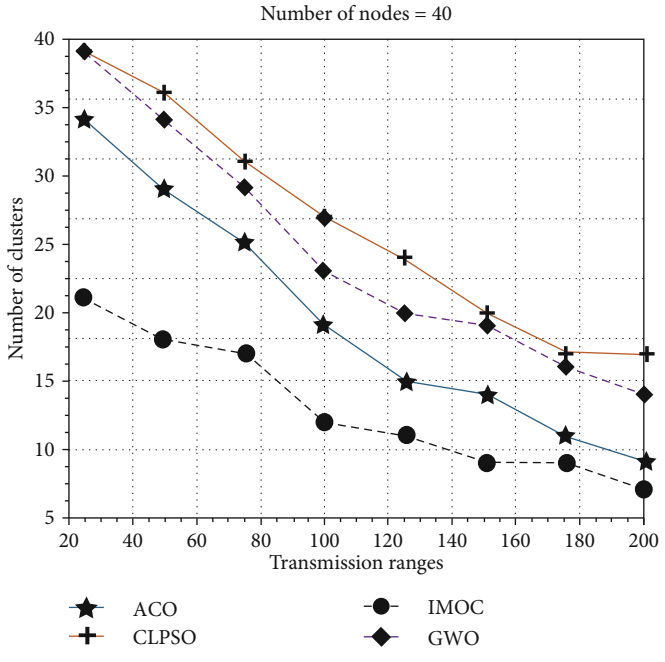
VANET which transfers the data among vehicles. A special channel is established between these two channels for transferring information such as road conditions. A UAV-assisted VANET routing protocol (UVAR) is a delay tolerant protocol [3] in which UAVs are used which have global knowledge of the network. UAR has two subcomponents: UVAR-G is responsible for transferring packets among connected vehicles by considering traffic density, whereas UVAR-S works by forwarding packets to the UAV. UVAR-S is an on-demand routing protocol that considers multipath toward UAVs and selects the most connected one as the preferred path.

To get reliable data delivery and guarantee robust paths, the flooding-based techniques are used in providing efficient routing solutions. The existing UAVs cooperate in an ad hoc fashion with vehicles [42]. The U2RV routing protocol is proposed in [43]. U2RV is a four-phase process. In the first phase, various paths are discovered; the paths can include any path established through the UAV; based on source and destination, a suitable path is selected from the set of paths discovered in the first phase. This is followed by the actual data delivery in the third phase. The final phase deals with the discovery of an alternative path which is necessary as the routes in VANETs are dynamic. UAVs are proposed to be used in VANETs [44] for help in finding the disconnected segments and work as relay nodes in VANET infrastructure. A central ground station is at the heart of the scheme which sends instructions to UAVs for storing and forwards the data and dispatches it towards disconnected segments. Disconnected segments are identified through the exchange of hello messages between vehicles and ground stations.

UADD, a protocol for smart transportation networks, is proposed in [45]. To provide communication between UAVs

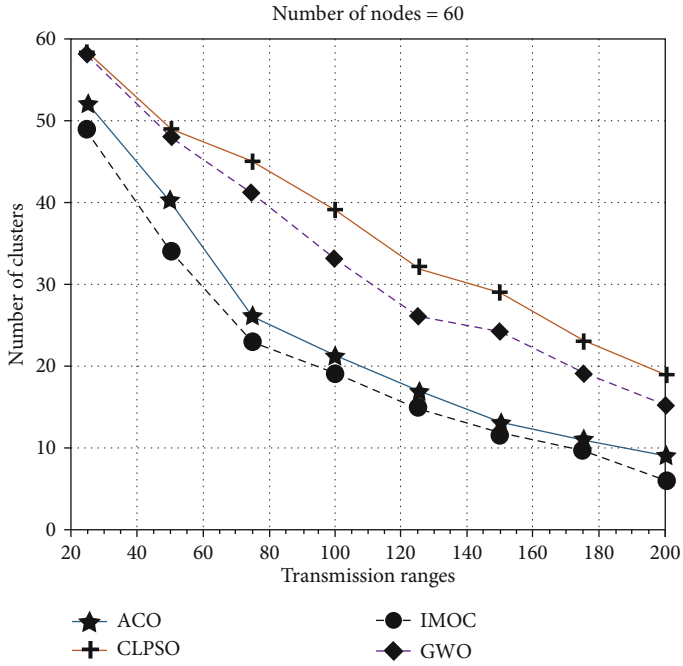


(a)

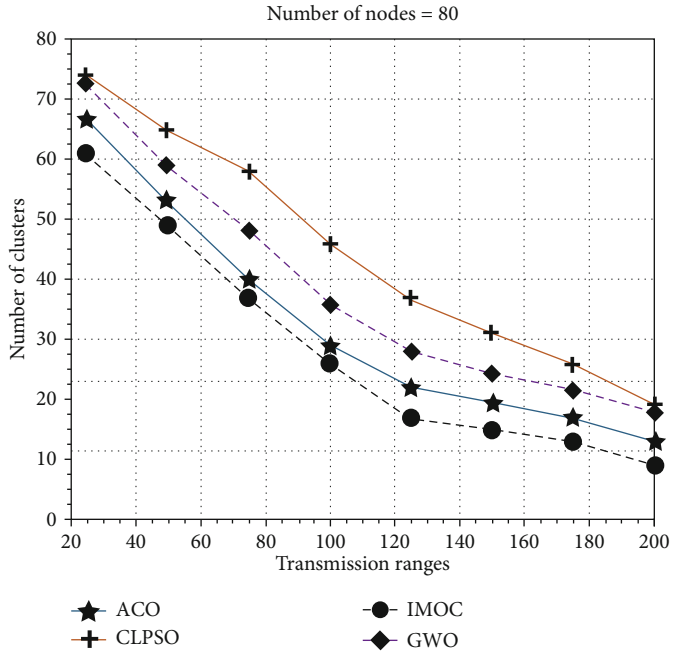


(b)

Figure 5: Continued.

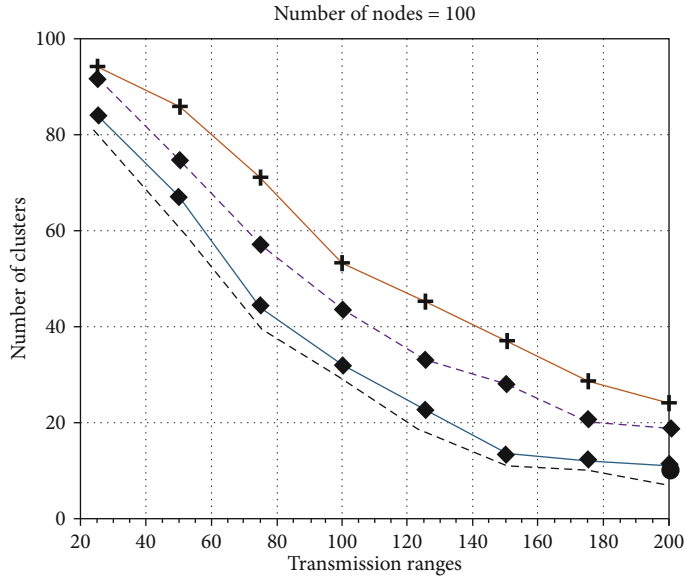


(c)



(d)

Figure 5: Continued.



(e)

Figure 5: Grid_size 1000 m × 1000 m, nodes 10 to 100.

and vehicles, an opportunistic virtual interaction scheme was introduced. The crux of this research work is forwarding the data on the optimal intersection to the UAV when the vehicle-to-vehicle communication is not possible. Like the previous scheme, the UAV is used to store and forward packets to the other vehicle. The use of UAV mitigates the effects of jamming in VANETs [46]. It has been observed that hackers observe the traffic pattern between OBUs and vehicles, so they launch a jamming attack. Authors have proposed to use UAVs to shield against this threat. A technique based on reinforcement learning was used to achieve an optimal relay policy adopted by UAVs to avoid the aforementioned attack.

One of the fundamental aspects of any ad hoc network is its mobility model. In [47], the authors proposed a mobility model for UAV- and VANET-based communication. In the proposed model, UAVs follow the movements of vehicles on the road. To maintain the connectivity, the received signal strength (RSSI) from the vehicle is used. The UAV selects the vehicle with the lowest RSSI value and tries to improve the RSSI, so that packets can be delivered successfully to the said vehicle.

To improve communication performance of VANETs between ongoing OBUs and UAVs against smart jammers, the UAVs are used to induce a specific strategy according to the jammer attack [48]. To enhance the network life time and mitigate the "hot spot" problem, a new algorithm is proposed, an asynchronous clustering and mobile data gathering based on timer mechanism (ACMDGTM) [49]. In the curved road scenario which overlays the node selection method, adaptive relay-node selection (ARNS) is used to redefine the optimal position of the node while considering obstacle distribution. The broadcasting characteristics of ARNS are used to classify the road structure [50].

2. IMOC-Proposed Methodology

The flow of the proposed IMOC algorithm is shown in Figure 2. During the initialization phase in solution space ($m \times n$), the random position assigned to each moth and moth array is equipped with fitness values. For flames, a similar ordered matrix and array are generated. The best value for the moth found so far is stored in the flame matrix. It is an iterative process, so the optimal number of flames in search space with the best moth against its flame is attained during each iteration, After each best find, it updates the moth position. A moth travels in the solution space until they have found an optimal solution or the searching operation is terminated.

In order of the $m \times n$ solution space, the random position assigned to each moth during the initialization phase and moth array is stored as fitness values. Similarly, the flame matrix and corresponding array are generated. The moth's best value found so far is stored in the flame matrix. To find an optimal solution or terminate the search operation, moths are moved in a solution space.

This operation used the dimension of lowerbound-upperbound of the search space. Further, it is used to evaluate the fitness value of each moth based on their location in the search space. The creation of a fitness matrix is an iterative process; updated values are stored in the matrix in ascending order. For each moth, the lower fitness value is provided by the fitness matrix. The optimal best score for the flame is calculated by combining the position of the moth and its fitness value and is used to update the moth position in the search space. For optimal solution, a linear decreasing factor "x" was used for convergence. For effective communication, the minimal number of clusters required is also obtained by using the same convergence technique.

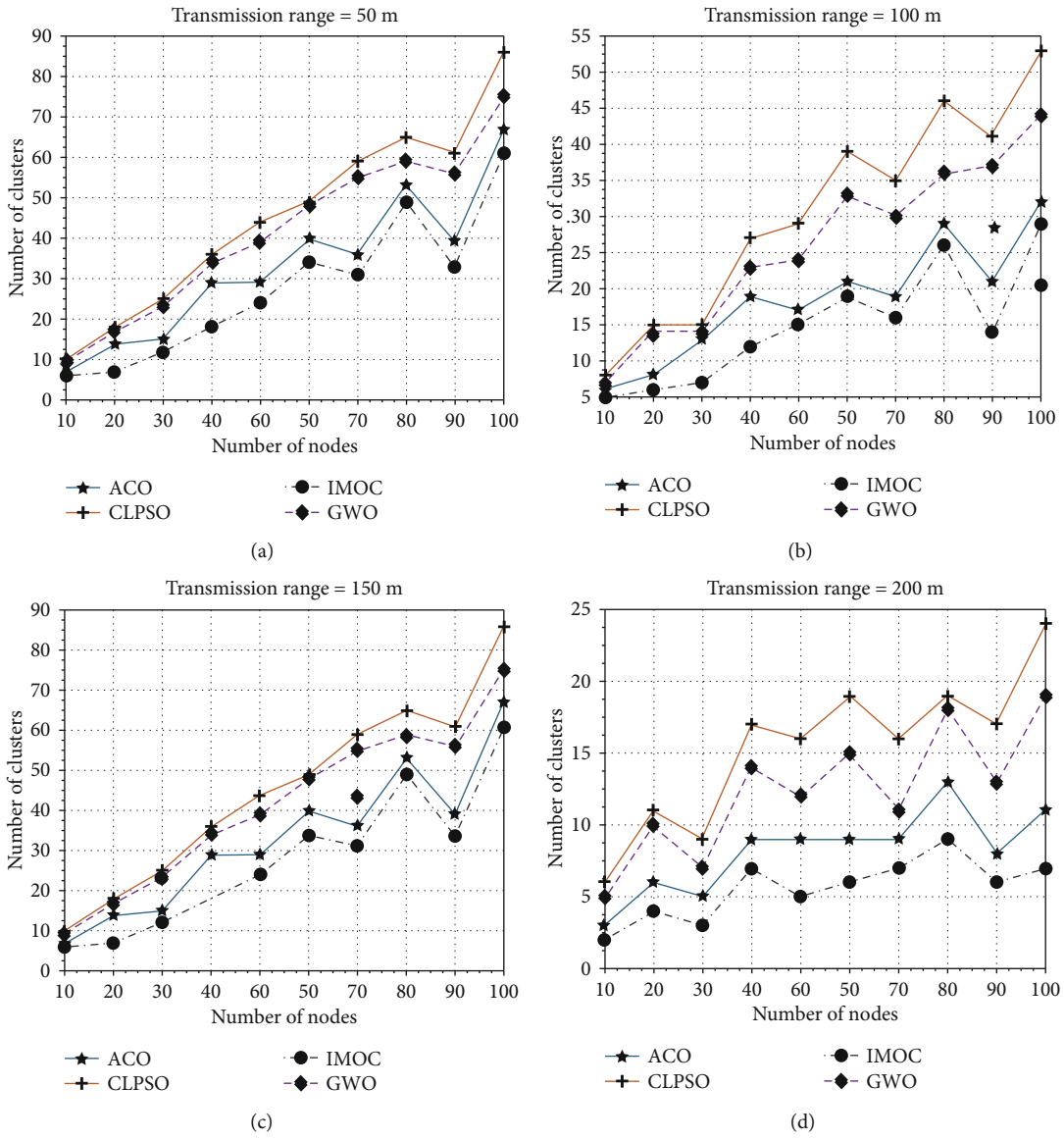


Figure 6: Grid_size 1000 m × 1000 m, transmission_ranges 50 to 200.

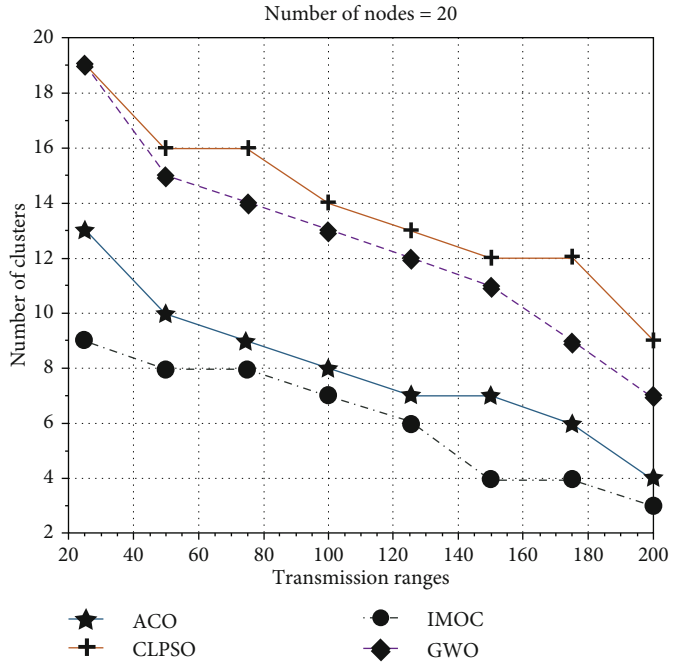
After creating clusters, selection of the cluster head (CH) is the next phase. Multiple parameters like grid size, node density, node connectivity, load balance factor, and transmission range are the parameters used in the CH selection process. These parameters are passed to fitness function with assigned weights. An important part of IMOC is to carry selection using a fitness function. Cluster lifetime increased by selecting the best CH resulting in minimizing network energy and limiting unnecessary broadcast overhead. The following equation (1) is used to calculate the fitness value for the IMOC algorithm:

$$\text{Fitness} = \frac{\text{W1} \times \text{Energy\_Resi}}{(\text{W2} \times \text{avg\_dis})(\text{W3} \times \text{delta\_diff})}. \quad (1)$$

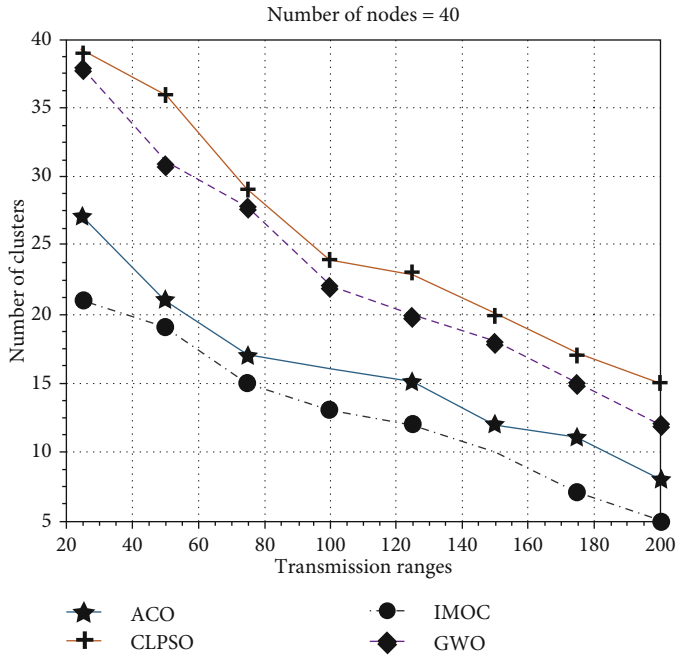
The residual energy of the vehicular node is denoted by Energy_Resi, the average distance between neighboring nodes is avg_dis, and the load balancing factor (LBF) is considered by delta_diff. Weight for energy is W1, the average distance is W2, and the delta difference is W3. Achieving an equal number of cluster members only results in an ideal scenario. In a real scenario, it is difficult to achieve as vehicular nodes change their positions and other parameters. The ideal node degree deviation of movement from its neighbors is computed by

$$\text{Delta\_Diff} = \text{ABS}(\text{Ideal\_Degree} - \text{Node Degree}). \quad (2)$$

Inappropriate selection of CH might result if selection criteria for CH are static and a single parameter might bias the fitness function [21–23, 38]. Depending on the scenario, weight is assigned to parameters dynamically by IMOC to counter the biasing problem and negatively impact the fitness values. In the first step, each value for the parameter normalized between the range of 0 and

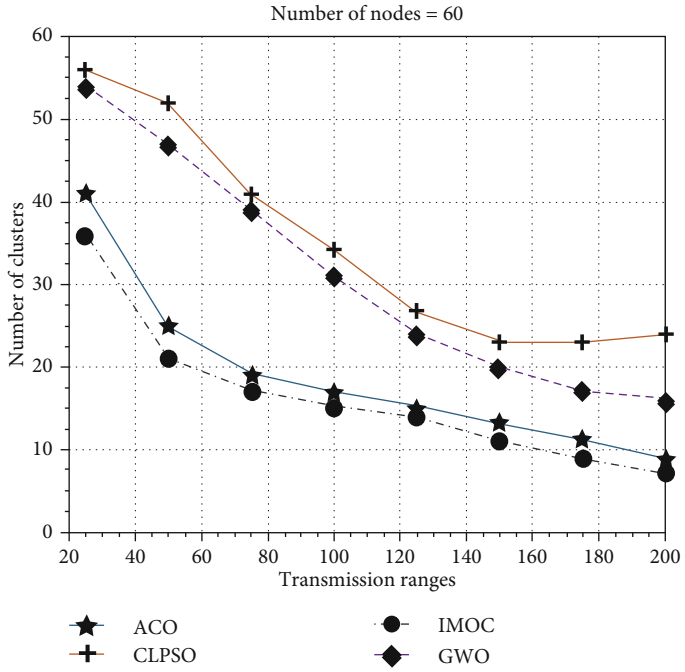


(a)

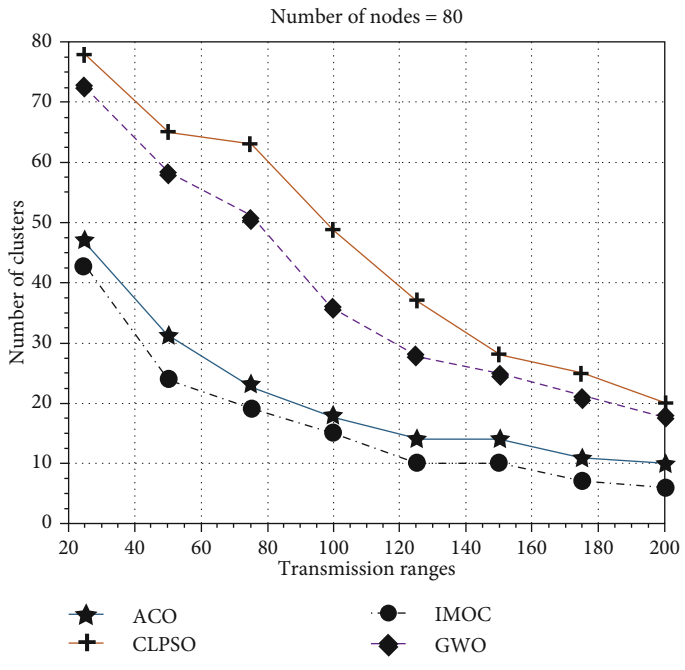


(b)

Figure 7: Continued.



(c)



(d)

Figure 7: Continued.

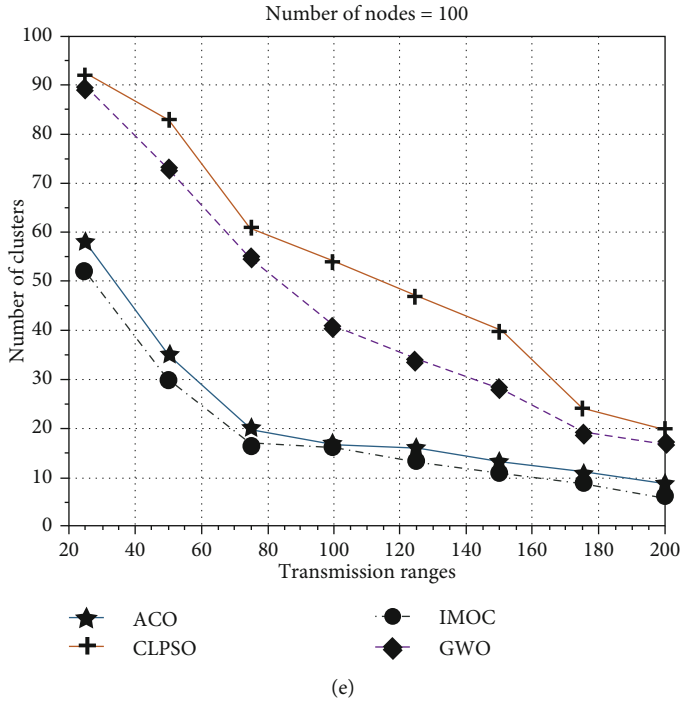


Figure 7: Grid_size 1500 m × 1500 m, nodes 10 to 100.

10. In equation (3), each parameter deviation is calculated based on negative impact.

$$\text{Dev\_}(p) = \text{ABS}(\text{\_mean} - \text{\_parameter}(p)). \quad (3)$$

Penalized outlier parameters are used in equation (3), to add penalty on weirdness from their mean, and are used to compute updated values for parameters. To penalize the outlier penalty, another equation is used with

$$w(p) = \frac{1}{\text{dev}(p)}. \quad (4)$$

The aggregated total of all weight essentially is equivalent to "1." Fitness for each node can be calculated for all parameters by equation (1).

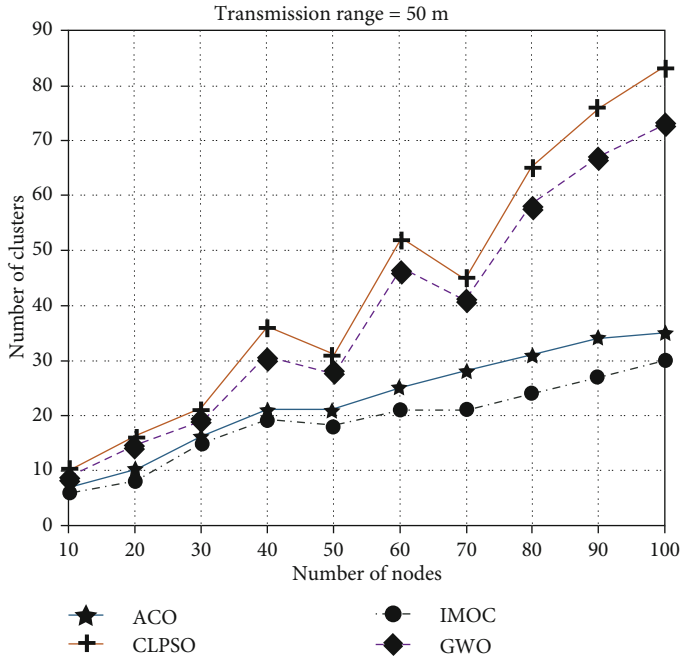
3. Experimental Results and Analysis

Table 3 shows the parameter setting for simulation; the total population for each algorithm set is to be 100 and 150, and the maximum iteration for each solution is set to ten, depending on the nature of the algorithms, inertia, weight, and evaporation rate used. Inertia is set as CLPSO 0.694, GWO 0.694, and IMOC 0.90, and evaporation rate is set as 0.5 for ACO. Four grid sizes are used to perform simulations 500 m × 500 m, 1000 m × 1000 m, 1500 m × 1500 m, and 2000 m × 2000 m. The vehicle maintains a minimum distance of 1.5 m in the simulations; the interval between vehicles is set to be 2 m. Transmission ranges for all simulations are considered from 25 m to 200 m and node density from 10 to 100. An assumption is considered wherein vehicle mobility remains fixed or is moved with constant velocity. IMOC experimentations were compared with ACO, GWO, and CLPSO which are some the state-of-the-art evolutionary clustering protocols.

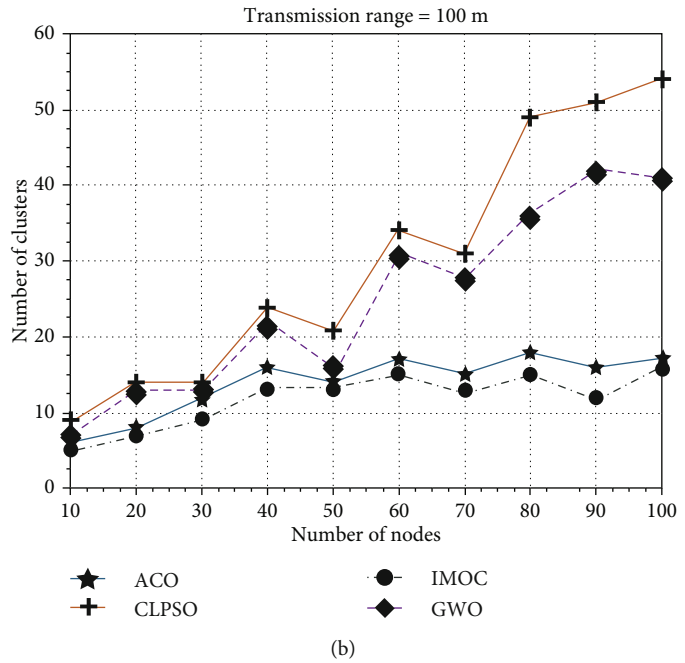
4. Results and Discussion

To measure the efficiency of the proposed IMOC algorithm with CLPSO, ACO, and GWO, numerous experimentations were performed. Their performance is presented in the following figures. To check the efficacy of IMOC, node density in the grid and transmission ranges for nodes were evaluated in multiple scenarios. IMOC maintained its supremacy and flexibility in the results. In Figure 3, transmission ranges for nodes were set from 25 m to 200 m keeping the grid size to 500 m × 500 m, and 10 to 100 vehicular nodes were deployed. In Figure 3(b) where the node density is 40 and the transmission size is 25 m, clusters created by CLPSO = 29, GWO = 27, and ACO = 23, but IMOC created only 19. When the transmission range was increased to 100 m, IMOC created only six clusters, in comparison to ACO 14, GWO 18, and CLPSO 21. The proposed technique showed consistent performance measure compared with the existing techniques. When the transmission range for the nodes was increased, few numbers of clusters were created. IMOC results outperformed other algorithms when nodes increased from 20, 40, 60, 80, to 100. IMOC created an optimal number of clusters for all sets of experimentation in the 500 m × 500 m grid size.

Figure 4 shows the optimal performance of IMOC in comparison with existing techniques. To strengthen this matter, an additional set of experimentation is performed,

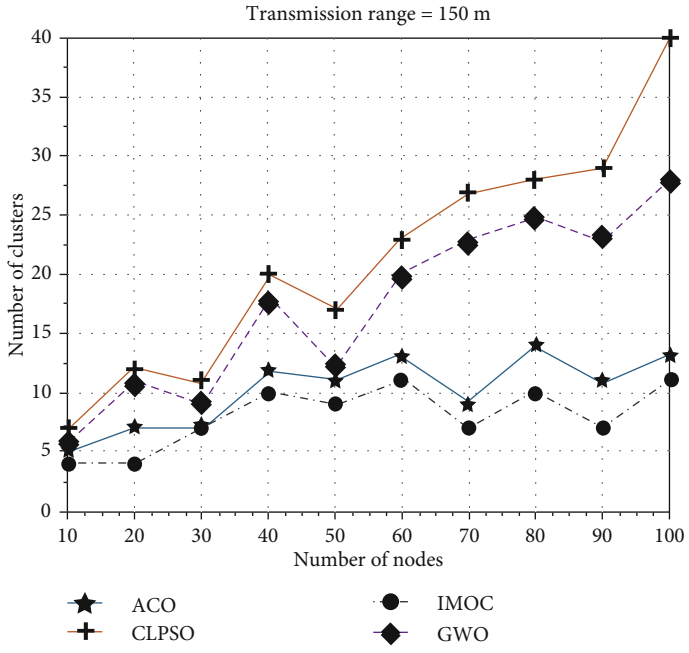


(a)

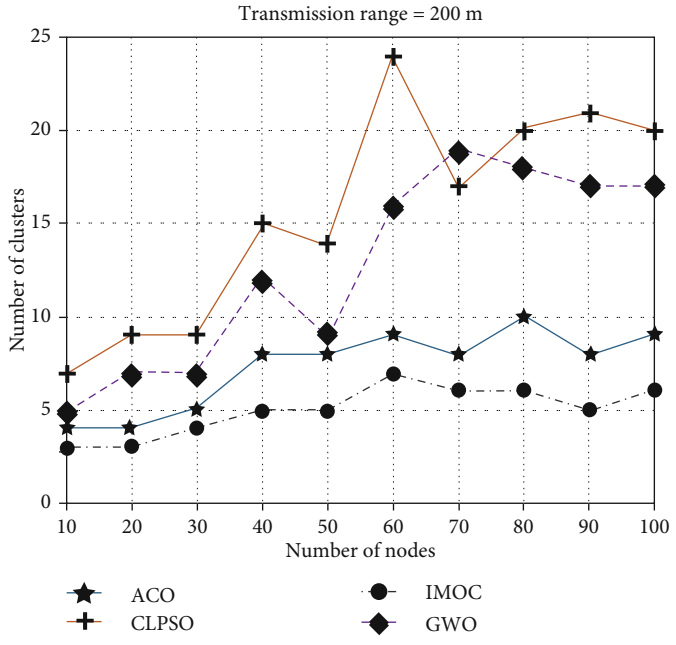


(b)

Figure 8: Continued.



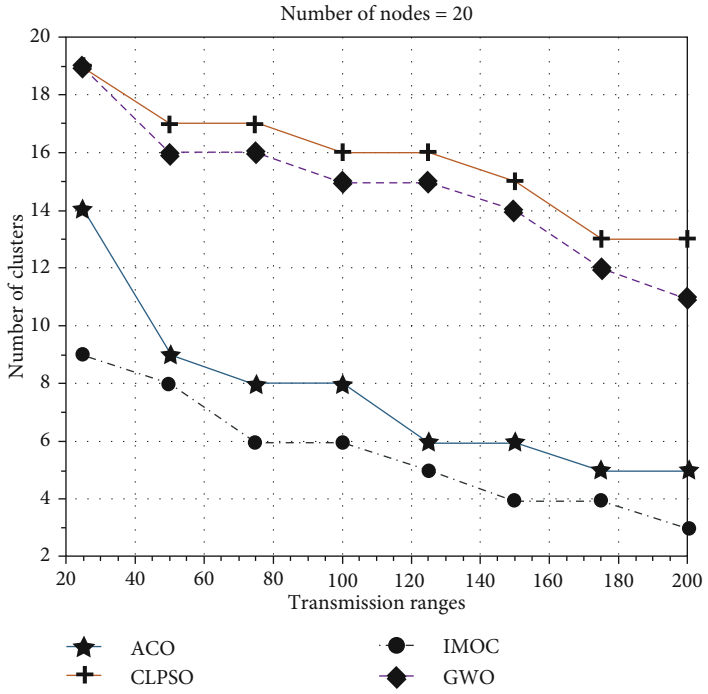
(c)



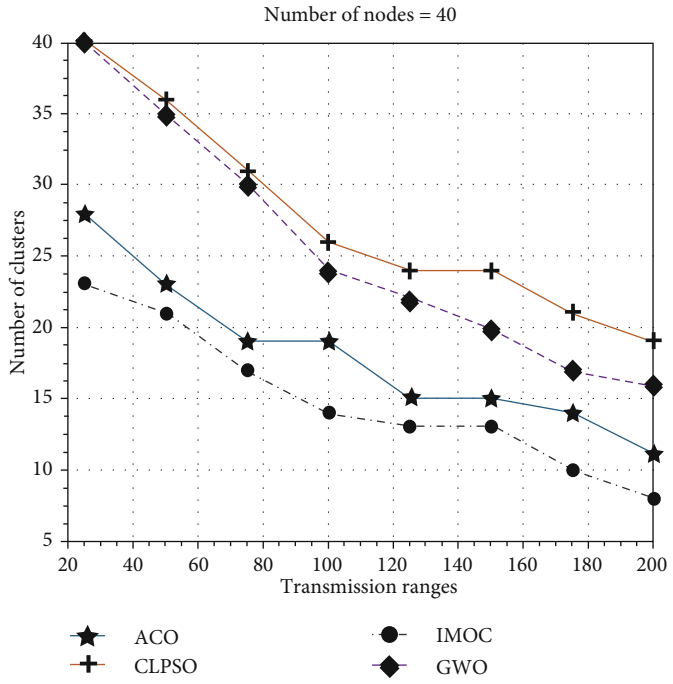
(d)

Figure 8: Grid_size 1500 m × 1500 m, transmission_ranges 50 to 200.

where nodes were fixed to 50 m and the cluster creation for grid size 500 m × 500 m was checked by varying node density from 10 to 100 nodes (Figure 4(a)). In these experiments, the tested transmission ranges were 50 m, 100 m, 150 m, and 200 m for node density from 10 to 100 nodes in Figure 4. With varying transmission ranges, it is noticed that IMOC performance was equally good and consistent; results for all instances produced were better, and ACO results were the closest to IMOC results in these settings. It was noticed that IMOC performance was also the best even if the grid size increased to 1 km × 1 km. Figure 5 depicts results for this scenario. The experimental setting was updated, and results for node density 50 for the 75 m transmission range show that IMOC created 18 clusters in comparison with 31, 36, and 23, respectively, for GWO, CLPSO, and ACO in Figure 5(a). The minimum number of clusters was created by IMOC. In Figure 5(e), to check the results, the accuracy for node density was increased to 100 and the transmission range to 125 m; GWO created 33 clusters, CLPSO 45, and ACO 23. In comparison, IMOC created only 18 clusters.

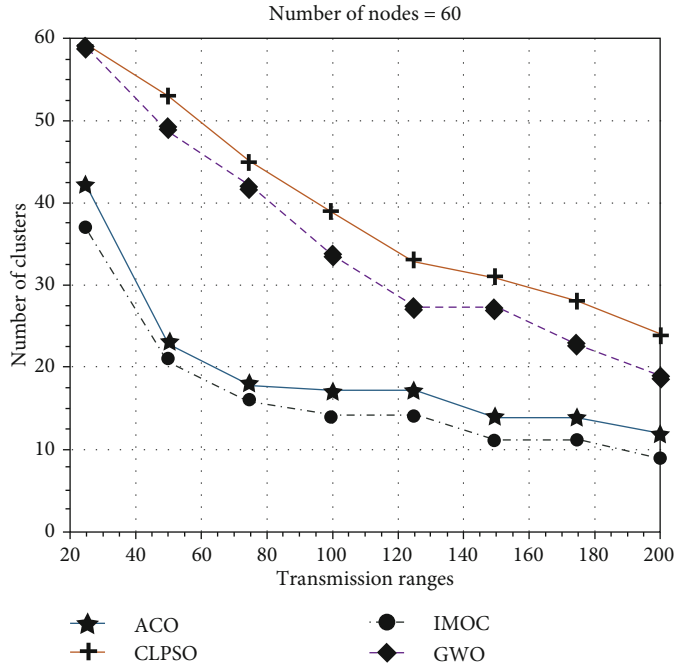


(a)

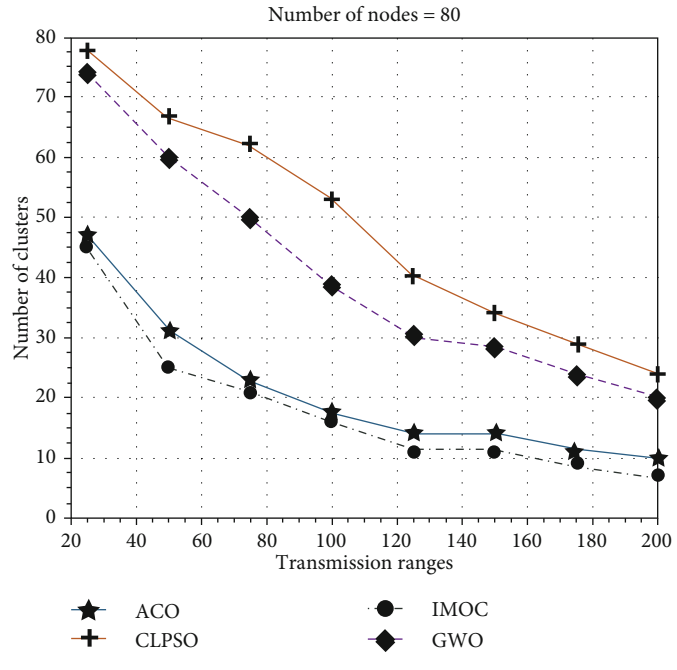


(b)

Figure 9: Continued.

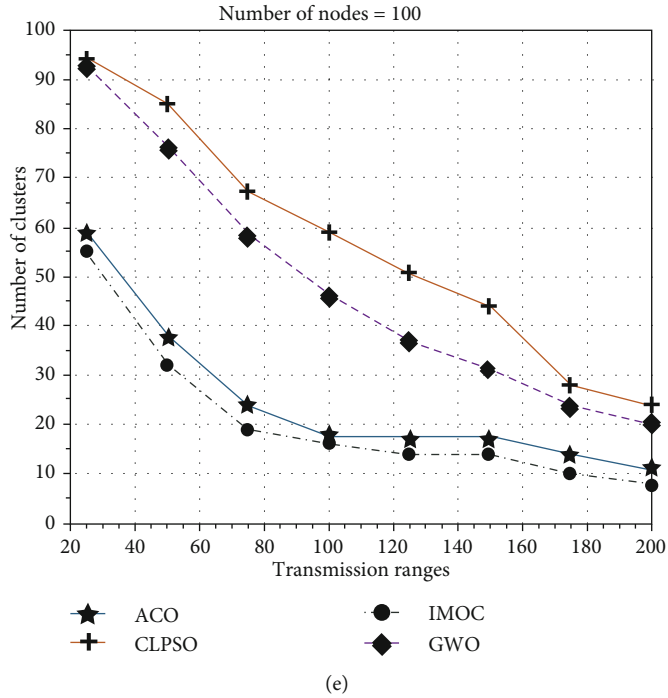


(c)



(d)

Figure 9: Continued.



(e)

Figure 9: Grid_size 2000 m × 2000 m, nodes 10 to 100.

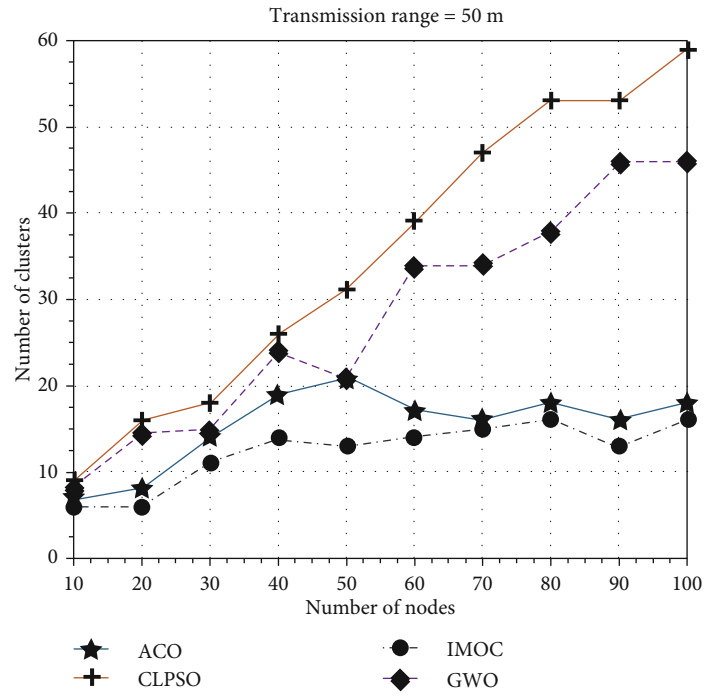
These results ensured that for any combination of nodes and transmission range for the same grid size, IMOC creates a minimum number of clusters in which its results optimized the said routing problem. The creation of a few clusters directly depends on the transmission size. The number of clusters increased with a short transmission range.

To analyze the IMOC performance with an existing algorithm, experimentations were performed with different transmission ranges. Experimentations were performed against transmission ranges of 50 m, 100 m, 150 m, and 200 m for node density from 10 to 100 nodes. Figure 6 displays the result that IMOC gives an optimal number of clusters within the 1000 m × 1000 m grid size. Simulation was performed to strengthen the results of the proposed algorithm by increasing the grid size to 1500 m × 1500 m. Figure 7 presents results for the subjected scenario for multiple nodes and transmission ranges. Figure 8(b) shows that when the transmission range is 100 and the node size 80, similar improved results for IMOC are presented.

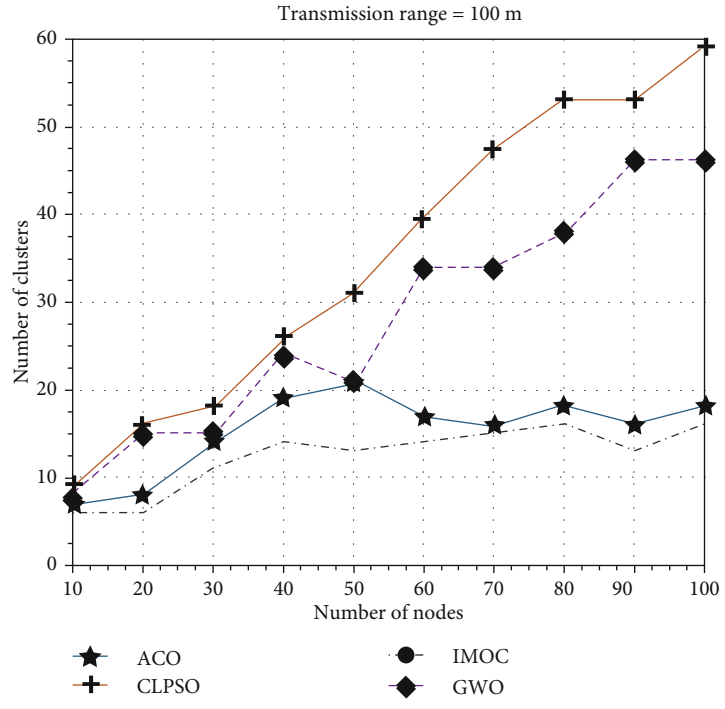
IMOC generated only 15 clusters in comparison with ACO 18, GWO 36, and CLPSO 49. In Figure 7, it can be visualized that for lower transmission range and higher transmission range, IMOC resulted in an optimal number of clusters in comparison to ACO, GWO, and CLPSO. Results have shown that IMOC performance improves for an increasing number of nodes in the grid. The efficiency and performance of IMOC remain optimal for any number of nodes and transmission range. IMOC produced 36 and 7 clusters at the transmission range of 25 m and 200 m respectively for 60 nodes, whileACO genrates 41 and 9 clusters for the same transmission ranges and number of node. ACO remains the closest minimal cluster producer for the problem under observation.

In Figure 8, the performance of the proposed algorithm was tested against multiple transmission ranges by changing node densities. On each point, IMOC performance shows better results. We can conclude that the higher the transmission range, the lesser the number of clusters; Figure 8(d) seconds this statement. It shows that our proposed technique generates 4 clusters at the transmission range of 200 m for 10 nodes, while for 100 nodes, it produces only 6 clusters. The efficiency of our proposed technique IMOC increases with the node density as compared to ACO, GWO, and CLPSO. Figure 8(d) shows that when we set a transmission range of 200 m for 10 and 100 nodes, IMOC produced only four clusters for ten number of nodes and only six for 100 nodes in Figure 9, when the grid size was increased to 2000 m × 2000 m, similar behavior of IMOC was observed by generating an optimal number of clusters. It can be seen that for extreme parameter settings for lower and higher nodes and transmission ranges, IMOC performance seems optimal. For 100 nodes and transmission range of 200 in Figure 9(e), IMOC created only 8, ACO 11, GWO 21, and CLPSO 24. Simulation results prove that IMOC for any given parameter setting creates an ideal number of clusters.

Efficient results were produced by IMOC compared with ACO, GWO, and CLPSO. Figure 10 depicts the optimal number of clusters by increasing the node density and keeping transmission ranges at 50 m, 100 m, 150 m, and 200 m. For any given parameter setting, IMOC performed best in producing an optimal number of clusters. In simulations, it can be observed that IMOC provided the best solution for any given scenario by creating a minimal number of clusters. To justify the performance of IMOC clustering in a more realistic scenario, in Figures 11 and 12, results are presented in the 3D grid by keeping the simulation setting as



(a)



(b)

Figure 10: Continued.

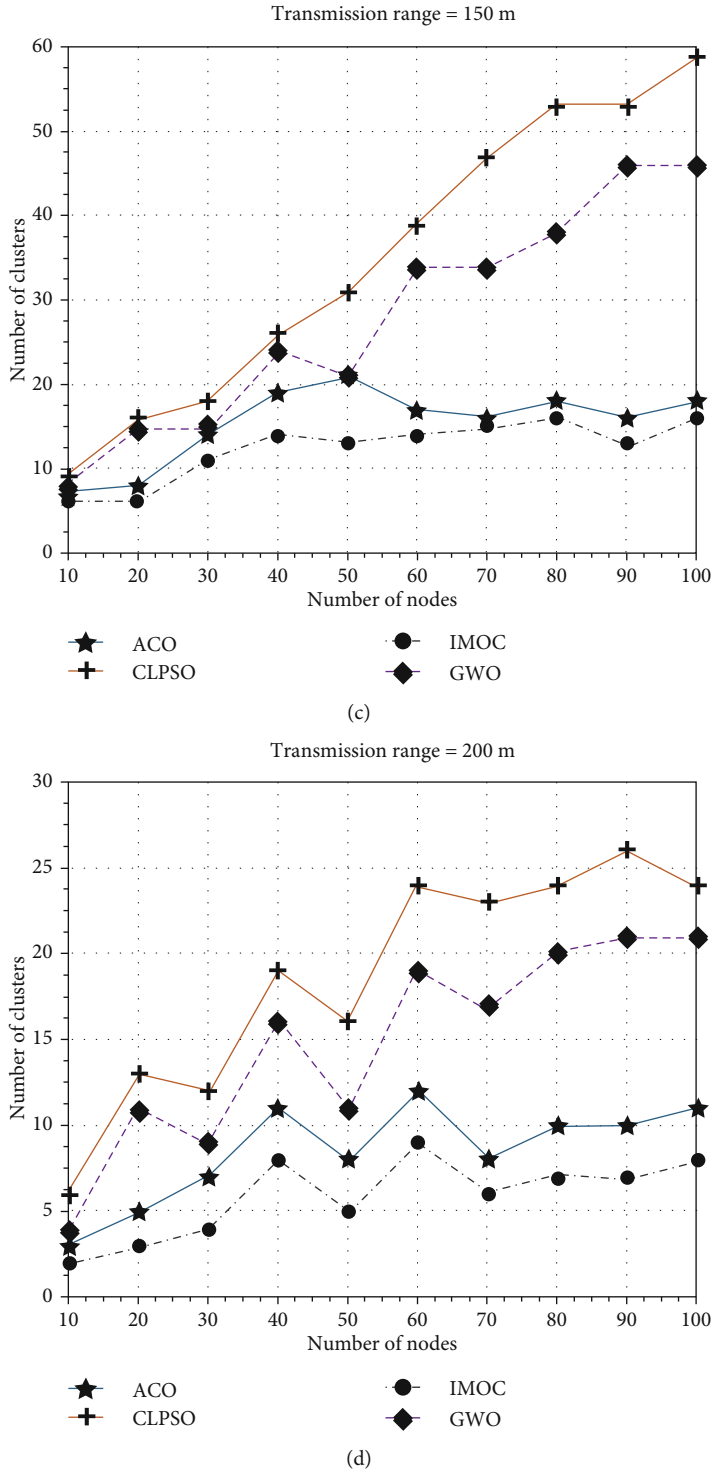


Figure 10: Grid_size 2000 m × 2000 m, transmission_ranges 50 to 200.

transmission ranges of 50 m, 100 m, 150 m, and 200 m and node density at 40 and 80 nodes for all sets of grid sizes from 500 m × 500 m to 2000 m × 2000 m.

Figure 11 depicts the results for node 40, and Figure 12 presents the results for node 80. The result justifies the performance of IMOC as optimal at any point for any scenario. The optimal number of clusters produced by IMOC for all grid sizes from 500 m × 500 m, 1000 m × 1000 m, 1500 m × 1500 m, and 2000 m × 20000 m.

4.1. Load Balance Factor. It is unrealistic to have an equal number of clusters for simulation. In some scenarios, one CH might be overloaded with a maximum number of cluster members in comparison to the cluster with a smaller number

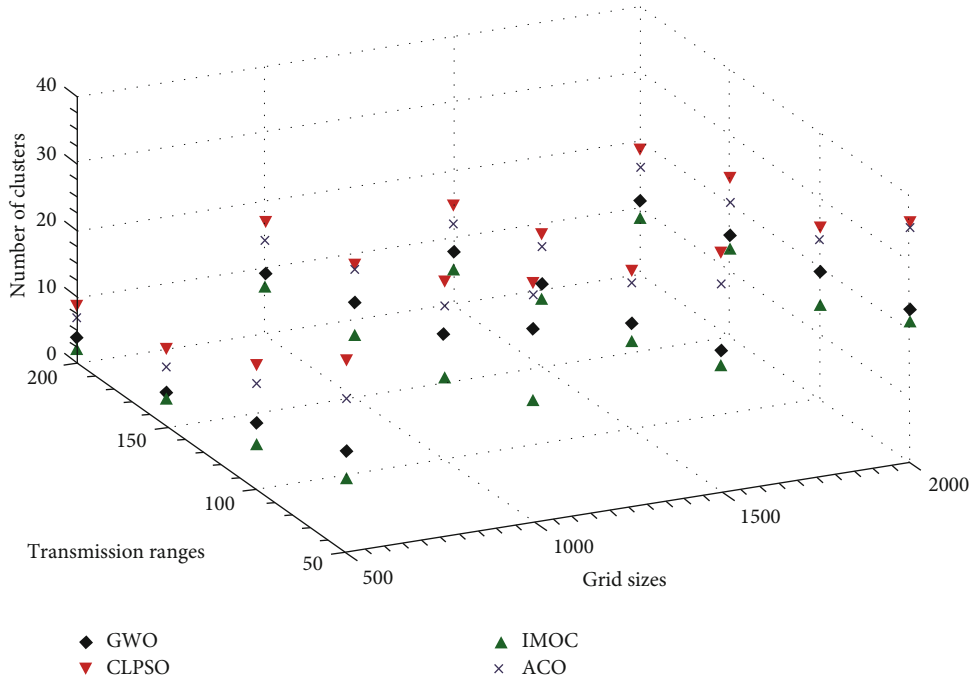


Figure 11: LBF for nodes 40.

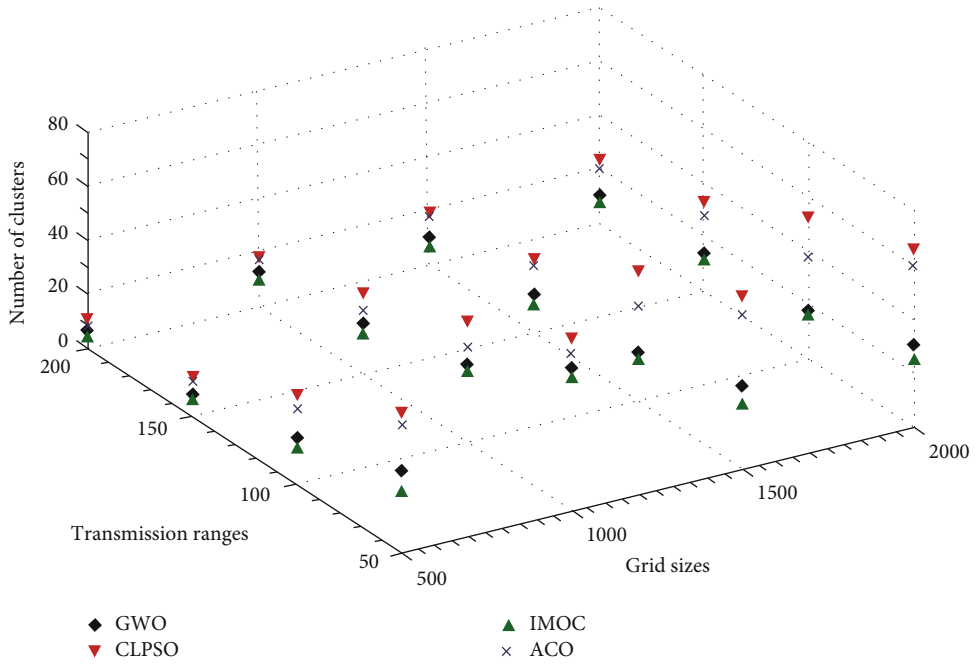


Figure 12: LBF for nodes 80.

of nodes. To mitigate the overloading of clusters, the load balance factor for each CH calculated by using

$$\text{Load balance factor} = \frac{1}{n_c \times \sum_i (x_i - u)^2}. \quad (5)$$

In equation (5), n_c represents the total number of X_i which is the cluster cardinality, and an average number of CH neighbor is $-u$. Figure 13 and 14 are for nodes 20 and 80 with varying transmission ranges from 25 m to 200 m for all grid sizes. IMOC performance is good in the case when the CH neighbor's number reached a threshold in terms of LBF.

A framework proposed to improve the routing efficiency of IMOC with support of UAV. In Figure 15, where grid size is 500 m × 500 m for node 60, IMOC created 19 clusters. CH

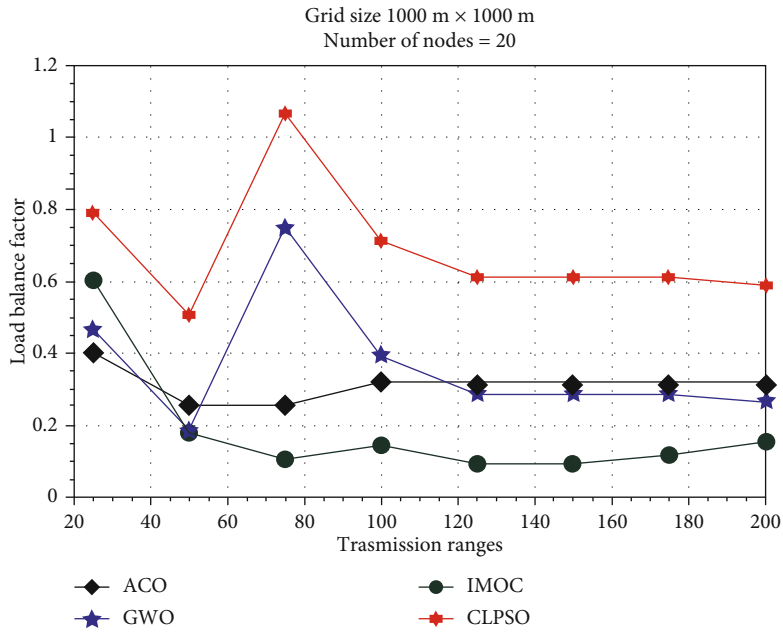


Figure 13: LBF for node 20.

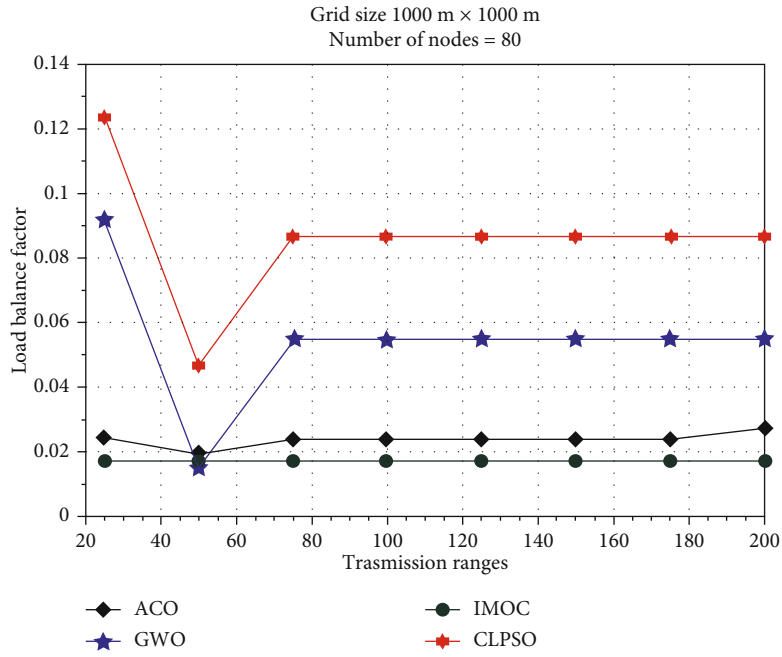


Figure 14: LBF For node 80.

and their associated members are depicted in Figure 15. Two kinds of communication existed between CH and cluster members (CM). Intercluster communication and intracluster communication can exist. To limit the overhead, CM cannot broadcast the message into the entire topology. Clustering ensured that CM can only send a request to its CH. In intra-cluster communication, the sender and receiver both are the members of the same CH. CM sends a request to CH to manage all services for cluster members. CH acts as a link between the sender and receiver to provide communication services. During inter-cluster communication, the sender and receiver both are not members of the same cluster. Conventionally, the sender node sends a request to its CH, and CH broadcasts this message into the entire topology. CHs receive this message and reply against the request, for path creation between destinations by connecting CHs. The path must be live and keep track of CHs for communication between the sender and receiver. In this scenario, hop count

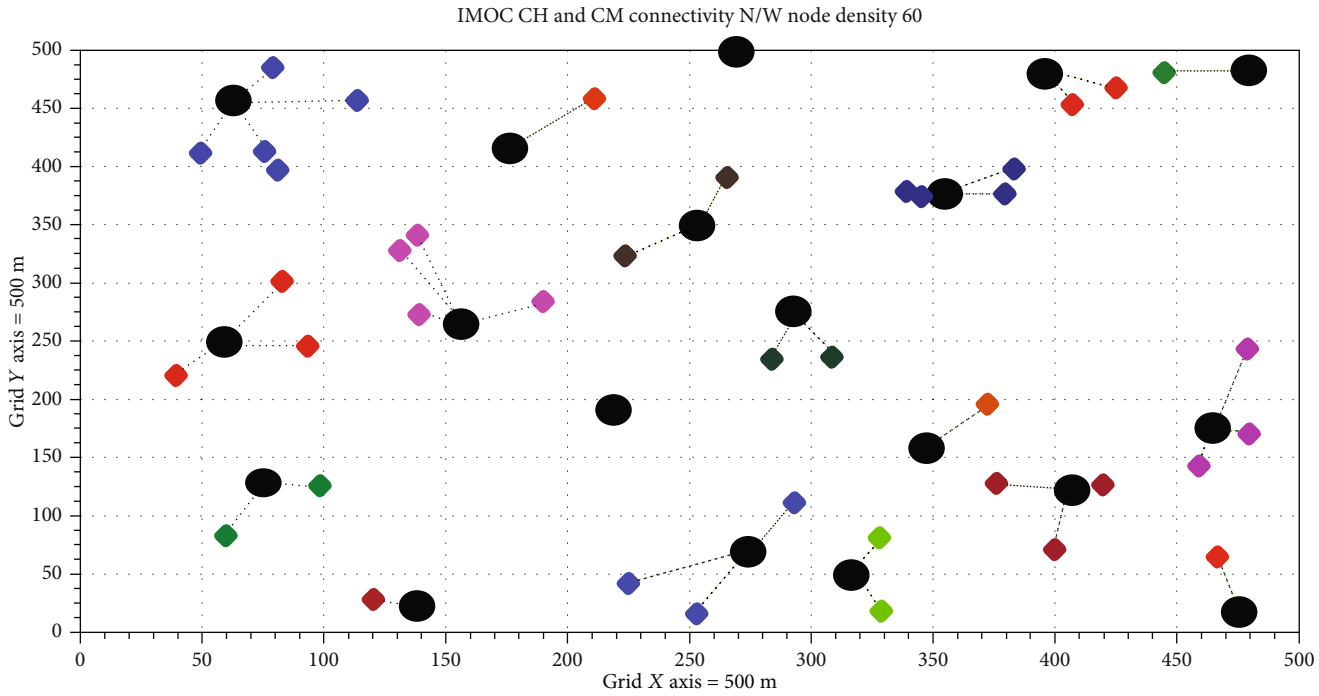


Figure 15: IMOC CH and cluster member association (500 m × 500 m).

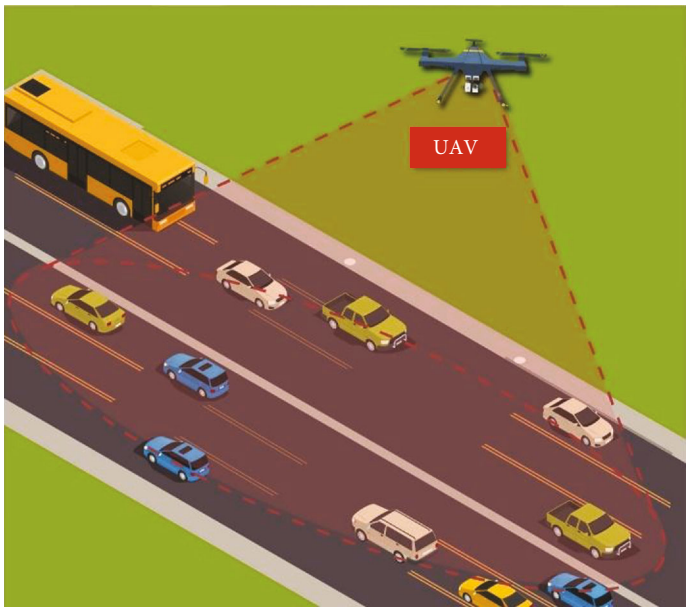


Figure 16: UAV and CH connectivity.

for multiple transmission varies from 1 to 18, depending upon the CH involved in communication. This will add additional computational complexity in routing. We proposed a solution to keep constant the hop count for all communications. Figures 16 and Figures 17 show that in the center of a topology-deployed AN (drone).

The complete grid is in the range of AN. All CHs are in the range of AN and listed. To reduce the intercluster communication problem and path construction and maintain it during communication are overhead between the sender and receiver. Considering conventional path construction based on broadcast request messages from the sender to its neighboring nodes, node reply to this request is the path members. In this problem, if the sender wants to send data outside the cluster, CH sends a request to AN, and AN sends a message to all CHs. Destination CH responds against AN request. The path will be established between the sender's CH to AN to destination CH. For all nodes, the hop count

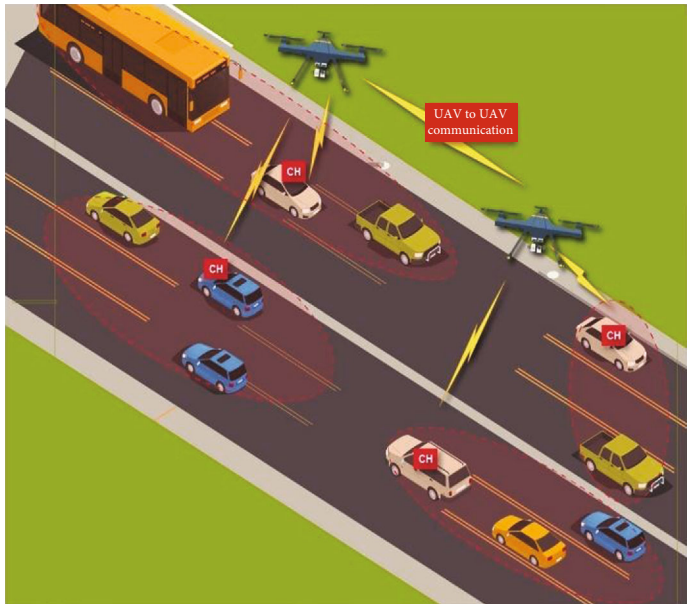


Figure 17: UAV assistance to VANET.

will be constant as two hop counts. As a result, unnecessary broadcast overhead is reduced for intercluster communication. This also improves the efficiency of the network. Clustering provides an optimal number of clusters to minimize broadcasting for intracluster communication. Additionally, AN with the IMOC clustering solution limits unnecessary broadcasting and improves network performance for intercluster communication.

5. Conclusion

To solve the VANET routing problem, IMOC solution presented an evolutionary algorithm based on cluster optimization. The MFO technique is used to find near-optimal solutions in search space. The IMOC algorithms work iteratively to find solutions from search spaces. The IMOC is an efficient algorithm for VANETs as a minimal number of clusters are achieved. It reduced unnecessary broadcasting and helped in minimizing the routing cost. The routing cost decreased as the near-optimal number of clusters was obtained from the search space. The FANET support in clustering also improved routing performance, by restricting unnecessary broadcasts and keeping the hop count constant for all communications. The proposed IMOC algorithm's efficiency was evaluated by performing a diverse set of simulations while varying topological parameters. To validate the optimization results, simulations were performed and monitored with multiple node densities in the search space with variable transmission ranges of vehicular nodes. In the said topological constraints, IMOC presented near-optimal solutions and creates a minimal number of clusters in the search space. Result comparison of renowned evolutionary algorithms, GWO, CLPSO, and ACO with IMOC shows that the proposed algorithm is the best solution for the problems under observation.

Data Availability

No dataset is utilized during the experimentation process.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A comparative survey of VANET clustering techniques," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 657–681, 2017.

[2] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3609–3626, 2009.

[3] O. S. Oubbati, A. Lakas, F. Zhou, M. Güneş, N. Lagraa, and M. B. Yagoubi, "Intelligent UAV-assisted routing protocol for urban VANETs," *Computer Communications*, vol. 107, pp. 93–111, 2017.

[4] M. A. Khan, A. Safi, I. M. Qureshi, and I. U. Khan, "Flying ad-hoc networks (FANETs): a review of communication architectures, and routing protocols," in *2017 First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT)*, pp. 1–9, Karachi, Pakistan, 2017.

[5] R. C. Biradar and S. S. Manvi, "Reliable ring based multicast routing scheme in MANET: an agent based approach," in *2009 IEEE International Conference on Automation Science and Engineering*, pp. 507–512, Bangalore, India, 2009.

[6] L. Junhai, Y. Danxia, X. Liu, and F. Mingyu, "A survey of multicast routing protocols for mobile ad-hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 78–91, 2009.

[7] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang, "Multiobjective evolutionary algorithms: a survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 32–49, 2011.

[8] A. Trivedi, D. Srinivasan, K. Sanyal, and A. Ghosh, "A survey of multiobjective evolutionary algorithms based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 21, pp. 440–462, 2017.

[9] N. Maslekar, M. Boussedjra, J. Mouzna, and H. Labiod, "A stable clustering algorithm for efficiency applications in VANETs," in *2011 7th International Wireless Communications and Mobile Computing Conference*, pp. 1188–1193, Istanbul, Turkey, 2011.

[10] S. J. Nanda and G. Panda, "A survey on nature inspired metaheuristic algorithms for partitional clustering," *Swarm and Evolutionary Computation*, vol. 16, pp. 1–18, 2014.

[11] A. Daeinabi, A. G. P. Rahbar, and A. Khademzadeh, "VWCA: an efficient clustering algorithm in vehicular ad hoc networks," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 207–222, 2011.

[12] N. Kumar, N. Chilamkurti, and J. H. Park, "ALCA: agent learning–based clustering algorithm in vehicular ad hoc networks," *Personal and Ubiquitous Computing*, vol. 17, no. 8, pp. 1683–1692, 2013.

[13] B. R. Senapati and P. M. Khilar, "Optimization of performance parameter for vehicular ad-hoc NETwork (VANET) using swarm intelligence," in *Nature Inspired Computing for Data Science*, pp. 83–107, Springer, 2020.

[14] K. D. Frank, C. Rich, and T. Longcore, "Effects of artificial night lighting on moths," in *Ecological Consequences of Artificial Night Lighting*, pp. 305–344, Island Press, 2006.

[15] P. Garg and A. Gupta, "Optimized open shortest path first algorithm based on moth flame optimization," *Indian Journal of Science and Technology*, vol. 9, p. 48, 2017.

[16] S. Mirjalili, "Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm," *Knowledge-Based Systems*, vol. 89, pp. 228–249, 2015.

[17] S. Jinyuan, Z. Chi, and F. Yuguang, "An ID-based framework achieving privacy and non-repudiation in Vehicular Ad Hoc Networks," in *MILCOM 2007 - IEEE Military Communications Conference*, pp. 1–7, Orlando, FL, USA, 2007.

[18] D. Zhu, G. Cui, and J. Huang, "The design of scheduling scheme ADCSA for mobile wireless sensor networks," *The Journal of Computer Information Systems*, vol. 8, pp. 7607–7618, 2012.

[19] J. Wu, M. Fang, H. Li, and X. Li, "RSU-assisted traffic-aware routing based on reinforcement learning for urban Vanets," *IEEE Access*, vol. 8, pp. 5733–5748, 2020.

[20] A. Dahiya and R. Chauhan, "A comparative study of MANET and VANET environment," *Journal of Computing*, vol. 2, pp. 87–92, 2010.

[21] B. Paul, M. Ibrahim, M. Bikas, and A. Naser, "Vanet routing protocols: pros and cons,," 2012, https://arxiv.org/abs/1204.1201.

[22] T. Clausen and P. Jacquet, "Rfc3626: optimized link state routing protocol (olsr)," *Experimental*, vol. 51, 2003https://www.ietf.org/rfc/rfc3626.txt.

[23] C. Perkins, E. Belding-Royer, and S. Das, *Ad hoc on-demand distance vector (AODV) routing*, University of Cincinnati, 2003.

[24] N. S. M. Usop, A. Abdullah, and A. F. A. Abidin, "Performance evaluation of AODV, DSDV & DSR routing protocol in grid environment," *IJCSNS International Journal of Computer Science and Network Security*, vol. 9, pp. 261–268, 2009.

[25] A. K. Gupta, H. Sadawarti, and A. K. Verma, "Performance analysis of AODV, DSR & TORA routing protocols," *International Journal of Engineering and Technology*, vol. 2, p. 226, 2010.

[26] D. Cao, Y. Liu, X. Ma et al., "A relay-node selection on curve road in vehicular networks," *IEEE Access*, vol. 7, pp. 12714–12728, 2019.

[27] S. Giordano and I. Stojmenovic, "Position based routing algorithms for ad hoc networks: a taxonomy," in *Ad hoc wireless networking*, pp. 103–136, Springer, 2004.

[28] S. A. Rao, M. Pai, M. Boussedjra, and J. Mouzna, "GPSR-L: greedy perimeter stateless routing with lifetime for VANETS," in *2008 8th International Conference on ITS Telecommunications*, pp. 299–304, Phuket, Thailand, 2008.

[29] Y. Yu, R. Govindan, and D. Estrin, *Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks*, UCLA Computer Science Department Technical Report, 2001.

[30] J. Bernsen and D. Manivannan, "Greedy routing protocols for vehicular ad hoc networks," in *2008 International Wireless Communications and Mobile Computing Conference*, pp. 632–637, Crete Island, Greece, 2008.

[31] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: a survey," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 12–22, 2007.

[32] S. Rosati, K. Krużelecki, G. Heitz, D. Floreano, and B. Rimoldi, "Dynamic routing for flying ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 1690–1700, 2015.

[33] W. Shi, H. Zhou, J. Li, W. Xu, N. Zhang, and X. Shen, "Drone assisted vehicular networks: architecture, challenges and opportunities," *IEEE Network*, vol. 32, no. 3, pp. 130–137, 2018.

[34] B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV assistance paradigm: state-of-the-art in applications and challenges," *Journal of Network and Computer Applications*, vol. 166, article 102706, 2020.

[35] B. Yin and X. Wei, "Communication-efficient data aggregation tree construction for complex queries in IoT applications," *IEEE Internet of Things Journal*, vol. 6, pp. 3352–3363, 2018.

[36] S. Kumar and A. Bansal, "Performance investigation of topology-based routing protocols in flying ad-hoc networks using NS-2," in *IoT and Cloud Computing Advancements in Vehicular Ad-Hoc Networks*, pp. 243–267, IGI Global, 2020.

[37] M. Moorkamp, J.-L. Wybo, and E.-H. Kramer, "Pioneering with UAVs at the battlefield: the influence of organizational design on self-organization and the emergence of safety," *Safety Science*, vol. 88, pp. 251–260, 2016.

[38] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (FANETs): a survey," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.

[39] O. K. Sahingoz, "Networking models in flying ad-hoc networks (FANETs): concepts and challenges," *Journal of Intelligent & Robotic Systems*, vol. 74, no. 1-2, pp. 513–527, 2014.

[40] O. S. Oubbati, A. Lakas, M. Güneş, F. Zhou, and M. B. Yagoubi, "UAV-assisted reactive routing for urban VANETs," in *Proceedings of the Symposium on Applied Computing*, pp. 651–653, New York, NY, USA, 2017.

[41] Y. Zhou, N. Cheng, N. Lu, and X. S. Shen, "Multi-UAV-aided networks: aerial-ground cooperative vehicular networking architecture," *IEEE Vehicular Technology Magazine*, vol. 10, pp. 36–44, 2015.

[42] O. S. Oubbati, N. Chaib, A. Lakas, P. Lorenz, and A. Rachedi, "UAV-assisted supporting services connectivity in urban VANETs," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3944–3951, 2019.

[43] O. S. Oubbati, N. Chaib, A. Lakas, S. Bitam, and P. Lorenz, "U2RV: UAV-assisted reactive routing protocol for VANETs," *International Journal of Communication Systems*, vol. 33, article e4104, 2019.

[44] H. Sedjelmaci, M. A. Messous, S. M. Senouci, and I. H. Brahmi, "Toward a lightweight and efficient UAV-aided VANET," *Transactions on Emerging Telecommunications Technologies*, vol. 30, article e3520, 2019.

[45] W. Li, X. Liu, X. Ma, X. Wang, and Y. Zhou, "UAV-aided data delivery scheme based on opportunistic virtual intersections for smart transportation networks," *Journal of Advanced Transportation*, vol. 2019, Article ID 1576908, 11 pages, 2019.

[46] Q. Wu, H. Wang, X. Li, B. Zhang, and J. Peng, "Reinforcement learning-based anti-jamming in networked UAV radar systems," *Applied Sciences*, vol. 9, no. 23, p. 5173, 2019.

[47] S. A. Hadiwardoyo, J.-M. Dricot, C. T. Calafate, J.-C. Cano, E. Hernandez-Orallo, and P. Manzoni, "UAV mobility model for dynamic UAV-to-car communications," in *Proceedings of the 16th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks*, pp. 1–6, New York, NY, USA, 2019.

[48] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV relay in VANETs against smart jamming with reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4087–4097, 2018.

[49] J. Wang, Y. Gao, W. Liu, W. Wu, and S.-J. Lim, "An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks," *Computers, Materials & Continua*, vol. 58, no. 3, pp. 711–725, 2019.

[50] D. Cao, Y. Jiang, J. Wang et al., "ARNS: adaptive relay-node selection method for message broadcasting in the internet of vehicles," *Sensors*, vol. 20, no. 5, p. 1338, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8822598, 15 pages
https://doi.org/10.1155/2020/8822598

Research Article

IoT-Based Healthcare Support System for Alzheimer's Patients

Rozita Jamili Oskouei,[1] **Zahra MousaviLou**,[2] **Zohreh Bakhtiari**,[3] **and Khuda Bux Jalbani**[3]

[1]*Department of Computer Science and Information Technology, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran*
[2]*Vali-e-Asr Hospital, School of Medicine, Zanjan University of Medical Science, Zanjan, Iran*
[3]*Riphah Institue of System Engineering, Riphah International University, Islamabad, Pakistan*

Correspondence should be addressed to Rozita Jamili Oskouei; rozita2010r@gmail.com

Received 10 May 2020; Revised 16 July 2020; Accepted 28 September 2020; Published 17 October 2020

Academic Editor: Farman Ullah

In the last decade, the Internet of Things (IoT) has become a new technology that aims to facilitate life and help people in all aspects of their lives. This technology is used for smart homes, smart grid stations, smart agriculture, health systems, transport services, smart cities, etc. The number of sensors and IoT devices along with applications is used for monitoring the health condition of patients. These devices will monitor the movement of targeted patients at home or out of their homes. Based on their behavior and movement, the treatment will be provided to Alzheimer's patients. The data will be collected from multiple sensors installed at patient's homes and smartwatches for checking their blood pressure level and temperature, which is too important in the current Corona Virus Disease (COVID-19) pandemic for these types of patients. On the other hand, due to the diminishing mobility of people around the world, increasing environmental pollution and stress which is caused by modern machine life and various brain and neurological diseases including Alzheimer's, Parkinson, etc. are widespread among people all over the world. The different types of communication protocols such as Message Queue Telemetry Transport (MQTT) and WebSocket (with authentication and autoclosing of connection) for sensors and the smartwatch have been used. The secure backend admin panel is used for tracing the location of doctors, patients, and ambulance. These protocols are implemented with security to protect the privacy of patients also.

1. Introduction

IoT technology has a huge impact on human life in all aspects like medicine, health, industry, transportation, education, and agriculture from the last decade. This technology uses sensors or actuators to understand the state of the surrounding environment. Most of them connected via these communication technologies such as WiFi and Global System for Mobile (GSM) to communicate with control centers and send data collected from the environment and to help for making decisions at remote control centers. Smart homes [1] are currently being developed with great acceptance by people around the world.

In this paper, the focus is on patients of Alzheimer's disease, which is the most common neurological disorder in the last decade. It is a type of dementia that occurs for most elderly people. In this type of illness, a person becomes oblivious as he/she is not able to perform his/her daily tasks independently and needs to have a person in the family always care for their behaviors and health. Therefore, for families with Alzheimer's patients, the cost of hiring a nurse or continuing care of this patient is high. However, it is expected that we will be able to remotely monitor the behaviors and health status of these patients using the facilities that the Internet of Things can provide, as to decrease the extra expenses and timely response to these patients. The IoT can play a vital role in this current COVID-19 pandemic situation. Different types of smartwatches, sensors, and actuators are installed at the home of these patients. These IoT devices are used to collect the data regarding their temperature, medicine intake timings, and movement. The different types of sensors and actuators have been used for their secure data transfer that the existing protocols have been used. These protocols have been used under the umbrella of Web of Things (WoT) like MQTT, WebSocket, and HTTP. The data collected from these IoT devices have been secured during

communication and as it is stored at the cloud servers. Several methods have been performed for monitoring the health conditions of patients. One of them is neural networks and Bayesian [2] for monitoring the skin in real-time with the help of IoT.

The research is organized in five sections. In the second part, we will review the background of researches which is done related to the use of the Internet of Things in the care of neurological patients, especially Alzheimer's patients. Section 3 briefly examines Alzheimer's disease medically and explains the symptoms, the causes of this disease, and the care needed for these patients. In the fourth section, we will introduce the Internet of Things technology and discuss the user of its facilities to help Alzheimer's patients who are living in smart homes. Finally, in Section 5, we will present the conclusions.

2. Related Words

The authors [3] have surveyed different types of sensors based on their potential benefits for the healthcare system. They have listed known sensors for observing physiology, health condition, intellectual, and full of feeling parts of individuals. Alzheimer's disease is named for the first time with the name of a doctor who first described it (Alzheimer) [4]. There are numerous researches which are done on Alzheimer's disease. However, in this part, we will discuss a few of these researches. The authors [5, 6] have primary scientific categorization; to be specific, the scientific classification of related work on the IoT of real-time monitoring health condition in telemedicine applications was introduced and considers the identification with the IoT issues that were examined and talked about in customer and server sides. The security impediment in portraying or understanding the components of real-time health systems observed prompted the examination of an extrascientific categorization layer and improved the security level.

Dieckmann and colleagues [7] investigated the tool that is used for Alzheimer's disease knowledge test (ADKT). The metrics used for these tests ADKT by these researchers showed that five important metrics for measuring the outcomes of caring of Alzheimer's patients have been proposed by different researchers: [8] ADKT, UAB-ADKT for health professionals [9], DQ [10], KAML-C [11, 12], and ADKS.

Karlin and Dalley [13] conducted a study on the relationship between people's concerns about dementia and Alzheimer's disease, the likelihood of doing tests and screenings, and applying methods to detect changes in cognitive status or patient performance, to detect the disease early. They made a descriptive study using data from Porter Novelli's SummerStyles 2013 online survey. They used chi squares with case-level weighting used for analyzing data of the 6,105 people over the age of 18 who were surveyed; 4033 (66% of all of these people who are selected) responded to the surveys.

The results of this study showed that 13% of survey respondents were very concerned about Alzheimer's disease. Women were more concerned about the disease than men. Those who looked after Alzheimer's patients were more concerned about the disease than others. Women were also more welcomed than men by tests to assess their probability of Alzheimer's. The information gained from screening can be useful in developing communication strategies to address public concerns about Alzheimer's disease and increase the likelihood of early detection of the disease.

Brian researched the detection of Alzheimer's disease [14]. The results of this study showed that the biggest challenge in this area is the lack of knowledge needed for the early diagnosis of this disease, not the mistake of diagnosis.

Jagadeeswari and colleagues [15] attempted to provide preliminary evidence for the acceptance, validation of new knowledge of the Alzheimer's Disease Knowledge Scale (ADKS), content updating, and psychometric testing to test knowledge of Alzheimer's disease of doctors and nurses. They used traditional scale development methods to create items and evaluate the psychometric properties of various types of collected samples. Finally, 30 items were selected for use in the diagnosis of the disease, which can be completed in just 5 to 10 minutes by the survey participants, and the results of their analysis can be used to diagnose, evaluate, and understand symptoms, course of the disease, its impact on the patient's life, the care and treatment, and the management of all stages of diagnosis until treatment. Preliminary results show that the ADKS has high reliability (in test phase) and reliability (in content, predictability, etc.). Finally, the ADKS can be designed for use in research and practice and can be used to increase knowledge of Alzheimer's disease and to help patients and their caregivers and physicians.

Srimathi et al. [16] conducted studies on cloud computing, fog computing, big data analytics, IoT, and mobile-based applications and emerging technologies in the field of personalized medical care systems. The researchers looked at the challenges of better designing a healthcare system for the early detection of diseases and explored possible solutions for delivering health-related care electronically and safely. This study emphasizes the need for developing high quality and precision electronic healthcare systems.

Smyth and colleagues [17] proposed a new idea to modify the existing access control system to detect medical conditions associated with the brain and receive a timely response from physicians in the time of medical emergencies. The idea is to improve the quality of medical services available to people around the world. In the follow-up to the study, the researchers developed a simple device that could help physicians to find brain abnormalities as quickly as possible in humans. The device can also be highly effective in the care of these patients remotely using IoT technology.

Hoe and colleagues [18] discussed on the physicians' and clinicians' level of knowledge and information about Alzheimer's disease or ADKS of physicians and clinicians across Australia. These researchers asked from 360 doctors and therapists some questions about their level of knowledge of Alzheimer's disease and the history of the practical care that they provided to their Alzheimer's patients, by emailing these questions to them. Results from the responses provided by these 360 individuals indicated that most of them had moderate information on the factors influencing Alzheimer's

disease, and only those who specialized in neuroscience or had practically encountered patients and those who attended the relevant workshops had relatively good information about this disease.

Yao and colleagues [19] conducted a study on the prevalence of Alzheimer's disease in the United Kingdom (UK). This study highlights the UK's top priority for tackling Alzheimer's disease, which is expected to become a major challenge for the world in 2020. There are more than 800,000 people in the UK with this disease. Since the disease has a huge impact on the lives of affected people and their families and need costs of £26 billion in a year to care for these patients in the UK and there is no cure for the definitive treatment of the disease, by 2050, nearly two million people in the UK are expected to have Alzheimer's disease. The main goals of this research are to develop and evaluate Maintenance Cognitive Stimulation Therapy (MCST) for Alzheimer's patients. A follow-up plan for the care of these patients Carer Supporter Program (CSP) was then proposed, and the effectiveness of this program was compared with the usual care provided to patients. Finally, a home treatment package (HTP) was developed for Alzheimer's patients, and the trial of this package was practically tested. The results showed that MCST can improve people's quality of life and reduce the cost of caring for sick people. Studies of CST implementation also show that many employees receive CST training over a one-day training period. This training course increases the capacity of medical staff to care for Alzheimer's patients. Management of care for Alzheimer's patients has also shown that it reduces long-term care for these patients and reduces their behavioral problems, creating an easy-to-use home user guide to help caregivers of these patients at home and also prevent hospitalization for dementia patients. Finally, the researchers point out that, due to the huge financial burden, they have not been able to fully put their proposed steps into practice.

Shaikh and colleagues [20] conducted a study to evaluate the knowledge, attitude, and approaches to care of Alzheimer's patients by Chinese medical professionals since China has the largest population in the world. A cross-sectional study was conducted on 450 randomly selected health professionals from Changsha, China. A questionnaire was sent to each of them for asking questions about their knowledge of Alzheimer's disease and their suggested ways to care for Alzheimer's patients and so on. 390 specialized out of a total of 450 selected one responded to this questionnaire, and the results of analyzing their responses showed that 87% of them had very poor knowledge of Alzheimer's disease and their knowledge was directly related to their experience of caring for Alzheimer's patients. Further, most of them were reluctant to provide practical care for Alzheimer's patients. Finally, in this study, the researchers suggested that a multilevel approach, including providing training courses for the community of health professionals and formulating policies and resources to meet the demands related to the delivery of Alzheimer's care services in China, is urgently necessary.

As in an early work, a lot of focus is on privacy and security issues of IoT related to healthcare systems. Instead, the solutions are also provided for the security of IoT devices and sensors from security breaches and privacy issues regarding patient information. To fill this gap in this paper, secure communication between IoT devices, sensors, and IoT apps along with the security of that information is stored. This is implemented with the help of WoT security standards and unique identification number also created for each installed device a client-side.

3. Overview of Alzheimer's Disease

Alzheimer's disease (AD), in 60-70% of cases, leads to dementia. The word dementia refers to a set of symptoms that can include memory loss and difficulty in thinking, forgetting words and stuttering, or delaying speech when speaking.

From a scientific perspective, Alzheimer's is a chronic neurological disorder that usually starts slowly and worsens over time. The most common primary symptom of this disease is difficulty in remembering recent events (so-called short-term memory loss). Alzheimer's is a progressive disease that affects the brain. This means that over time, more parts of the brain are damaged. Currently, it affects about 6% of people who are in 65 years of age and older, and unfortunately, no definitive treatment has been offered so far [21].

In other words, Alzheimer's is a type of memory disorder which occurred by brain cell death. Therefore, dementia patients are unable to continue their normal social lives. The disease can be controlled if it is diagnosed at an early stage. Therefore, early detection of Alzheimer's is essential for the treatment of patients. Usually, people living with the patient can diagnose Alzheimer's due to changes that are happening in the patient's behaviors and cognitive impairment (loss of cognitive abilities). However, it is difficult to diagnose Alzheimer's in the early stages of the elderly people who are living alone [22]. The Journal of the American Medical Association has acknowledged that Alzheimer's is diagnosed only when the patient's memory and cognitive function are severely affected and impair the individual's ability to perform daily tasks.

Alzheimer's patients often do the same task repeatedly, make the same gesture, say the same words, or ask the same question over and over again [23]. Repetition of previous actions and words in Alzheimer's is common due to memory loss and general behavioral changes. One may repeat daily tasks such as shaving or may be tempted to move home appliances [24]. Alzheimer's patients often see changes in their sleep patterns. For example, 20 minutes of sleep per day may increase to several hours in a day [25].

Due to the gradual loss of memory, one of the biggest threats to those suffering from Alzheimer's disease is being confused about what to do. The most common symptoms of Alzheimer's disease include increased daytime sleepiness, nocturnal confusion, confusion, and anxiety. All of these behavioral abnormalities caused by Alzheimer's disease are called "Sunset".

More recently, in 2019, a comprehensive study of Alzheimer's disease has been conducted in the United States for finding the most common cause of dementia [26], with details on how to diagnose the disease, the prevalence of

the disease, its mortality rates, and how to present it. Primary care for patients, the cost of maintaining these patients, and how to provide long-term care to patients at home and in the hospital are discussed in this research.

Alzheimer's disease is a type of brain disease. It is also a degenerative disease which means it worsens over time. Alzheimer's disease is thought to become hidden for 20 years or more, and the patient cannot detect minor changes in the patient's behavior. Over the years, people have noticed symptoms such as memory loss and speech problems. The cause of these symptoms is that nerve cells are damaged or destroyed in areas of the brain that are involved in thinking, learning, and memory (cognitive function). In the advanced stages of the disease, one loses the ability to perform daily activities and is said to be suffering from dementia or Alzheimer's disease. The brain of a healthy adult has about 100 billion neurons, each with long and branching branches. These branches enable individual neurons to communicate with other neurons. Such connections are called synapses. The brain has about 100 trillion synapses. They allow signals to move quickly in the neural circuits of the brain, which is related to emotional and emotional messages, memories, thoughts, movements, and skills. Beta-amyloid protein fragment accumulation (also known as beta-amyloid plaque) outside the neurons and accumulation of abnormal tau protein (called tau tangles) in the neurons are two of the changes caused by Alzheimer's disease. Beta-amyloid plaques may interfere with the neuronal communication of neurons in synapses, leading to cell death, while the tau tangles block the transport of nutrients and other essential molecules into the neurons. By increasing the amount of beta-amyloid, we reach a stage where tau is abnormally distributed throughout the brain [27].

3.1. Symptoms of Alzheimer's Disease. The main symptoms of Alzheimer's disease include the following [27]:

(i) Loss of memory that disrupts human daily life

(ii) Occurring problems in planning or solving problems

(iii) Difficulties in doing chores at home and work

(iv) Confusion on choosing a time or place

(v) Difficulties in understanding visual images and spatial relations

(vi) Difficulties in using new words in conversation or writing

(vii) Missing objects and being unable to retrieve those objects

(viii) Decreasing decision-making power in work or social activities

(ix) Changes in the patient's mood and personality

3.2. How to Diagnose Alzheimer's Disease. It is also possible to diagnose Alzheimer's disease by a general practitioner. The steps to identify this disease by a general practitioner (GP) are as follows [27]:

(i) The doctor examines the patient's family and medical history (such as history of psychology and cognitive and behavioral changes)

(ii) In the next step, the physician will obtain information about his or her mood through a close relative of the patient

(iii) Next, the physician will perform cognitive tests, physical exams, or neurological tests

(iv) Then, the physician may request that the person be subjected to magnetic resonance imaging. Magnetic resonance imaging can help identify brain abnormalities such as the presence of a tumor or evidence of a stroke or brain attack

3.3. Important Factors in Alzheimer's Disease. The factors affecting Alzheimer's disease include the following:

(i) Family history: people with one of their parents, brothers, or sisters with Alzheimer's disease are at greater risk of developing the disease in comparison to people with no prior history of relatives [28]. Even people who have more than one member of their first-degree family have been at increased risk of developing the disease [29]. When the disease develops in the family, it can be due to genetic, environmental or lifestyle factors, or a combination of them. Alzheimer's disease may be one of the hereditary causes of the inheritance of the $\varepsilon 4$ Apolipoprotein E (APOE) gene

(ii) APOE $\varepsilon 4$ gene: this gene provides a blueprint for a protein that carries cholesterol into the bloodstream. Each individual inherits a form of the APOE gene—$\varepsilon 2$, $\varepsilon 3$, or $\varepsilon 4$ from each parent. The $\varepsilon 3$ form the most common gene transmitted between individuals of a family [30], which inherits approximately 60% of the US population $\varepsilon 3$ from both parents [31]. The forms of $\varepsilon 2$ and $\varepsilon 4$ are much less common. It is believed that having $\varepsilon 3$ form does not increase or decrease the risk of AD, while having $\varepsilon 2$ may reduce the risk of Alzheimer's disease. However, the $\varepsilon 4$ form increases the risk of AD at a young age. Those who inherit the two $\varepsilon 4$ genes have a much higher risk of developing Alzheimer's disease. Researchers estimate that between 40 and 65 percent of people with AD have one or two copies of the APOE $\varepsilon 4$ gene [32–34]. However, the inheritance of the APOE $\varepsilon 4$ gene does not guarantee that a person will have AD. It is believed that many nongenetic factors are involved in the development of AD

(iii) MCI (Mild Cognitive Impairment): this is a condition in which a person has mild but observable changes in his or her intellectual abilities that are

understood by the affected person and their family members and friends, but it does not affect the person's ability to perform daily activities. People with MCI who have memory problems are more likely to develop Alzheimer's disease than those without MCI [35, 36]. However, MCI does not always lead to dementia and may be due to the use of a particular drug and can be resolved after a short time

(iv) Cardiovascular disease risk factors: much evidence suggests that brain health is closely linked to overall cardiovascular health [27]. The brain is nourished by one of the largest networks of blood vessels in the body. A healthy heart helps to get enough blood through the blood vessels to the brain, and healthy blood vessels ensure that the brain receives oxygenated and nutrient-rich blood for its normal function. Many factors (such as smoking [37, 38], obesity (especially in middle age) [39, 40], diabetes [41, 42], high cholesterol in adolescence [42], and hypertension in adolescence [43, 44]) can increase the risk of AD. Also, new evidence suggests that having a proper diet, such as a diet with a vegetable meal and vegetable oil use, may reduce AD and the risk of dementia

(v) Education: uneducated or illiterate people are at a higher risk of developing Alzheimer's and dementia compared to those with higher education [45–47]. Some researchers believe that when a person is educated for many years, a cognitive reserve is created in his/her brain that enables him/her to largely neutralize the changes that lead to Alzheimer's disease [48–50]. According to cognitive storage, having years of training enhances the communication between neurons in the brain and enables the brain to compensate very early brain changes that are happening due to Alzheimer's by using alternate pathways to complete a cognitive task

(vi) Social and cognitive engagement: many studies have shown that having social interactions can reduce Alzheimer's disease [51, 52]. Being socially and cognitively active may help to create a cognitive reserve, but the precise mechanism of this fact is unknown. Fewer studies have been conducted on the relationship between social and cognitive interactions and the likelihood of AD and dementia, and more research currently is necessary to be done to fully understand how social and cognitive interactions with biological processes reduce the risk of Alzheimer's disease

(vii) Traumatic brain injury (TBI): moderate or severe brain injury can increase the risk of Alzheimer's disease [53]. TBI is a disorder of the normal functioning of the brain caused by blows on head or skull injuries. Not all blows to the head cause impaired brain function. TBI is caused by a brain injury that results in loss of consciousness or forgetfulness for more than 30 minutes. If the loss of consciousness or posttraumatic forgetfulness lasts more than 24 hours, severe injury is considered. Half of all moderate or severe TBIs are caused by motor vehicle accidents [54]. Moderate TBI, on average, doubles the risk of AD, but if TBI is severe it can increase the risk of AD by 4.5 times [55]. Groups that suffer from repeated injuries, such as boxers, soccer players, and warriors, are at a higher risk of developing dementia, cognitive impairment, and neurological illnesses than others [56, 57]

3.4. Nursing Care for Alzheimer's Patients. In general, the care of Alzheimer's patients is classified into two types.

3.4.1. Pharmacological Treatment. Pharmacological treatments are treatments that prescribe medications to stop the disease or treat its symptoms. There is currently no drug that can stop the death and neuronal defects in the brain which are caused by Alzheimer's disease. Several drugs have been introduced to reduce or stop brain cell death, but only 5 of them have been approved by the US Food and Drug Administration. These drugs improve the symptoms of Alzheimer's by increasing the number of chemicals called neurotransmitters in the brain [58].

3.4.2. Nonpharmacological Treatment. Nonpharmacological therapies are methods in which some methods such as cognitive training and behavioral interventions are used to treat patients. The nonpharmacological treatment cannot reduce the amount of neuronal death and defect in the brain which is caused by Alzheimer's disease. In general, the goals of nontherapeutic approaches include the following:

(i) The goal of some of them is to compensate for abnormalities to maintain cognitive function or help the brain

(ii) The goal of some nonpharmacological approaches is to improve the quality of life of patients

(iii) Other goals of nontherapeutic approaches include reducing behavioral symptoms such as depression, anorexia, wandering, sleep disorders, anxiety, and aggression

Previous studies have shown that few nonpharmacological therapy methods may improve or stabilize cognitive function, daily activities, behavior, mood, and quality of life of Alzheimer's patients [59].

4. IoT Technology

The Internet of Things (IoT) provides how intelligent objects can be interconnected in computing environments everywhere [60]. Internet infrastructure as a global platform plays an important role in the formation of the Internet of Things and enabling it to communicate between physical objects. Innovations in the IoT are accomplished by incorporating sensors into the objects that make them intelligent and allowing physical infrastructures to be integrated around the world and able to connect using communication technologies.

The overall architecture of the Internet of Things is illustrated in Figure 1.

The term "Internet of Things" refers to an Internet-based architecture that facilitates the exchange of services, information, and data between billions of predominantly intelligent objects. The idea of the Internet of Things was first proposed by Kevin Ashton in 1998 and has been the focus of many universities and industries in the short term [61].

In some literature and research, the Internet of Things by the name of the Internet which has been mentioned everywhere and at all times provides a link between all of these objects to facilitate and make life easier for people in all situations.

From Figure 1, we can see that hardware and software solutions work together to create an Internet of Things object. The Internet of Things must be able to communicate between billions or trillions of nonhomogenous devices over the Internet, so there is a critical need to create a layered architecture for flexible Internet of Things.

The IoT domain covers a wide range of standard or nonstandard technologies, software platforms, and applications. Therefore, a reference architecture alone cannot be used as a blueprint for all possible implementations. In Figure 1, we define the architecture of the Internet of Things as a framework in which objects, people, and cloud services interact to facilitate the delivery of functional tasks. Therefore, Figure 1 can be considered as a reference model for IoT [60].

According to the Institute of Electrical and Electronics Engineers (IEEE) definitions and standards, an IoT system is a network of networks that typically connect a large number of objects/sensors/devices through information communications and infrastructures to process value-added services and services through processing. Smart to provide data and various application managements.

IoT is a computational concept that envisions a future in which physical objects will be connected to the Internet and able to identify themselves to other devices. It was first introduced with RFID technology as a communication method, and subsequently, sensor technologies and other wireless technologies were added to ICT.

According to the Internet of Things European Research Cluster (IERC) definition, the Internet of Things is a dynamic global network infrastructure with capabilities for self-regulation based on collaboration standards and communication protocols, as physical and virtual objects and physical properties and virtual characters can be identified and utilized intelligent interfaces and integrated seamlessly into information networks [62].

IoT technology has been developed in the last decade. By using different sensors, we can understand the environment and communicate with different objects in remote environments with the help of one of the communication technologies (such as WiFi and GSM). In the case of Alzheimer's patients, by installing sensors in various locations of the smart home, we can fully monitor the movements and activities of these patients indoors. We can also monitor the movement of these patients anywhere outside the home by installing sensors on the clothing or body of patients.

By analyzing the data collected by these sensors, which are sent to embedded control centers through one of the communication technologies (such as WiFi), we will be able to extract useful information and this information can be used by physicians and health professional to be effective in making appropriate decisions in the care or treatment of these patients [63].

4.1. Application of Internet of Things in Medicine. Some of the applications of IoT technology in the field of healthcare include tracking objects, staff, and patients; identifying and authenticating individuals; and collecting data of patients' physical health automatically [64].

Remote patient status monitoring systems are used by physicians and medical staff to control blood pressure, temperature, heart rate, respiratory rate, etc. It has been the focus of many medical centers in recent years. Other uses of the Internet of Things in the health and medical world include identification and authentication including unique patient identification to reduce sudden and harmful events for that patient, comprehensive electronic medical record-keeping, identification of neonatal vital signs, and identifying them in hospitals to prevent them from coordinating their delivery to their parents.

Automated data collection and transfer are mainly used to reduce form processing time, automation, automated care, rapid audit, and inventory management in medical centers. The sensors are capable of focusing on patients, in particular detecting the condition of the patient, providing real-time information to the patient and their companions to improve patients' health indicators. Other applications include telemedicine, monitoring compliance of patients' prescriptions with their current conditions, and alerting patients when critical health status and vital signs are present.

It is noteworthy that the Internet of Things (IoT) uses communication technologies (such as RFID, NFC, WSN, Wi-Fi, and Bluetooth) to communicate between different objects in the medical world. To continuously monitor patients' vital functions (such as body temperature, blood pressure, heart rate, cholesterol levels, and blood sugar), sensors send data to doctors or physicians from long distance.

The dependence on using IoT to provide health and medical care to patients is increasing day by day. This is due to the low cost of patient care and the high quality of services that are provided by IoT [64]. Based on the unique biological, behavioral, social, and cultural characteristics of each patient, the integrated function of providing comfort and well-being healthcare and protecting each patient from a critical situation is defined as personal healthcare. Such care allows physicians and nurses to follow the principle of primary care, "delivering the right care for the right person at the right time," resulting in better outcomes and higher satisfaction for patients and their families. In other words, in this way, the provision of health and medical care at a much lower cost is provided and the primary pathological diagnosis is made by providing preventive services to worsen the patient's condition.

An appropriate service refers to the prevention, early detection of disease and primary pathology, and home care

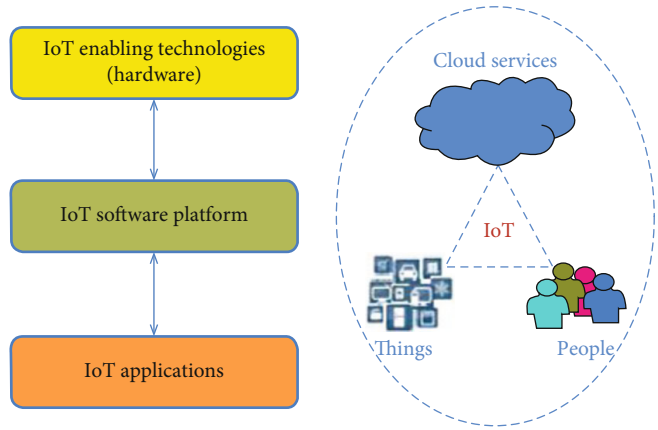


Figure 1: Architecture of the Internet of Things architecture [60].

rather than hospitalization and very expensive care in the clinic by checking patients' status to ensure their total health status. It is expected that using the Internet of Things (IoT) will be able to manage personalized care services and create a digital ID for each person. Various tools are used in health centers, they can communicate with each other and provide system-to-system communication services in any location, and it will have a significant impact on early identification and easier treatment of patients. The classification of IoT-based personalized healthcare systems is summarized in two categories including clinical care and remote monitoring, which are briefly described below [65].

4.2. Clinical Care. IoT-based surveillance systems are used for hospitalized patients whose physiological status requires ongoing attention. These surveillance systems use sensors to collect patients' physiological information, which is analyzed using gateways and applications in cloud computing, and ultimately, the results are stored on the cloud server. This information is then transmitted wirelessly to professionals for analysis on a digital basis. It is unnecessary to control the vital signs of patients at regular intervals by a health professional using the Internet of Things. In other words, the Internet of Things creates an automatic, continuous flow of health information and vital signs for patients. Thus, the quality of medical care is enhanced through continuous attention, which reduces the cost of care and eliminates the need to permanently monitor the patient [64].

4.3. Remote Monitoring. Lack of easy access to effective health surveillance systems can lead to many risks to patients' health or some illnesses remain unknown for long periods which causes disease progression, which is currently one of the problems in the health field all around the world. Some small but powerful wireless devices that interconnect via IoT and monitor patients' health status are currently being developed as a viable solution to this problem. An example of a patient remote control system using the Internet of Things is shown in Figure 2.

As can be seen in Figure 2, patients' health data can be safely collected using these solutions. Various types of sophisticated sensors and algorithms are used to analyze data and then to share data and results through wireless communications. Medical professionals can remotely provide patients the necessary recommendations to maintain their health.

4.4. Health and Treatment Networks in the Internet of Things. The IoT Health Network or the IoT Network for Health and treatment (abbreviated as IoThNet) is one of the essential elements of the IoT Health Internet. This network provides access to the IoT backbone and provides access to healthcare-related communications.

4.5. Application of Internet of Things to Improve the Quality of Life of Alzheimer's Patients. Three conditions can be attributed to Alzheimer's disease: wandering, dementia, and severe memory loss. Several types of research have been done by various researchers about IoT application in improving the life of Alzheimer's patients. In this section, we discuss some of these researches.

Ashfaq and colleagues [4] proposed a mobile-based system for Alzheimer's patients. They used the smartphone app to guide Alzheimer's patients and assist them in their daily activities. IoT technology can play a major role in helping Alzheimer's patients. This researcher developed a special Android application to help relatives and guide Alzheimer's patients. The program has various games and competitions to enhance the patient's brain functions and display progress reports. It also provides tips on where to place different objects and daily reminders of the food and medicine to Alzheimer's patients. It also utilizes GPS location capabilities to provide location care for Alzheimer's patients. IoT technology is used to measure patient status using wireless communication technology. The main purpose of the system is to create an environment for patient care at home and reduce the costs of patient healthcare.

Sindhu and colleagues [66] provided information to family members conducting Alzheimer's care in homes that are equipped with smart home automation. They were offered a solution to take care of their patients full-time. People who care for Alzheimer's patients in nursing homes can also use this system. The results of the proposed solutions by the

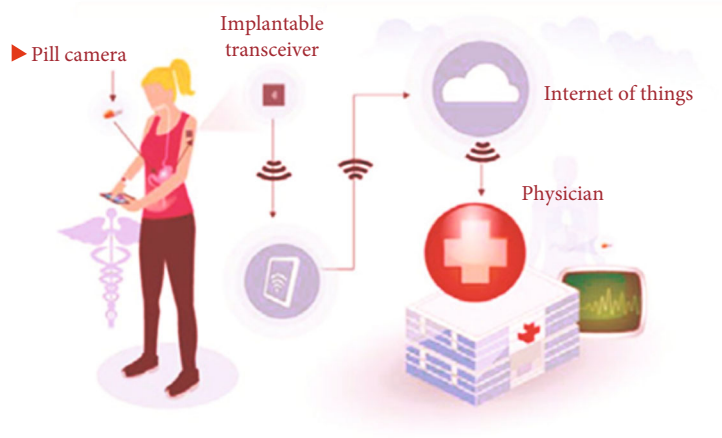


Figure 2: Telecommunications of patient care system using IoT [64].

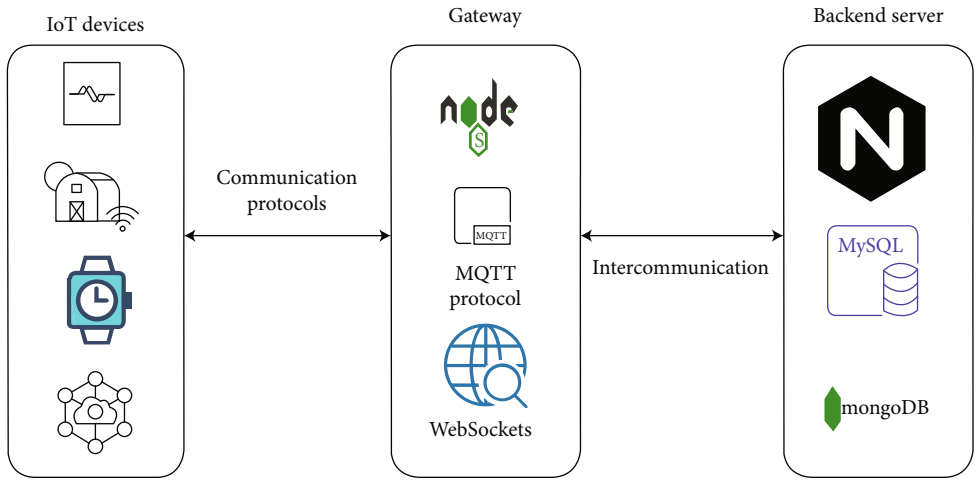


Figure 3: Proposed system diagram for improving Alzheimer patients.

researchers showed that, by applying the proposed solutions, the patients' quality of life has been improved.

Haruka and colleagues [67] proposed a low-cost GPS tracking system for Alzheimer's patients that could track and locate these patients in real-time. The main focus of this study is one of the uses of health monitoring technology by caregivers of Alzheimer's patients. According to reports released by this research group, some dementia symptoms are unfortunately often seen by doctors and patients' families as signs of aging, so care does not start on time. It is also reported that the number of Alzheimer's patients in India in 2010 was approximately 3.7 million, which is expected to increase to 6 million by 2040. Therefore, it is necessary to adopt methods for accurate identification of these patients in the early stages of the disease.

Haruka and colleagues [68] have introduced a system using the machine-to-machine (M2M)/IoT platform to help Alzheimer's patients who are living alone. For this purpose, the researchers installed sensors in the patients' homes that can detect early signs of behavior change and mental disorder and dementia. Data from these sensors are also used to analyze the behaviors of Alzheimer's patients. Also, a questionnaire was developed and distributed among these patients. The results of these questionnaires were used as data for their characteristics. Then, analyzing this data and comparing it with the data collected by the sensors can determine the presence or absence of dementia.

Enshaeifar and colleagues [69] developed a localized product for controlling the activities of Alzheimer's patients using real-time image processing that is capable of monitoring patient activities and managing emergencies. The purpose of this product is to help patients to maintain their independence while reducing the demand for the physical presence of physicians and health professionals to check their health status. In this study, a safety assistant was called the Path Tracking and Fall Detection System (PTFaD) wandering tracking system, a smartphone-based system that can monitor patients in and out of the home and make alert notifications in emergency time to medical services. To effectively detect Alzheimer's patients, PTFaD uses a smartphone camera to take pictures of the patient while s/he is moving. The photos will be delivered to the cloud computing system along with the time they were taken and the GPS location information. If needed, doctors or carers can use that data to quickly

```
let server = net.createServer(function (connection) {
    console.log('Client connected');
    connection.on('data', function (data) {
        cdata = data.toString();
        parts = cdata.split(";");
        imei_number = parts[1];
        dbConnecct();
        if (!empty(imei_number)) {
            if (checkDeviceRegistered(imei_number)) { // Checking the device already registered in DB
            // True
            //Do something if device already registered
            logDeviceConnection(imei_number); // Save the request in the file
        } else {
            // False
            //If device is not registerd this code will execite
            registerDevice(imei_number); // Registering/storing device information in System/DB
            logNewDevice(imei_number); // Save the new device connection is file
            notifySystemAdmin(imei_number); // Notifying sytem admin about device connection
        }
            connection.write("Device connection establised");
        } else {
            connection.write("Device did not recognised");
        }
    });
        connection.on('end', function () {
            console.log('Client disconnected');
        });
});
server.timeout =0;
server.listen(port, function () {
    console.log('Server is now listening on ' + port);
});
server.on('error', function (err) { console.log(err);
});
function dbConnecct() {
    con.connect(function (err) {
        if (err) throw err;
    });
}
function checkDeviceRegistered(imei_number) {
    con.query("SELECT * FROM devices WHERE imei=" + imei_number, function (err, result, fields) {
        if (err) throw err;
        if (result) {
            return true;
        } else
            return false;
    });
}
function logDeviceConnection() {
}
function registerDevice() { // Registering/storing device information in System/DB
}
function logNewDevice() { // Save the new device connection is file
}
```

Algorithm 1: The WebSocket connection function

find a way to help the patient. The study also suggested a method for detecting the time when a patient falls; in that case, a message would be sent to the caregiver or physician, and if the response was not received within the specified timeframe from this caregiver or physician, then a message would be sent to the emergency departments in the city of the emergency health center of a hospital.

Enshaeifar and colleagues [69] investigated a technical design called Integrated Health Management Technology (TIHM). TIHM generates notifications about patients' health

```
<?php
$data = json_decode(file_get_contents("php://input"));
if (isset($data) && !empty($data)){
   deviceConnection($data);
}
echo json_encode(array('success' =>1, 'response' => 'Device connection establised'));
function deviceConnection($data)
{
   $imei = isset($data['imei']) ? $data['imei'] : ''; // Getting the imei number from request
   if (!empty($imei)) {
      if ($this->checkDeviceRegistered($imei)) { // Checking the device already registered in DB
         // True
         //Do something if device already registered
         $this->logDeviceConnection($imei); // Save the request in the file
      } else { // False
         //If device is not registerd this code will execite
         $this->registerDevice($imei); // Registering/storing device information in System/DB
         $this->logNewDevice($imei); // Save the new device connection is file
         $this->notifySystemAdmin($imei); // Notifying sytem admin about device connection
         }
         echo json_encode(array('success' =>1, 'response' => 'Device connection establised'));
   } else {
      echo json_encode(array('status' =>0, 'response' => 'Device did not recognised'));
   }
}
function checkDeviceRegistered($imei)
{
   $res = $this->db->from('devices')->where(array('imei' => $imei))->get()->row(); // Checking/Fetching in DB
   if (!empty($res))
      return false;
   else
      return true;
}
function registerDevice($imei)
{
   $this->db->insert('devices', array('imei' => $imei)); // Adding/Inserting new entery to DB
}
```

Algorithm 2: HTTP RestAPI for device connection

using the Internet of Things, IoT devices and solutions and interoperable standards, a set of machine learning algorithms, and data analysis. This information is monitored continually by a team of physicians at the healthcare center who make appropriate decisions about how best to care for Alzheimer's patients, based on the data collected and the warnings generated. This research discusses the technical design of TIHM and explains why the combination of patient-centered design and human experience should be an integral part of technology design.

5. Proposed Research Methodology

The IoT instead of existing security challenges and privacy issues has played a vital role in a different area of daily life, such as transportation, smart homes, smart agriculture, smart grid stations, and healthcare systems. In this research paper, the proposed solution is for mentally ill persons. This system as shown in Figure 3 will help to monitor activities of these persons, health conditions, and movements and will be responded as quickly as possible in case of emergency by concerned hospitals.

In the above system, different types of devices are used like sensors, smart stickers on cloths of those peoples, smartwatches, smart cameras, and smartphone apps. As the heterogeneous devices are being used so that all have different ways of communication, some of them are using Message Queuing Telemetry Transport (MQTT), HyperText Transfer Protocol (HTTP), or WebSockets for fetching the required information. All the methods are developed in NodeJS for communication between these IoT devices. The smartwatches are used for the fetching of heartbeat blood pressure, temperature level, and diabetes level information into the proposed system. Based on the defined thrash hold for each mentioned disease, the alert will be generated to concerned persons against the patient's identification number and name.

The Nginx web server is used as a reverse proxy for the NodeJS program which communicates locally at the Ubuntu server for the security and protection of user privacy information. This web server is used for the public face domain

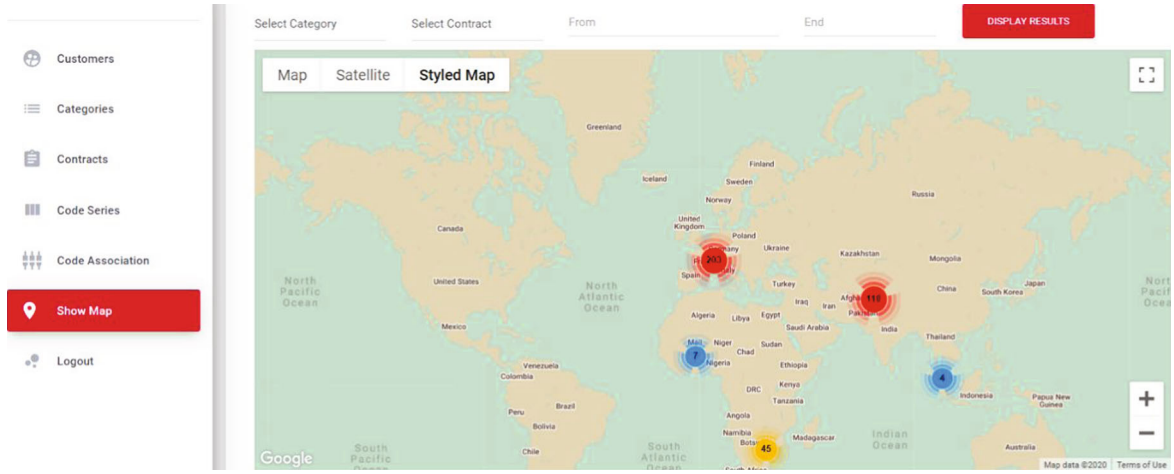


Figure 4: Live locations of Alzheimer patients.

also for the management dashboard and showing publically viewable information. The MySQL database server is used for storing the information of users for the data analysis regarding mental issues and disease records for future use. For real-time data storage, mongoDB has been used and this server is hosted at a cloud.

5.1. Method for IoT Device Communication Protocols. As the various devices are in use for tracking and monitoring the healthcare of Alzheimer patients at the same time, these are supporting different types of communication protocols. To fulfill those requirements, the communication method for devices has been developed in NodeJS for the MQTT, HTTPS, and WebSockets protocols, as it is depicted in Algorithm 1.

This method will first look into the communication method or request type from devices. In that condition, the communication protocol will be selected from this method, which can be MQTT, HTTP, or WebSocket protocol for gathering the information from Alzheimer monitoring IoT devices that are connected with this network via WiFi or data network from respective country cellular network company. As in the above function, we have developed a method for the WebSocket connection with sensors or with other IoT devices. This method has been created for database connection and store required information into it. For the security of patient's information and to avoid fake devices, this method will look into the database also for the registration of these devices. It is providing the functionality of data analysis and activities of patients to the system administrators or concerned authorities for the development of more facilities for them. The same type of method is also developed for the MQTT protocol-supported device for communication.

In Algorithm 2, the RestAPI method is developed for the communication over HTTP-supported devices and smartphone apps. This function also creates a database connection and looks for registered devices and apps. This function mainly proposes it to provide the service of transportation to Alzheimer patients as they need it. And it will be used for the delivery of medical facilities to them at the doorstep, and they can be registered via this app with simple steps. These two methods are described here just for the proof of concept to our improved solution for these patients, and some information has not been shared here due to the privacy of users.

6. Results and Discussions

In this paper, the different types of sensors, smartwatches, actuators, and IoT applications are used for the monitoring movement of Alzheimer patients. The various types of IoT devices are used for the health monitoring of these patients for the accuracy of information regarding their health. Due to the COVID-19 pandemic, it is very hard to take care of Alzheimer patients because there is a need for social distance. The communication between these devices has been made secure by implementing existing IoT protocols under the umbrella of WoT. To overcome the unavailability of authentication in WebSocket, it is managed via programming for each device authentication token that has been stored in the database at a cloud. This is implemented for the security of a patient's privacy and data collected from these IoT devices. By following, these authors [70] recommended a method against device forging at the physical layer and security of data at transit or rest. Due to this system, it will be easy for the health department of any country to locate these patients or any other patients with different diseases. The better service can be provided via ambulant or doctor advise remotely. As already mentioned, security or privacy issues in their survey by authors [4] have been fixed in this research paper.

6.1. Tracking and Management Dashboard. The critical part of any application or Alzheimer patients is monitoring, as the activities are monitoring and stored in the database that can be used for analysis and develop a good treatment solution for these patients. At the same time, these improved services can be utilized for the old people at the doorstep with the help of IoT. By this, any government body for health services or any private company can get information regarding those patient's area which is affected by this disease, so that medical facilities can be provided in that location and more transportation in that area. The tracking of vehicles and

Figure 5: List of events processed.

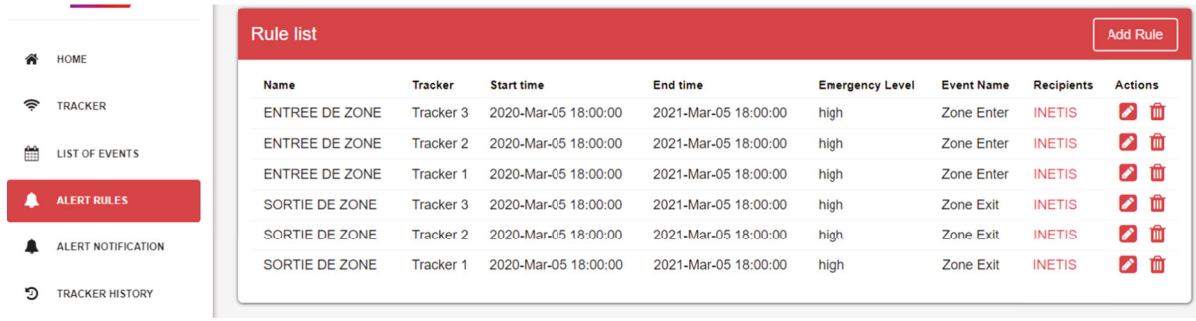


Figure 6: Alert rules defined.

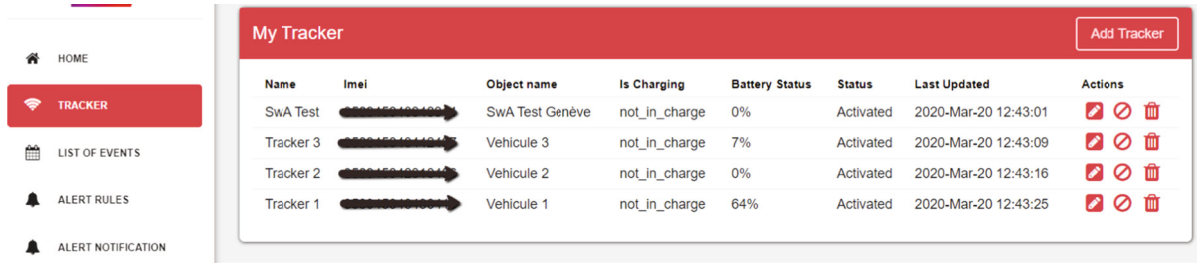


Figure 7: Status devices at patient's area.

locations of these patients is shown in Figure 4. These locations are dummy currently as per our testing of the developed application with the help of NodeJS and other tools.

The test of different events has been done on real Alzheimer patients as depicted in Figure 5. These events are related to the ambulance services to them in an emergency and also in normal routine for their medical checkup. These events are processed on the bases of alerts generated by those sensors or smartwatches. But we have not displayed those events here due to the privacy protection of patients suffering from this disease. The resulting ratio of this improved solution is 95% accurate in the shape of alerts generated by installed sensors, stickers on cloths, or smartwatches.

The list of alert rules has been designed which can be defined by the administrator. These rules are based on real issues faced by Alzheimer patients and categorized on that condition. Few rules are shown in Figure 6. These are categorized with added trackers as per respective area and medical service providers.

Another main section of this solution is the status of devices installed at the client-side. We have tried our best to get useful information as much as possible for the betterment of these patients and easiness for them. In this status section, the device name, IMEI number, charging status, battery status, and active or not are as shown in Figure 7.

7. Conclusions

One of the greatest human problems is the development of a variety of diseases, including Alzheimer's, in the old age of people all around the world. In the last decade, for unknown reasons, the number of people with this disease has been rising worldwide. Numerous studies have been conducted by researchers around the world, but the real reasons which cause the disease are still unknown. On the other hand, despite the pharmacological and nonpharmacological treatments which are suggested by physicians and researchers to help these patients, virtually none of those treatments can completely prevent the disease from progressing. Since caring for these patients is needed at all times (24 hours per day), it takes a lot of patience for the family and ultimately is economically costly. On the other hand, the Internet of Things (IoT) has rapidly received popularity throughout the world in the last decade. This technology can continuously monitor Alzheimer's patient's behaviors at home and abroad and inform the geographical location and occurrence of the accident and critical conditions to family members and healthcare personnel. In this paper, we have proposed a novel solution for tracking activities and monitoring the health condition of patients with the help of IoT devices. This

devices' communication has been secured with the recommended standards for the MQTT, WebSocket, and HTTP for the IoT application. The data has been collected from different types of devices and sensors to get accurate information regarding the patient's health condition. To keep in view the security and privacy issues of data, it is secured during transit and at rest. With this solution, transportation and medical facilities can be provided to them. The state-of-the-art administration dashboard has been developed for monitoring, device status alert generation rules, and live tracking of patients and vehicles for them with the help of Google maps. Our proposed system has a 95% accuracy ratio for the emergency alerts and condition of patients regarding their blood pressure, heartbeat, or sugar level. This will help health departments of any country to provide health facilities rapidly and perfectly to their patients.

Data Availability

All are simulation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. A. Zaidan, B. B. Zaidan, M. Y. Qahtan et al., "A survey on communication components for IoT-based technologies in smart homes," *Telecommunication Systems*, vol. 69, no. 1, pp. 1–25, 2018.

[2] A. A. Zaidan, B. B. Zaidan, M. A. Alsalem, O. S. Albahri, A. S. Albahri, and M. Y. Qahtan, "Multi-agent learning neural network and Bayesian model for real-time IoT skin detectors: a new evaluation and benchmarking methodology," *Neural Computing and Applications*, vol. 32, pp. 8315–8366, 2020.

[3] P. P. Ray, D. Dash, and N. Kumar, "Sensors for internet of medical things: state-of-the-art, security and privacy issues, challenges and future directions," *Computer Communications*, vol. 160, pp. 111–131, 2020.

[4] A. S. Ashfaq, S. G. Nitin, D. M. K. Abid, and T. A. Husain, "Android and internet of things (IoT) based Alzheimer care/rehabilitation system to monitor and progress patient health condition," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 3, 2017.

[5] M. Talal, A. A. Zaidan, B. B. Zaidan et al., "Smart home-based IoT for real-time and secure remote health monitoring of triage and priority system using body sensors: multi-driven systematic review," *Journal of Medical Systems*, vol. 43, no. 3, p. 42, 2019.

[6] A. Spector, M. Orrell, A. Schepers, and N. Shanahan, "A systematic review of 'knowledge of dementia' outcome measures," *Ageing Research Reviews*, vol. 11, no. 1, pp. 67–77, 2012.

[7] L. Dieckmann, S. H. Zarit, J. M. Zarit, and M. Gatz, "The Alzheimer's disease knowledge test," *The Gerontologist*, vol. 28, no. 3, pp. 402–408, 1988.

[8] J. J. Barrett, W. E. Haley, L. E. Harrell, and R. E. Powers, "Knowledge about Alzheimer disease among primary care physicians, psychologists, nurses, and social workers," *Alzheimer Disease and Associated Disorders*, vol. 11, no. 2, pp. 99–106, 1997.

[9] C. Gilleard and F. Groom, "A study of two dementia quizzes," *British Journal of Clinical Psychology*, vol. 33, no. 4, pp. 529–534, 1994.

[10] D. Kuhn, S. P. King, and B. R. Fulton, "Development of the knowledge about memory loss and care (KAML-C) test," *American Journal of Alzheimer's Disease and Other Dementias*, vol. 20, no. 1, pp. 41–49, 2005.

[11] B. D. Carpenter, S. Balsis, P. G. Otilingam, P. K. Hanson, and M. Gatz, "The Alzheimer's disease knowledge scale: development and psychometric properties," *Gerontologist*, vol. 49, no. 2, pp. 236–247, 2009.

[12] W. Tang, K. Kannaley, D. B. Friedman et al., "Concern about developing Alzheimer's disease or dementia and intention to be screened: an analysis of national survey data," *Archives of Gerontology and Geriatrics*, vol. 71, pp. 43–49, 2017.

[13] N. J. Karlin and M. Dalley, "Alzheimer's disease knowledge: a comparison study," *Journal of Clinical Geropsychology*, vol. 4, pp. 211–218, 1998.

[14] D. Brian, "The Alzheimer's disease knowledge scale: development and psychometric properties," *The Gerontologist*, vol. 49, no. 2, pp. 236–247, 2009.

[15] V. Jagadeeswari, V. Subramaniyaswamy, R. Logesh, and V. Vijayakumar, "A study on medical internet of things and big data in the personalized healthcare system," *Health Information Science and Systems*, vol. 6, no. 14, pp. 1–20, 2018.

[16] R. Srimathi, D. Mukul, S. Robin, S. Shikhar, V. Gomathi, and G. Kanimozhi, "Detection of brain abnormalities using internet of things," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 18, pp. 2003–2008, 2018.

[17] W. Smyth, E. Fielding, E. Beattie et al., "A survey-based study of knowledge of Alzheimer's disease among health care staff," *BMC Geriatrics*, vol. 13, no. 1, pp. 1–8, 2013.

[18] J. Hoe, M. Orrell, G. Charlesworth et al., "Support at Home: Interventions to Enhance Life in Dementia (SHIELD) – evidence, development, and evaluation of complex interventions," *Program Grants for Applied Research*, vol. 5, no. 5, pp. 1–184, 2017.

[19] Y. Wang, L. D. Xiao, Y. Luo, S. Y. Xiao, C. Whitehead, and O. Davies, "Community health professionals' dementia knowledge, attitudes, and care approach: a cross-sectional survey in Changsha, China," *BMC Geriatrics*, vol. 18, no. 1, p. 122, 2018.

[20] A. A. Shaikh, N. S. Gupta, A. D. M. Khan, and H. T. Artist, "Android and internet of things (IoT) based Alzheimer care/rehabilitation system to monitor and progress patient health condition," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 3, 2017.

[21] H. Ishii, K. Kimino, M. Aljehani, N. Ohe, and M. Inoue, "An early detection system for dementia using the M2 M/IoT platform," *Procedia Computer Science*, vol. 96, pp. 1332–1340, 2016.

[22] B. Carmi, "How IoT solutions help Alzheimer's patients stay independent," November 2019, https://blog.aeris.com/neo/how-iot-solutions-help-alzheimers-patients-stay-independent.

[23] N. Khachiyants, D. Trinkle, S. J. Son, and K. Y. Kim, "Sundown syndrome in persons with dementia: An update," *Psychiatry Investigation*, vol. 8, no. 4, pp. 275–287, 2011.

[24] Tuck et al., "Dementia and sleeping disorders," November 2019, https://www.tuck.com/dementia/#dementia_and_sleep_apnea.

[25] Z. H. K. Chong, Y. X. Tee, L. J. Toh et al., "Predicting potential Alzheimer medical condition in elderly using IOT sensors-

case study," in *IRC Conference on Science Engineering, and Technology*, Singapore, August 2017.

[26] Alzheimer's Association, "Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 15, pp. 321–387, 2019.

[27] R. C. Green, L. A. Cupples, R. Go et al., "Risk of dementia among white and African American relatives of patients with Alzheimer disease," *JAMA*, vol. 287, no. 3, pp. 329–336, 2002.

[28] N. T. Lautenschlager, L. A. Cupples, V. S. Rao et al., "Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: what is in store for the oldest old?," *Neurology*, vol. 46, no. 3, pp. 641–650, 1996.

[29] Alzheimer's Disease Education and Referral Center, *Alzheimer's disease genetics: fact sheet*, National Institutes of Health, Bethesda, MD, USA, 2011.

[30] J. Raber, Y. Huang, and J. W. Ashford, "ApoE genotype accounts for the vast majority of AD risk and AD pathology," *Neurobiology of Aging*, vol. 25, no. 5, pp. 641–650, 2004.

[31] A. M. Saunders, W. J. Strittmatter, D. Schmechel et al., "Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease," *Neurology*, vol. 43, no. 8, pp. 1467–1472, 1993.

[32] L. A. Farrer, L. A. Cupples, J. L. Haines, and B. Hyman, "Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease," *JAMA*, vol. 278, no. 16, pp. 1349–1356, 1997.

[33] K. J. Anstey, C. von Sanden, A. Salim, and R. O'Kearney, "Smoking as a risk factor for dementia and cognitive decline: a meta-analysis of prospective studies," *American Journal of Epidemiology*, vol. 166, no. 4, pp. 367–378, 2007.

[34] C. R. Jack Jr., M. S. Albert, D. S. Knopman et al., "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's Dement*, vol. 7, no. 3, pp. 257–262, 2011.

[35] G. M. McKhann, D. S. Knopman, H. Chertkow et al., "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 263–269, 2011.

[36] G. M. McKhann, M. S. Albert, and R. A. Sperling, "Changing diagnostic concepts of Alzheimer's disease," in *Alzheimer's disease — Modernizing concept, biological diagnosis, and therapy*, H. Hampel and M. C. Carrillo, Eds., pp. 115–121, Karger, Basel, Switzerland, 2012.

[37] C. Groot, A. M. Hooghiemstra, P. G. H. M. Raijmakers et al., "The effect of physical activity on cognitive function in patients with dementia: a meta-analysis of randomized control trials," *Ageing Research Reviews*, vol. 25, pp. 13–23, 2016.

[38] N. Farina, J. Rusted, and N. Tabet, "The effect of exercise interventions on cognitive outcome in Alzheimer's disease: a systematic review," *International Psychogeriatrics*, vol. 26, no. 1, pp. 9–18, 2014.

[39] M. Brasure, P. Desai, H. Davila et al., "Physical activity interventions in preventing cognitive decline and Alzheimer-type dementia," *Annals of Internal Medicine*, vol. 168, no. 1, pp. 30–38, 2018.

[40] J. De Reuck, C. A. Maurage, V. Deramecourt et al., "Aging and cerebrovascular lesions in pure and in mixed neurodegenerative and vascular dementia brains: a neuropathological study," *Folia Neuropathologica*, vol. 56, no. 2, pp. 81–87, 2018.

[41] *National Institute on Aging. What are frontotemporal disorders?*December 2018, https://www.nia.nih.gov/health/what-are-frontotemporal-disorders.

[42] M. Cherubini and R. Wade-Martins, "Convergent pathways in Parkinson's disease," *Cell and Tissue Research*, vol. 373, no. 1, pp. 79–90, 2018.

[43] I. Stojkovska, D. Krainc, and J. R. Mazzulli, "Molecular mechanisms of α-synuclein and GBA1 in Parkinson's disease," *Cell and Tissue Research*, vol. 373, no. 1, pp. 51–60, 2018.

[44] National Down Syndrome Society, *Aging and Down syndrome: A health and well-being guidebook*National Down Syndrome Society, New York, NY, USA.

[45] L. E. Hebert, J. L. Bienias, N. T. Aggarwal et al., "Change in risk of Alzheimer disease over time," *Neurology*, vol. 75, no. 9, pp. 786–791, 2010.

[46] L. E. Hebert, J. Weuve, P. A. Scherr, and D. A. Evans, "Alzheimer disease in the United States (2010-2050) estimated using the 2010 census," *Neurology*, vol. 80, no. 19, pp. 1778–1783, 2013.

[47] A. M. Saunders, W. J. Strittmatter, D. Schmechel et al., "Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease," *Neurology*, vol. 43, no. 8, pp. 1467–1472, 1993.

[48] L. A. Farrer, L. A. Cupples, J. L. Haines et al., "Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer Disease," *JAMA*, vol. 278, no. 16, pp. 1349–1356, 1997.

[49] R. Mayeux, M. Sano, J. Chen, T. Tatemichi, and Y. Stern, "Risk of dementia in first-degree relatives of patients with Alzheimer's disease and related disorders," *Archives of Neurology*, vol. 48, no. 3, pp. 269–273, 1991.

[50] M. Tang, Y. Stern, K. Marder et al., "The APOE-e4 allele and the risk of Alzheimer's disease among African Americans, whites, and Hispanics," *JAMA*, vol. 279, no. 10, pp. 751–755, 1998.

[51] H. C. Hendrie, J. Murrell, O. Baiyewu et al., "APOE $\varepsilon 4$ and the risk for Alzheimer disease and cognitive decline in African Americans and Yoruba," *International Psychogeriatrics*, vol. 26, no. 6, pp. 977–985, 2014.

[52] C. Reitz, G. Jun, A. Naj et al., "Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E $\epsilon 4$, and the risk of late-onset Alzheimer disease in African Americans," *JAMA*, vol. 309, no. 14, pp. 1483–1492, 2013.

[53] F. J. Wolters, S. J. van der Lee, P. J. Koudstaal et al., "Parental family history of dementia in relation to subclinical brain disease and dementia risk," *Neurology*, vol. 88, no. 17, pp. 1642–1649, 2017.

[54] J. E. Maye, R. A. Betensky, C. M. Gidicsin et al., "Maternal dementia age at onset in relation to amyloid burden in non-demented elderly offspring," *Neurobiology of Aging*, vol. 40, pp. 61–67, 2016.

[55] Institute of Medicine, *Cognitive Aging: Progress in Understanding and Opportunity for Action*, The National Academies Press, Washington, DC, USA, 2015.

[56] M. Rusanen, M. Kivipelto, C. P. Quesenberry, J. Zhou, and R. A. Whitmer, "Heavy smoking in midlife and long-term risk of Alzheimer's disease and vascular dementia," *Archives of Internal Medicine*, vol. 171, no. 4, pp. 333–339, 2012.

[57] D. Choi, S. Choi, and S. M. Park, "Effect of smoking cessation on the risk of dementia: a longitudinal study," *Annals of*

Clinical and Translational Neurology, vol. 5, no. 10, pp. 1192–1199, 2018.

[58] D. Miorandi, S. Sicari, F. de Pellegrini, and I. Chlamtac, "Internet of things: vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[59] G. Santucci, "From the internet of data to the internet of thing," in *Paper for the International Conference on Future Trends of the Internet*, vol. 28, 2009.

[60] I. G. Smith, O. Vermesan, P. Friess, and A. Furness, "Europe's IoT Strategic Research Agenda 2012 article," in *The Internet of Things 2012 New Horizons Chapter: 2*, G . Ian, Ed., pp. 22–117, 2012.

[61] American Society of Heating, *Refrigerating and Air-Conditioning Engineers, Fundamentals Handbook*, ASHRAE, Atlanta, GA, 2013.

[62] A. M. Vilamovska, E. Hattziandreu, R. Schindler, C. Van Oranje, H. De Vries, and J. Krapelse, *RFID Application in Healthcare – Scoping and Identifying Areas for RFID Deployment in Healthcare Delivery*, RAND Europe, 2009.

[63] N. David, *How the Internet of Things Is Revolutionizing Healthcare, Freescale Semiconductors?*, 2013.

[64] S. Li, L. D. Xu, and S. Zhao, "The internet of things: a survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.

[65] E. Thierry and D. Jules, "IoT-Enabled Health Monitoring and Assistive Systems for in Place Aging Dementia Patient and Elderly," *Internet of Things (IoT) for Automated and Smart Applications*, 2019.

[66] B. Sindhu, N. R. Shravani, K. J. Pooja Rao, H. Y. Yashaswini, and C. G. Nayana, "Real-time health and security monitoring device for dementia affected elders," *International Journal of Engineering Research & Technology*, vol. 6, no. 13, pp. 1–5, 2018.

[67] I. Haruka, K. Keisuke, A. Maher, O. Nobuhiro, and I. Masahiro, *20th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, Procedia Computer Science, 2016.

[68] D. K. Shende, M. S. Madrewar, M. S. Bhongade, and M. S. Dugade, "Dementia patient activity monitoring and fall detection using IoT for elderly," *International Journal of Trend in Scientific Research and Development*, vol. 3, no. 4, pp. 363–367, 2019.

[69] S. Enshaeifar, P. Barnaghi, S. Skillman et al., "The internet of things for dementia care," *IEEE Internet Computing*, vol. 22, no. 1, pp. 8–17, 2018.

[70] K. B. Jalbani, A. H. Jalbani, and S. S. Soomro, "IoT security," in *Industrial Internet of Things and Cyber-Physical Systems: Transforming the Conventional to Digital*, pp. 98–118, IGI Global, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 9280198, 9 pages
https://doi.org/10.1155/2020/9280198

WILEY | Hindawi

Research Article

A New Computing Paradigm for Off-Grid Direction of Arrival Estimation Using Compressive Sensing

Hamid Ali Mirza,[1] **Laeeq Aslam**,[1,2] **Muhammad Asif Zahoor Raja**,[3,4] **Naveed Ishtiaq Chaudhary**,[1] **Ijaz Mansoor Qureshi**,[5] **and Aqdas Naveed Malik**[1]

[1]*Department of Electrical Engineering, International Islamic University, Islamabad 44000, Pakistan*
[2]*School of Engineering & Applied Sciences, ISRA University, Islamabad 44000, Pakistan*
[3]*Future Technology Research Center, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliu, Yunlin 64002, Taiwan*
[4]*Department of Computer and Electrical Engineering, COMSATS University Islamabad, Attock Campus, Attock, Pakistan*
[5]*Department of Electrical Engineering, Air University Islamabad, 44000, Pakistan*

Correspondence should be addressed to Hamid Ali Mirza; hmdmirza@gmail.com

Received 24 February 2020; Revised 10 July 2020; Accepted 23 July 2020; Published 25 August 2020

Academic Editor: Hing Cheung So

In this paper, a method for solving grid mismatch or off-grid target is presented for direction of arrival (DOA) estimation problem using compressive sensing (CS) technique. Location of the sources are at few angles as compare to the entire angle domain, i.e., spatially sparse sources, and their location can be estimated using CS methods with ability of achieving super resolution and estimation with a smaller number of samples. Due to grid mismatch in CS techniques, the source energy is distributed among the adjacent grids. Therefore, a fitness function is introduced which is based on the difference of the source energy among the adjacent grids. This function provides the best discretization value for the grid through iterative grid refinement. The effectiveness of the proposed scheme is verified through extensive simulations for different number of sources.

1. Introduction

The direction of arrival (DOA) has been under investigation for a long time [1–3], and with the advancement in wireless communications, it has applications in a variety of fields like radar, acoustic signal processing, medical imaging, and seismology. The goal of DOA is to estimate the location of the closely spaced sources in the presence of a noise. It is common to use antenna arrays with different structure [4, 5], and there are many algorithms for estimation of the DOA [6, 7]. The performance of these algorithms depends upon the number of the samples and signal-to-noise ratio (SNR). These algorithms can be divided in to three main categories [8], conventional beamforming, subspace techniques, and maximum likelihood technique. The most notable algorithms used for DOA are multiple signal classification (MUSIC) algorithm, estimation of signal parameter via a rotational variant technique (ESPIRIT), and CAPON. These algorithms require the sampling to be on the Nyquist rate, i.e., sampling frequency should be at least twice the highest frequency present in the signal.

Compressive sensing (CS) has gained a lot of attention over the years because of its ability to exploit the concept of sparsity [9, 10]. The CS has applications in the fields like radars [11], image reconstruction and restoration [12, 13], blind source separation [14, 15], beamforming, and source localization [16–18]. A signal can be reconstructed only from a small number of linear measurements if it is sparse in certain domain. It means that the information rate of the signal is much smaller than the suggested signal bandwidth. As most of the real-time signals are sparse, therefore, the signal can be reconstructed using a system of equations such that the smaller number of samples is used. It is important that it satisfies the restricted isometric property (RIP). A recursive

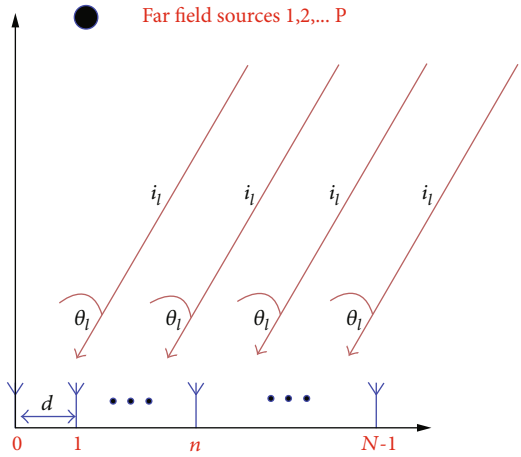


Figure 1: Signal model for uniform linear array.

weighted minimum norm with focal underdetermined system solver (FOCUSS) is used to achieve sparsity in the problem of source localization [19]. In [20], a hardware has been developed for spectral estimation using CS framework.

Normally, in the CS, we are required to find a sparse signal using an overcomplete dictionary. This concept can also be applied to source localization problem where spatial sparsity is exploited; the source is sparse in the spatial domain means that if it is not present at every angle, hence, the concept of CS can be applied due to the sparsity of source. The source localization can be done using single or multiple measurements to achieve super resolution. For multiple measurements, the computational complexity increases with the amount of data growing in a system. Different methods are present for solving multiple measurement problems.

One of the main issues with the source localization using CS is to define the resolution of the grid [21]. It is assumed that the location of the sources falls on the resolution of the grid defined. However, this is not always possible or practical. There may be a scenario when the location of the source does not coincide with the resolution of the grid defined. This creates a grid mismatch or off-grid targets. There are different methods available to solve this problem; one of them is iterative grid refinement. In iterative grid refinement, the grid resolution is changed until the dense grid mitigates the grid mismatch. One of the main drawbacks of using iterative grid refinement is that decreasing the grid resolution may not comply with the RIP condition. The other approach to resolve grid mismatch problem is off-grid sparse method [22]. In this process, initial grid resolution is still defined. The DOA of the sources is not restricted to the grid. A bias is added to the signal model using first order approximation of the manifold matrix. The new model may be nonconvex and difficult to solve. Iterative grid refinement is one of the most used methods. However, it requires a method to best select the discretization value for the grid. Therefore, in this paper, a method based on the distribution of the source energy due to grid mismatch is presented to calculate the discretization value for the grid. The main contributions of the paper are summarized as follows.

(i) A novel framework for solving grid mismatch or off-grid target is presented for DOA estimation problem using CS technique

(ii) A fitness function is introduced based on the difference of source energy between adjacent grids due to grid mismatch in DOA estimation

(iii) An approach for finding the best discretization value using the designed objective function is presented for iterative grid refinement

(iv) The proposed scheme is viably tested for multiple sources with different energy and spatial resolution-based scenarios in DOA estimation

The rest of the article is organized as follows. The mathematical background for DOA estimation using MUSIC algorithm is presented in Section 2, while in Section 3, compressive framework for DOA estimation is presented. In Section 4, the proposed algorithm is provided, and results along with necessary discussion are presented in Section 5. The concluding remarks are given in Section 6.

2. DOA Estimation

In this section, a general model for direction of arrival (DOA) estimation based on a subspace technique for DOA estimation is presented. Let us consider a uniform linear array (ULA) as shown in Figure 1 with N number of antenna elements, and the distance between the antenna element is $\lambda/2$, while there are P number of far field sources at different angles, θ_l. Then, the received signal at the m^{th} antenna element is given as

$$y_m(t) = \sum_{i=1}^{P} s_i(t) e^{j(m-1)kd \sin \theta_i}, \tag{1}$$

where s_i is the amplitude of the signal that is received. k and d are wave number and distance between the antenna elements.

Then, Equation (1) can be written as

$$\mathbf{y} = \mathbf{A}\mathbf{s}. \tag{2}$$

In the presence of noise, the received signal is updated as

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{n}, \tag{3}$$

where $\mathbf{n}$ is the Gaussian noise, $\mathbf{A}$ is the steering vector, and $\mathbf{y}$ is the received vector. As mentioned earlier, there are number of ways to solve Equation (3) like MUSIC and MVDR. If it satisfies the Nyquist sampling rate, the correlation of the received matrix is given as

$$\begin{aligned} \mathbf{R} &= E\left[\mathbf{y}\mathbf{y}^H\right], \\ \mathbf{R} &= E\left[(\mathbf{A}\mathbf{s} + \mathbf{n})(\mathbf{A}\mathbf{s} + \mathbf{n})^H\right], \end{aligned} \tag{4}$$

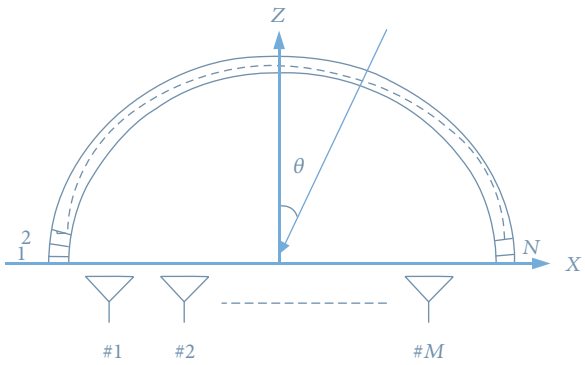


Figure 2: DOA model for CS.

which is

$$\mathbf{R} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_n. \tag{5}$$

Then, accordingly, the spatial spectrum is given as

$$p(\theta) = \frac{1}{a^H(\theta)\mathbf{e}_n\mathbf{e}_n^H a(\theta)}, \tag{6}$$

where $\mathbf{e}_n$ is the eigen vector orthogonal to the steering vector. The performance of the MUSIC algorithm degrades with the reduction in the number of the samples and presence of a noise. In the next section, we will look at the CS approach to the DOA estimation.

3. Compressive Sensing

The concept and mathematical development of compressive sensing for DOA estimation are briefly presented in this section. In CS framework, a sparse representation of the signal can be reconstructed only from a small number of samples. Let us consider a signal $\mathbf{s}$, a discrete signal which is sparse in certain domain $s\varepsilon\mathbb{C}^{N\times 1}$ and $\mathbf{y}$ is the received signal of dimension M such that $y\varepsilon\mathbb{C}^{M\times 1}$. Then, received signal without noise is given as

$$\mathbf{y} = \mathbf{A}\mathbf{s}. \tag{7}$$

Here, $\mathbf{A}$ is a sensing matrix of dimension $A\varepsilon\mathbb{C}^{M\times N}$, where $M \ll N$. We count the number of nonzeros elements in a signal $\mathbf{s}$, which is given by $\|\mathbf{s}\|_0$ also known as l_0—norm. This leads to a nonlinear programming (NP) hard problem. To solve NP problem, many approximation methods have been developed. One of the methods is to use l_1 or l_p relaxation. If the unknown signal $\mathbf{s}$ is considered sparse, then the optimization problem can be mathematically represented as

$$\mathbf{s} = \min \|\mathbf{s}\|_p s.t \mathbf{y} = \mathbf{A}\mathbf{s}, \tag{8}$$

considering $p = 0$, then the above equation will be an NP hard problem. While considering $p = 1$, we can recast it as an l_1—norm problem and solve it using Equation (9).

$$\mathbf{s} = \min \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda\|\mathbf{s}\|_p. \tag{9}$$

In practical scenario, there is always a noise. Now, if the received signal is contaminated with a noise $\mathbf{n}$, then, (7) is rewritten as

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{n}, \tag{10}$$

and the optimization problem becomes

$$\min \|\mathbf{s}\|_1 s.t \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 < \varepsilon, \tag{11}$$

where ε is a parameter that specifies how much noise is allowed. To formulate the problem in the CS framework, consider Figure 2 [23], where the spacing between the antenna elements is d which is $\lambda/2$. As the goal is to find the location of the sources, we consider a ULA with narrow band signal for K number of sources and M number of antenna elements. The received signal is given in (1). As seen in Figure 2, to cast this in the sparse representation problem, an overcomplete dictionary of array steering vector $\mathbf{A}$ is introduced, where $\mathbf{A} = [\theta_1, \theta_2, \cdots, \theta_N]$, N is the sampling of the grid. The N will be much higher than K. Therefore, the matrix $\mathbf{A}$ is given as

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{jkd\sin\theta_1} & e^{jk\sin\theta_2} & \cdots & e^{jk\sin\theta_N} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ e^{j(m-1)kd\sin\theta_1} & e^{j(m-1)kd\sin\theta_2} & \cdots & e^{j(m-1)kd\sin\theta_N} \end{bmatrix}. \tag{12}$$

One of the main issues with application of CS in the DOA problem is definition of the grid resolution. The resolution depends upon the sampling grid formulation, and the sampling grid is uniform. If the grid size is defined very fine, it increases the computational requirements. If the size of the sampling grid is large, then the resolution decreases, and close targets cannot be detected.

4. Proposed Methodology

In this section, we propose a methodology for the selection of discretization value for the grid. Considering Figure 3, the upper grid represents the resolution of the grid with discretization value of $r = \theta_{n+1} - \theta_n$, while θ_L represents the location of the source.

The first step is to estimate the vector $\mathbf{s}$ using the overcomplete dictionary defined for the iteration

$$\mathbf{s}^i = \min \left\|\mathbf{y} - \mathbf{A}^i\mathbf{s}\right\|_2^2 + \lambda\|\mathbf{s}\|_1. \tag{13}$$

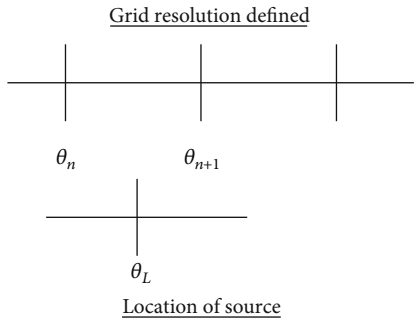


Figure 3: DOA model for CS.

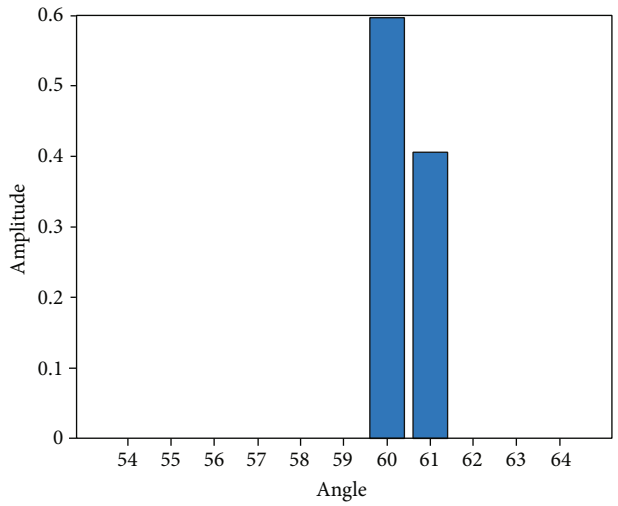


Figure 4: Ambiguity of single source location.

SetInitial grid resolution of 1°
  Calculate the over complete dictionary
  Select fitness function equal zero
  Estimate regularization term λ using GCV
While ($F_i \geq F_{i+1}$)
  Estimate $\mathbf{s}^i$ according to Equation (13)
  Calculate F_i according to Equation (15)
  Change grid resolution
  Calculate dictionary $\mathbf{A}^{i+1}$
  Estimate $\mathbf{s}^{i+1}$
  Calculate F_{i+1}
end while
Output:r^i

Algorithm 1: Proposed Algorithm.

Due to grid mismatch, the energy of the source is distributed among the adjacent grids as shown in Figure 4. The discretized grid resolution is 1° defined, and the location of the source is 60.4°. The energy of the source distributed among the adjacent grids is mathematically presented as

$$E_i = \left| E_{\theta_K} - E_{\theta_{K+1}} \right|. \quad (14)$$

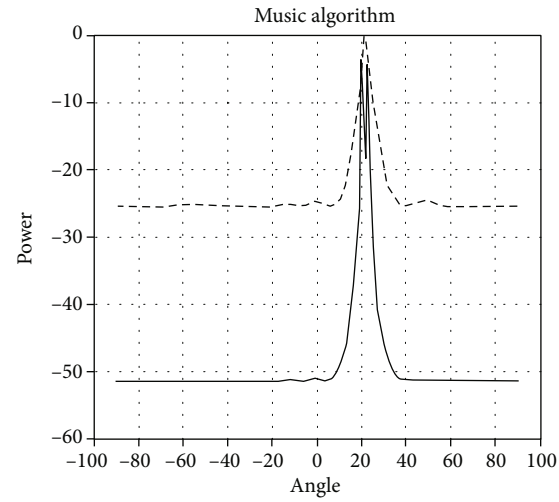


Figure 5: DOA estimation using MUSIC.

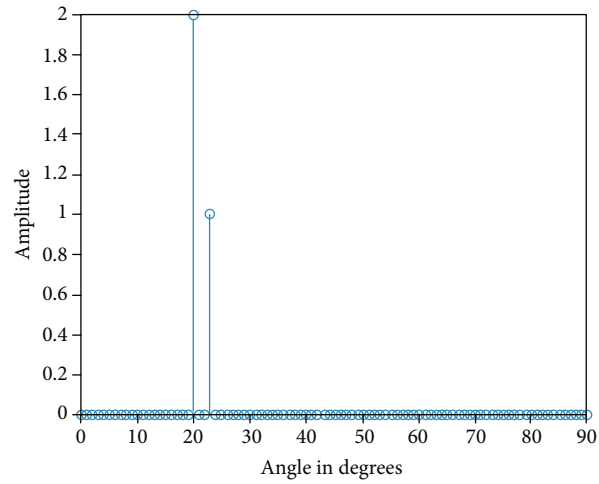


Figure 6: DOA estimation using CS.

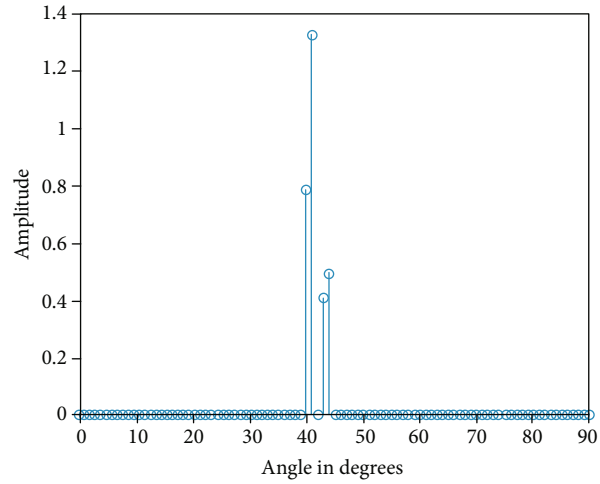


Figure 7: Ambiguity due to grid Mismatch Problem.

Table 1: Grid mismatch case for two sources.

Amplitude and location		Grid resolution 1°		Grid resolution 0.5°		Grid resolution 0.1°	
Amplitude	Source # 1 = 2	A_1	0.7830	A_1	1.999	A_1	1.998
		A_2	1.3208				
	Source # 2 = 1	A_3	0.4061	A_2	0.999	A_2	0.998
		A_4	0.4905				
Location	40.5°	θ_1	40°	θ_1	40.5°	θ_1	40.5°
		θ_2	41°				
	43.5°	θ_3	43°	θ_2	43.5°	θ_2	43.5°
		θ_4	44°				
Fitness		$F_1 = 0.6221$		$F_2 = 2.999$		$F_3 = 2.998$	

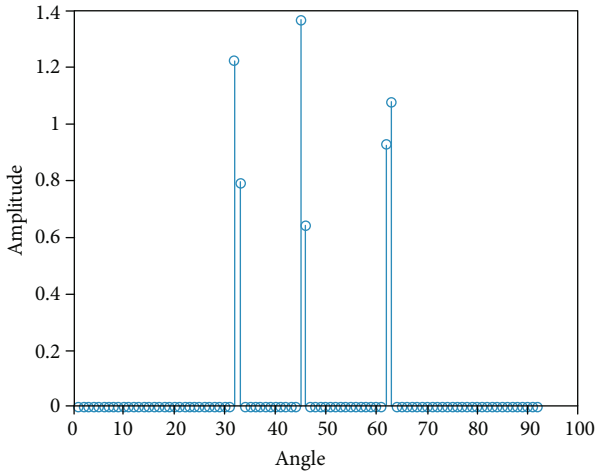


Figure 8: Three sources of grid mismatch case.

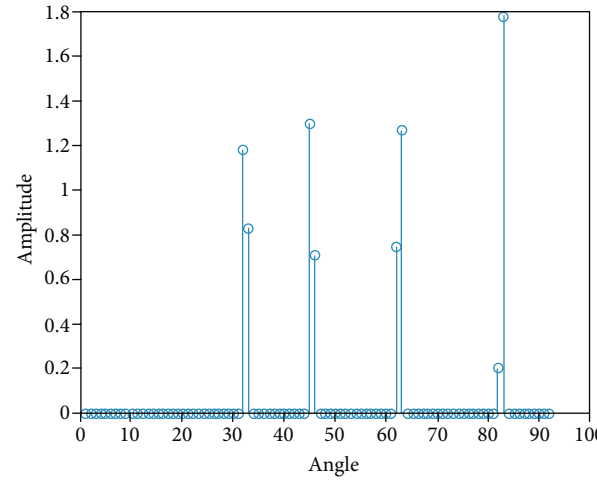


Figure 9: Four sources of grid mismatch problem.

Let i be the iteration index in which the discretization value is r^i. The peaks in the vector $\mathbf{s}$ are detected, and the difference is taken as in Equation (14).

$$F_i = E_i. \tag{15}$$

Then, the process is repeated with $i + 1^{\text{th}}$ iteration with finer grid discretized value and with dictionary defined with new discretization value. The fitness function is calculated as

$$F_{i+1} = E_{i+1}. \tag{16}$$

A termination criterion is defined asset as

$$F_i \geq F_{i+1}. \tag{17}$$

If it satisfies the termination criteria, then the discretization value in the i^{th} iteration is the best value for the grid to discretize at. For the case of multiple sources, Equation (14) can be generalized for sum of the difference of adjacent peaks and sum of individual peaks if there is no adjacent peak. As mentioned, i is the iteration number. In each iteration, the discretization value is reduced. It is selected by the user. In our simulations, we have selected a discretization value of 1, 0.5, 0.1, and 0.01. The main steps involved in the proposed algorithm are given as follows.

The regularization term λ in (13) plays important role for the accuracy of the solution and must be estimated. It is a compromise between finding a solution that is sparse as possible and has lower error as possible. Two methods for estimating the regularization term are L curve and generalized cross validation (GCV). In GCV, it is more convenient as compared to L curve where we must find the corner [24]. It can be computed using the following relations.

$$\text{GCV}(\lambda) = \frac{\|\mathbf{A}\mathbf{s}_\lambda - \mathbf{y}\|^2}{\text{trace}\left(I - \mathbf{A}\mathbf{A}^{\#}\right)^2}, \mathbf{A}^{\#} = \left(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I}\right)^{-1}\mathbf{A}^T, \mathbf{s}_\lambda = \mathbf{A}^{\#}\mathbf{y}. \tag{18}$$

The GCV estimate is variant of the above equation which is obtained by applying necessary calculation and results in GCV function [25]. This technique estimates λ by assuming that the optimum value of λ should be chosen to minimize GCV value.

Table 2: Grid mismatch case for two sources with finer resolution.

Amplitude and location		Grid resolution 1°		Grid resolution 0.5°		Grid resolution 0.1°		Grid resolution 0.01°	
Amplitude	Source # 1 = 2	A_1	0.9797	A_1	0.3381	A_1	1.9988	A_1	1.9916
		A_2	1.1367	A_2	1.6696				
	Source # 2 = 1	A_3	0.5718	A_3	0.5718	A_2	0.9987	A_2	0.991
		A_4	0.3312	A_4	0.3122				
Location	40.4°	θ_1	40°	θ_1	40°	θ_1	40.4°	θ_1	40.4°
		θ_2	41°	θ_2	40.5°				
	43.3°	θ_3	43°	θ_3	43°	θ_2	43.3°	θ_2	43.3°
		θ_4	44°	θ_4	43.5°				
Fitness		$F_1 = 0.4167$		$F_2 = 1.4934$		$F_3 = 2.9975$		$F_4 = 2.9827$	

Table 3: Grid mismatch case for four sources.

Amplitude and location		Grid resolution 1°		Grid resolution 0.5°		Grid resolution 0.1°		Grid resolution 0.01°	
Amplitude	Source # 1 = 2	A_1	1.1805	A_1	0.3951	A_1	1.9998	A_1	1.9991
		A_2	0.8286	A_2	1.6061				
	Source # 2 = 2	A_3	1.2950	A_3	0.7604	A_2	1.9997	A_2	1.9977
		A_4	0.7080	A_4	1.2409				
	Source # 3 = 2	A_5	0.7446	A_5	1.9054	A_3	1.9990	A_3	1.9923
		A_6	1.2663						
	Source # 4 = 2	A_7	0.2037	A_6	1.0556	A_4	1.9951	A_4	1.9631
		A_8	1.7768	A_7	0.9408				
Location	30.4°	θ_1	30°	θ_1	30°	θ_1	30.4°	θ_1	30.4°
		θ_2	31°	θ_2	30.5°				
	43.3°	θ_3	43°	θ_3	43°	θ_2	43.3°	θ_2	43.3°
		θ_4	44°	θ_4	43.5°				
	60.5°	θ_5	60°	θ_5	60.5°	θ_3	60.5°	θ_3	60.5°
		θ_6	61°						
	80.7°	θ_7	80°	θ_6	80.5°	θ_4	80.7°	θ_4	80.7°
		θ_8	81°	θ_7	81°				
Fitness		$F_1 = 3.0338$		$F_2 = 3.7118$		$F_3 = 7.9938$		$F_4 = 7.9523$	

5. Simulation Results

We start our simulations with MUSIC algorithm. In Figure 5, the results of the two sources based on DOA estimation are presented that are at an angle of 20° and°. The received signal SNR is 20 dB, and the number of antenna elements is 10. We consider the performance of the algorithm for different number of samples. The number of samples is 50 and 3. It is observed that as the number of the samples decreases, the detection performance degrades, and the targets are not distinguishable.

Next, we consider CS techniques for solving the DOA problem. Considering a single sample at $T = 1$, we create an overcomplete dictionary having a resolution of 1°. The location of the targets can be resolved by solving Equation (9) with the help of linear programming. We use convex optimization toolbox for solving this problem. For simplicity, we consider a noiseless case. It is shown in Figure 6 that two targets with amplitude of 2 and 1 on normal scale at location 20° and 23° are resolved.

Next, we consider a scenario in which the target locations are not aligned with the grid resolution. Considering two targets, one is at 40.5° with amplitude 2, and the other is at 43.5° with amplitude 1. In Figure 7, it is shown that the four targets are detected. This creates ambiguity about the location of the targets and the number of the targets. However, it is observed

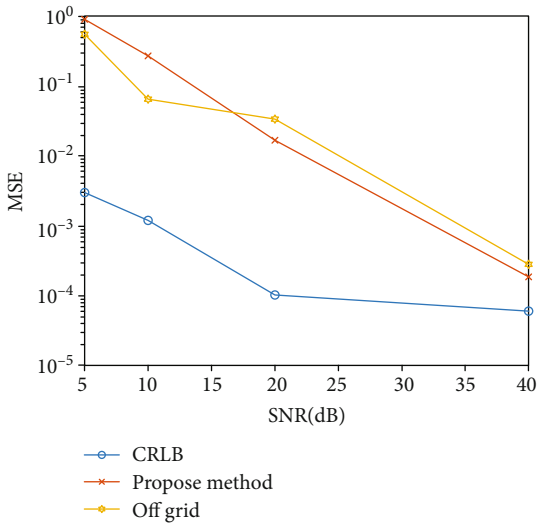


Figure 10: MSE comparison.

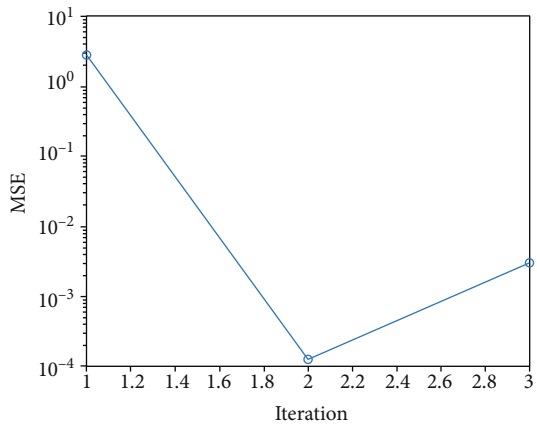


Figure 11: MSE vs. iteration.

that the amplitude of the received signal is distributed in the adjacent grid. To solve this, the proposed algorithm is used.

The fitness function is calculated for the pair of adjacent peaks. Then, the grid resolution is changed and again solved with the new dictionary elements, and the fitness function is calculated. The new fitness function calculated is compared with the previous one. If the fitness function increases, the steps are repeated; otherwise, we conclude that the optimized value has been reached, and the true location of the targets have been estimated as shown in Table 1.

Next, we consider a scenario for three and four numbers of sources in the DOA estimation problem. Initially, the grid resolution is taken to be 1°. It is assumed that the targets are not aligned with the grid resolution which again creates ambiguity about the number of the sources as shown in Figures 8 and 9. Table 1 shows the iterations for solving the grid mismatch problem using the proposed algorithm.

Next, we consider a different case, where finer resolution is required to detect the sources. Two sources are considered with locations at 40.4° and 43.3° and amplitude of 2 and 1 shown in Table 2. Similarly, four off-grid sources are considered in Table 3, where we present the results for a scenario of four sources at locations 30.4°, 43.3°, 60.5°, and 80.7°. The results show that the location of the sources is accurately estimated using the proposed algorithm.

As shown in the tables with the help of the fitness function, best discretization value for the grid is calculated. As mentioned, to address the basis mismatch, off-grid sparse method is also used. The signal model for the off-grid target with bias is given in [26, 27]. The minimization problem can be written as

$$\min_{s,\delta} \|\mathbf{y} - (\mathbf{A} + \mathbf{B}\Delta)\mathbf{s}\|_2^2 + \lambda\|\mathbf{s}\|, \tag{19}$$

where $\mathbf{B} = [b(\theta_1), \cdots, b(\theta_N)]$ and $b(\theta_{nk})$ is the derivative of $a(\theta_k)$ with respect to θ_{nk}. $\Delta = \text{diag}(\delta) = [\delta_1, \cdots, \delta_N]^T$. δ is the difference between the nearest grid point and the direction of the k^{th} signal.

In Figure 10, we consider two targets located at 30.5° and 60.5°. The regularization parameter for each SNR level is calculated using the GCV method. The mean square error (MSE) of the DOA estimation of the proposed method is compared with Cramer Rao lower bound (CRLB) for DOA estimation and off-grid method with bias. For a greater SNR-based scenario, the proposed algorithm is reasonable accurate. Additionally, the proposed approach is simpler and requires less computations.

In Figure 11, we compare the iterations for grid refinement with MSE. We consider two targets at 30.5° and 60.5°, where the grid is refined in each iteration. The resolution of the grid for each iteration is 1°, 0.5°, and 0.1°. It is observed that the MSE is the lowest for 0.5° grid resolution. Although the targets are detected for grid resolution of 0.1°, the MSE increases due to the RIP condition.

In Figure 12, we show the overall process flow structured diagram of the proposed study.

6. Conclusion

Compressive sensing is an interesting technique for finding the location of the sources using sensor array. However, the definition of the grid is a challenge. If the location of the target coincides with the grid resolution, it is detected, whereas if not, then the signal power of the source is distributed among the adjacent grids. This creates ambiguity about the location of the target. In this paper, we presented an iterative grid refinement mechanism based on a fitness function. The fitness function governs the extent of the grid refinement. When the maximum value of the fitness function is reached, we conclude further refinement of the grid is not required, and the ambiguity about the location of the target is removed. Thus, the best discretization value for the grid is calculated. In the future work, this technique can be applied for multiple input multiple output radar system [28].

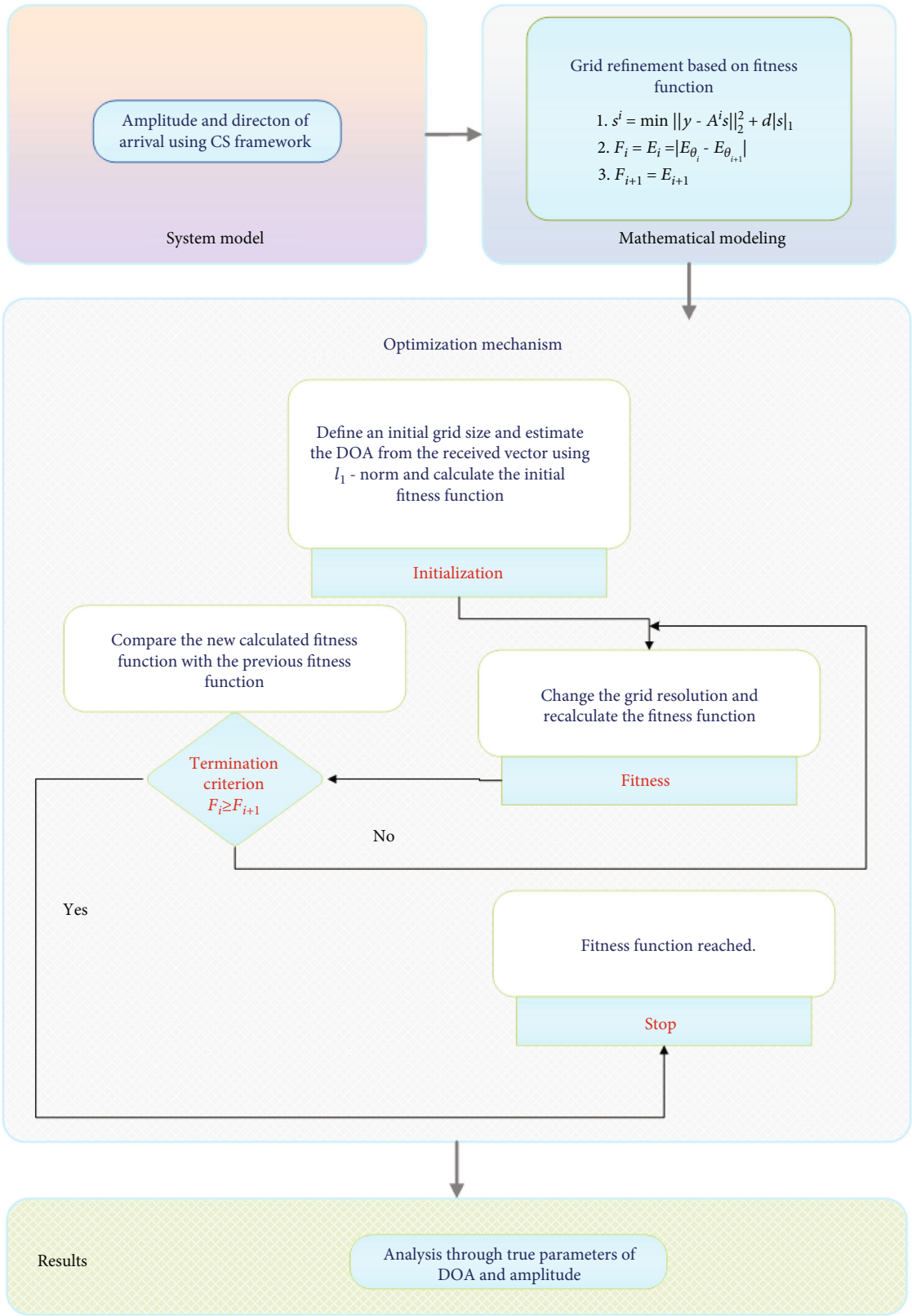


Figure 12: Graphical abstract of the proposed study.

Data Availability

There is no data associated with this manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[2] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, 1993.

[3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[4] Z. Zheng, Y. Huang, W.-Q. Wang, and H. C. So, "Direction-of-arrival estimation of coherent signals via coprime array interpolation," *IEEE Signal Processing Letters*, vol. 27, pp. 585–589, 2020.

[5] Z. Zheng and S. Mu, "Two-dimensional DOA estimation using two parallel nested arrays," *IEEE Communications Letters*, vol. 24, no. 3, pp. 568–571, 2020.

[6] H. Chen, W. Wang, and W. Liu, "Augmented quaternion ESPRIT-type DOA estimation with a crossed-dipole array," *IEEE Communications Letters*, vol. 24, no. 3, pp. 548–552, 2020.

[7] Z. Zheng, W.-Q. Wang, H. Meng, H. C. So, and H. Zhang, "Efficient beamspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11776–11789, 2018.

[8] Y. Liao and A. Abouzaid, "Resolution improvement for MUSIC and ROOT MUSIC algorithms," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, no. 2, pp. 189–197, 2015.

[9] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[10] R. Baraniuk, "Compressive sensing [Lecture Notes]," *IEEE Signal Process Magazines*, vol. 24, no. 4, pp. 118–121, 2007.

[11] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *2007 IEEE Radar Conference*, pp. 128–133, Boston, MA, USA, June 2007.

[12] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[13] N. Anselmi, G. Oliveri, M. A. Hannan, M. Salucci, and A. Massa, "Color compressive sensing imaging of arbitrary-shaped scatterers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 6, pp. 1986–1999, 2017.

[14] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, no. 11, pp. 2353–2362, 2001.

[15] J. Lu, W. Cheng, D. He, and Y. Zi, "A novel underdetermined blind source separation method with noise and unknown source number," *Journal of Sound and Vibration*, vol. 457, pp. 67–91, 2019.

[16] J. J. Fuchs, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 702–709, 2001.

[17] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[18] A. C. Gurbuz, V. Cevher, and J. H. Mcclellan, "Bearing estimation via spatial sparsity using compressive sensing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1358–1369, 2012.

[19] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[20] M. Mishali, Y. C. Eldar, and A. J. Elron, "Xampling: signal acquisition and processing in union of subspaces," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4719–4734, 2011.

[21] M. A. Hadi, S. Alshebeili, K. Jamil, and F. E. Abd El-Samie, "Compressive sensing applied to radar systems: an overview," *Signal, Image and Video Processing*, vol. 9, no. S1, pp. 25–39, 2015.

[22] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[23] T. Terada, T. Nishimura, Y. Ogawa, and T. Ohgane, "DOA estimation of multi-band signals using a sparse signal reconstruction method," in *2013 IEEE Antennas and Propagation Society International Symposium (APSURSI)*, pp. 866-867, Orlando, FL, USA, 2013.

[24] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.

[25] R. Zdunek and A. Cichocki, "Improved M-FOCUSS algorithm with overlapping blocks for locally smooth sparse signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4752–4761, 2008.

[26] Q. Liu, H. C. So, and Y. Gu, "Off-grid DOA estimation with nonconvex regularization via joint sparse representation," in *Signal Processing*, vol. 140, pp. 171–176, Elsevier, 2017.

[27] Z. Wei, X. Li, B. Wang, W. Wang, and Q. Liu, "An efficient super-resolution DOA estimator based on grid learning," *Radioengineering*, vol. 28, no. 4, pp. 785–792, 2019.

[28] Q. Liu and X. Wang, "Direction of arrival estimation via reweighted l_1 norm penalty algorithm for monostatic MIMO radar," *Multidimensional Systems and Signal Processing*, vol. 29, no. 2, pp. 733–744, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8887589, 11 pages
https://doi.org/10.1155/2020/8887589

Research Article

Management of Load-Balancing Data Stream in Interposer-Based Network-on-Chip Using Specific Virtual Channels

Mona Soleymani,[1] **Midia Reshadi**,[1] **and Ahmad Khademzadeh**[2]

[1]*Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*
[2]*Education and International Scientific Corporation Department, Iran Telecommunication Research Center, Tehran, Iran*

Correspondence should be addressed to Midia Reshadi; midia.reshadi@gmail.com

Received 11 April 2020; Revised 8 July 2020; Accepted 27 July 2020; Published 25 August 2020

Academic Editor: Farman Ullah

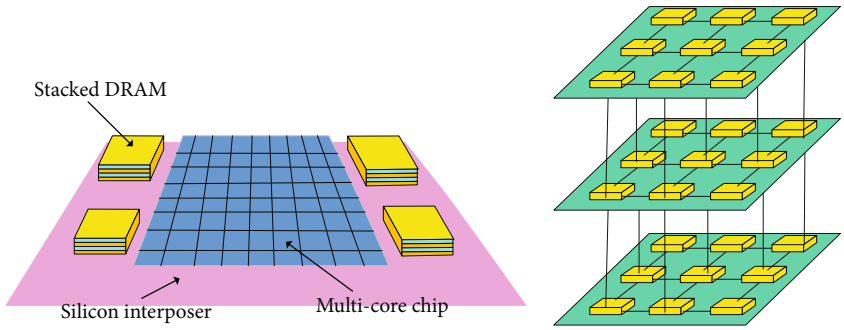
The interaction between cores and memory blocks, in multiprocessor chips and smart systems, has always been a concern as it affects network latency, memory capacity, and power consumption. A new 2.5-dimensional architecture has been introduced in which the communication between the processing elements and the memory blocks is provided through a layer called the interposer. If the core wants to connect to another, it uses the top layer, and if it wants to interact with the memory blocks, it uses the interposer layer. In a case that coherence traffic at the processing layer increases to the extent that congestion occurs, a part of this traffic may be transferred to the interposer network under a mechanism called load balancing. When coherence traffic is moved to the interposer layer, as an alternative way, this may interfere with memory traffic. This paper introduces a mechanism in which the aforementioned interference may be avoided by defining two different virtual channels and using multiple links which specifically determines which memory block is going to be accessed. Our method is based on the destination address to recognize which channel and link should be selected while using the interposer layer. The simulation results show that the proposed mechanism has improved by 32% and 14% latency compared to the traditional load-balancing and unbalanced mechanisms, respectively.

1. Introduction

With the silicon interposer-based technology, a new architecture called 2.5D has been introduced in which an increasing number of memory blocks may be merged throughout multiprocessor chips [1]. What spotlights this architecture is that it increases the amount of memory, which is horizontally located around the processing chip. Unlike the 3D NoC, which the capacity of memory in a package is banned by the size of a processor chip [1–4], in 2.5D technology, the size of the interposer layer determines how much memory blocks are able to be integrated [1]. Figure 1 shows a 2.5D stacking technology with four DRAM stacks on the interposer. Placing on the two horizontal sides, more memory stacks are integrated through the silicon interposer area. For this reason, higher capacities and higher bandwidth are achievable compared to 3D technology [5–7]. In this figure, the differences between 2.5D and 3D technologies have been shown.

In 2.5D architecture, there are two layers: the conventional processing layer, including processing elements, routers, and links, which is located on top layer, and a newly emerged layer called interposer placed below it. The connection between processing cores are regularly established on the top layer; in the case that a core wants to interact with a memory block, it may use the interposer layer to send/receive its packet [8].

In Figure 1(a), it may be seen that memory blocks are located on the silicon interposer and connected with the interposer layer through interface nodes. There are also two disparate types of traffic: the first one is the core-to-core traffic associated with the interaction between cores and the second is the core-to-memory traffic, which is responsible



(a) 2.5D architecture (b) 3D architecture

Figure 1: 2.5D technology vs. 3D technology.

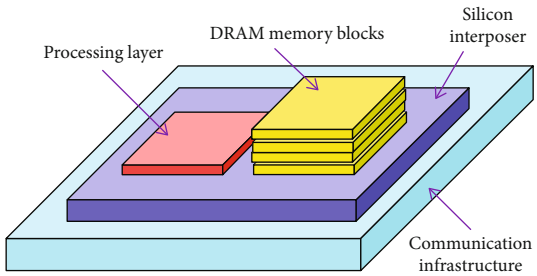


Figure 2: 3D view of 2.5D network-on-chip.

for transferring packets between memory blocks and cores [1, 6]. Core-to-core traffic, which is known for coherence traffic, should be distributed across all the networks in order to avoid protocol-level deadlock [1]. Each of aforementioned types of traffic has its own characteristics [1]. Coherence traffic is established based on peer to peer communication patterns while core-to-memory traffic generates a many-to-few traffic pattern [1]. As each of these traffic patterns is utilized for different purposes, they should be segregated from each other. Figure 2 shows another view of the 2.5D network-on-chip that memory blocks are just presented on the right side of the processing layer.

One of the characteristics of the interposer layer is the support of different heterogeneous integrations and structures [9–11].

As it may be seen in Figure 3, disparate components with different functions may be integrated through the interposer layer. Internet of Things may be a favorable heterogeneous utilization in which multifunctional and compact devices with high performance and low energy consumption are needed. Figure 4 demonstrates two disparate traffics in two different paths. In a conventional 2D network-on-chip, a processing core may send its packet to another processing core using blue routers; similarly, a processing core may interact with a memory block through green routers. The internal structure of a tile, containing an element and a router alongside cache memories, is also presented in Figure 4.

Considering the upper layer, when the volume of core-to-core coherence traffic is exceedingly increasing, congestion is potentially created due to heavy workload while the lower layer seems to be underutilized [1]. As a result, the performance will be weak [8]. When the congestion occurred in the processing layer, nodes cannot potentially send/receive their packets easily and there are lots of time to be wasted as the routers. This will spoil the performance of routing and switching since some packets are stuck to some routers and there is no way to go ahead. Latency is the most valuable parameter which is negatively affected due to this phenomenon. At this time, a mechanism called load balancing has been proposed to alleviate the congestion and related problems [8].

What the load-balancing mechanism does is to balance loads of traffic on the existing network layers and enhance the utilization of available resources [8]. With regard to the volume of traffic between memory-core and intercore communication, it seems logical that the number of packets streamed across the cores is higher than the memory traffic. This means that the data from one processing node to another similar one flowed more frequently than the request of reaching memory blocks [1]. This is the reason why congestion is supposed to be seen in the processing layer while the interposer layer does hardly face congestion, if not at all. In order to recognize whether the load-balancing method should be utilized or not, the volume of the coherence traffic at the top layer is supposed to size up. Some methods have been proposed to achieve the congestion information which may be categorized into two different groups: buffer-aware and latency-aware. In the former method, according to the filling capacity of each neighbor router, the information of congestion may be detected, which is not globally reliable because of the locality [12]. So, the buffer-aware strategy is not appropriate in 2.5D NoC. In a study proposed by Chen et al. [8], known as a dynamic latency-aware load balancing (DLL), the latency of congestion packets may be properly recorded. This is mainly because this information is tracked by the clock in every router until it arrives at the destination nodes. We use the DLL congestion detection method to recognize when the load-balancing mechanism should be used according to the congestion data; this strategy is designed based on latency-aware architecture.

In the load-balancing method, after detecting congestion in the upper layer, it is possible to send some packets randomly to the interposer layer so as to move and reach their destinations. Actually, they use the interposer layer as an alternative way to escape from the heavy workload they are confronted with at the top layer. As mentioned earlier, the

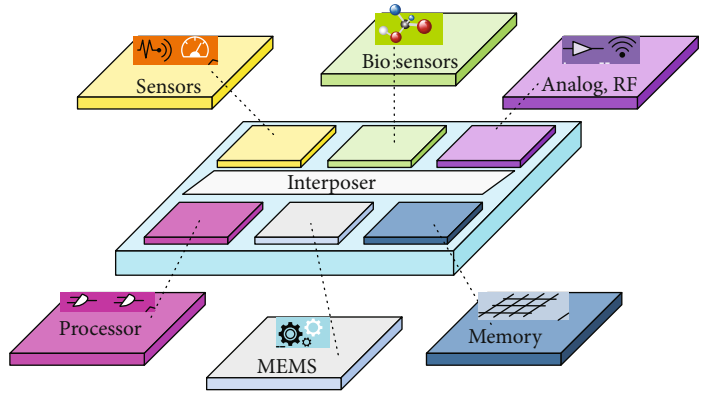


Figure 3: The interposer layer which provides different integrations with various technologies on a single system.

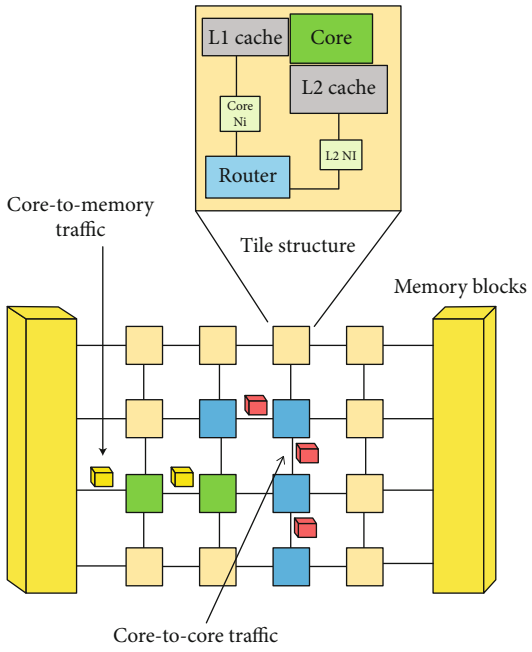


Figure 4: Different types of traffic.

interposer layer is targeted to use for core-to-memory traffic, so by sending the core-to-core traffic to the lower layer, these two types of packets are susceptible to interfere. In this paper, we aim to propose a way in which the load-balancing mechanism does not result in conflict. This happens due to the separation of traffics by utilizing virtual channels and multiple links.

First, we want to explain that the communication between the different components is managed differently in our work. Core-to-core communication is different from core-memory communication. Communication between core-to-core occurs via routers. The router is provided with an entry in the form of a kernel message or whatever adjacent components. The router is partitioned into two layers, i.e., the packet layer and the circuit layer. By the type and the size of the data to be transferred via the routers, the router layer is used. A packet-switched NoC is used to transfer small amounts of data and in a synchronized communication, whereas to transfer data transfers such as multimedia and big data applications, the circuit-switched router (NoC) is used, which is flexible and configures all combinatorial cable connections, resulting in less latency and better energy efficiency. The core-memory communication is done via the memory pipeline access ports dedicated by a wired connection. All memory is shared and accessible by pipeline from all cores within clusters.

The rest of the paper is arranged as follows: Section 2 tends to represent related work. In section 3, we attempt to assume a target structure and show how our proposed method may effectively manage and control the load-balancing mechanism. Experimental results are drawn in Section 4, and finally, Section 5 presents the conclusion.

1.1. Related Work. In order to take advantage of 3D die stacking, several papers have been recently published [13–16]. However, traffic has not been differentiated in many of the proposed methods [1], which means traffic from any types may be passed through each of the layers. Xu et al. [17] proposed that long links may be utilized to customize every 3D layer and this also depletes the hop counts for all traffics. Disparate physical layers have been proposed in [18] in which virtual channels are used to categorize coherence traffic types to hinder protocol-level deadlock. Furthermore, 3D NoC architecture faces some issues, namely, thermal and cooling troubles, the lack of EDA tools, and many problems related to the test [19]; for this reason, 2.5D NoC, which is based on silicon stacking, has recently been popular [5, 20]. That is, many design tools [1, 6, 7, 21] support 2.5D stacking architecture and likewise, many use it for further GPU designs [1, 22]. It is worth mentioning that this newly emerged technology is already seen in many commercial products [1, 23, 24]. Jerger et al. and Li et al. were the pioneers to conduct research and investigate the effect of the interposer in 2.5D NoC in favor of space [1, 8]. Their work first focused on how different topologies may affect the load-balancing strategy in 2.5D interposer-based architecture and then illustrated the destination latency-aware method as an effective way of detecting information related to the congestion [1]. In the aforementioned work, packets are sent through the paths and received by other paths, and as a result, an unsustainable congestion detection method was used for network selection [8]. Figure 5 shows an example of an interleaved-based system,

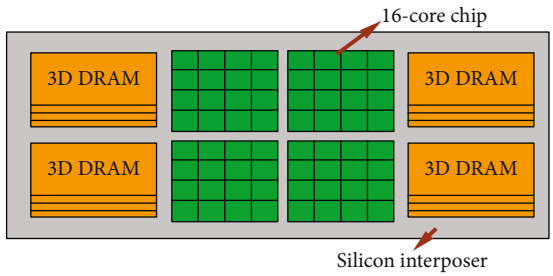


Figure 5: 64-core system composed with DRAM [25].

including four 16-core processing possibilities, an interconnect replica, and a quartet of HBM DRAM stacks located along the left and right sides of the processors.

For recognizing the area related to the congestion, another suitable strategy called Dynamic Latency-aware Load-balancing (DLL) mechanism [8] was proposed. Like Enright's work, the DLL method has used the latency-aware strategy, and it is implemented into different stages: detecting and collecting congestion data and distributing it towards source nodes. Apart from topologies evaluated in [1], other topologies in 2.5D network-on-chip have been proposed such as [25]. The topology named ClusCross used in [25] defines any small chip as a cluster and utilizes long links to easily have access to memory blocks, which boost the bisection bandwidth and deplete average hop count.

Other works like [26] has addressed some issues which results in avoiding bottleneck by looking into the new interposer design space of passive and active interposer technologies, its topologies, and clocking schemes to determine the cost-optimal interposer architectures. The work in [27] has proposed a method called EIRs (Equivalent Injection Routers) which transforms the few-to-many traffic pattern to many-to-many pattern along with the interposer links. EIR scheme solves the bottleneck problem as well as enhancing throughput among manycore processors.

The load-balancing method has been introduced by Jerger et al. to be used in the 2.5D interposer network-on-chip, but this is not controlled and managed. In our previous paper, we have introduced a limited method called CLBM (Controlled Load-Balancing Mechanism) which controls the load-balancing method by defining a forbidden area [28]. Furthermore, in the aforementioned paper, a multicast ring has been introduced to propagate the latency packet across the source nodes. Although in our previous work we have controlled the load-balancing mechanism, it yet does not separate various types of traffic and just prevents the edge grids from being the hotspot. In this paper, we try to use the interposer layer as a second option, when congestion is detected in the processing layer, through a manageable and controllable way that ensures there is no conflict and interference between coherence and core-memory traffic.

2. Proposed Method

In this section, we introduce our 2.5D target structure and the conventional load-balancing strategy. Then, the potential challenges in relation with the load-balancing strategy on interposer-based NoC are explored and we explain how using virtual channels may manage and control traffic interference.

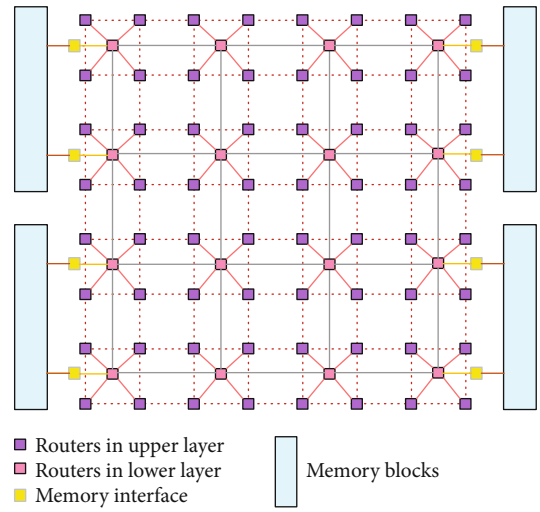


Figure 6: The view of the processing layer and the interposer layer topology.

2.1. Target Structure. Our target structure, in this paper, is based on a 64-core processing unit and 4 stacked memory blocks integrated on the 2.5D network-on-chip architecture. Our processing layer is designed by the mesh topology, containing all the processing cores, while the lower layer embodies the interposer routers which are responsible for memory interaction. TSVs (Through Silicon Via) and μbumps connect these two layers. Because μbumps take a heavy toll on the network [29], the concentrated mesh has been proposed to be used for the interposer routers and this decreases the count of routers. According to the concentrated mesh, every four processing cores are attached to an interposer router. The memory blocks are placed on the horizontal sides of the interposer layer.

Figure 6 illustrates all the routers in a view. There are some yellow nodes known as memory interface nodes, which make a connection between edge interposer routers and the memory blocks.

As mentioned earlier, we may divide the traffic into two different categories: the first one is related to the interaction between processing cores with each other and we know them as CtC. The other is about the connection between processing cores and memory blocks known as CtM in this paper.

2.2. Load-Balancing Mechanism. As usual, in the interposer-based network-on-chip, CtC packets go through the CPU layer and the interposer layer is utilized for CtM traffic [1]. The volume of CtC traffic sometimes appears to be exceeding from a specific threshold, whereas a less number of cores have a request for interacting with memory [1, 8]. In such a scenario, congestion is going to happen on the processing network, meaning that the workload of CtC is more than the CtM traffic. Detecting congestion in the CPU layer is synonymous with the use of the load-balancing mechanism. In

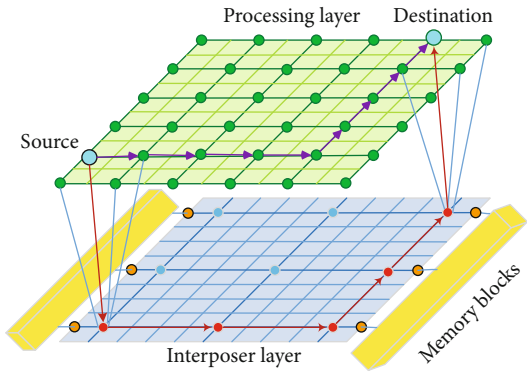


Figure 7: An illustration of load balancing.

this method, some CtC-related packets may choose the interposer layer to move and route. When they approach the nearest node of their destination, they tend to be transferred to the upper layer again and reach their destination node. Figure 7 illustrates how the load-balancing mechanism works. According to this figure, source and destination nodes want to communicate with each other and both of them as a part of CtC traffic are located on the processing layer. Conventionally, all CtC packets are supposed to be routed on the CPU layer in purple links through eight hops. Given the fact that congestion has occurred in the upper layer, the source node sends its packet to the interposer router attached to it. This packet is routed in the interposer layer and after crossing six hops back to the top layer and reach its destination. Even though the length of links has increased, the number of hops has reduced.

We empirically consider 10 cycles as a basic threshold to recognize congestion for interlayer networks. There are still other criteria established when traffic is permitted to move across the second network [30]. The nodes in the interposer layer have been called grids. In the load-balancing mechanism, the CtC traffic in the upper layer is passed through the interposer layer and this will interfere with CtM which is basically associated with the interposer layer. This interference not only may disturb the process of having access to memory blocks for CtM but also may create undesirable and mixed congestion in the lower layer. This paper is aimed at segregating these two types of traffic when a load-balancing mechanism is used.

2.3. Proposed Strategy. This paper is aimed at separating CtC traffic from CtM ones when both try to use the interposer layer as a network to reach their destination. Our strategy is implemented on the routers associated with the interposer layer so as to recognize how to send each individual packet. As mentioned earlier, routers in the interposer layers are known as grids which are numbered as shown in Figure 8.

Every four routers from the processing layer are connected to one specific router from the lower layer, so the number of interposer routers is one-quarter of the processing layer. Figure 9 shows routing mechanism of our proposed method.

We propose that using virtual channels at the interposer layer may be an effective solution to allocate two different paths for each type of traffic. In Figure 10, it is clearly presented.

When a core wants to send a packet from the processing layer to the interposer layer in the light of the load-balancing mechanism, in a grid attached to that core, it is decided which virtual channel should be selected to go through. This may be possible just by considering the address of destination which that packet carries. Since we have 64 cores in the upper layer, if the destination address is more than 63, this is synonymous with the proof that the packet is CtC and should be transmitted throughout VC1 in the interposer layer. This channel segregation in the interposer layer in which two different types of traffic may be differentiated is the solution to avoid the possible conflict as well as potential congestion.

An appropriate channel is initially chosen based on arbitration and switching mechanism [29]. As it can be seen in Figure 9, if we look at a tile in detail, first, according to the destination grid address, it is decided in arbitration block which virtual channel should be selected. Then, in switching part, through input/output ports, the coming packet goes ahead through the ideal virtual channel which is either VC1 associated to CtC or VC2 connected to CtM. This selection happens once the packet is sent from the upper layer to the lower layer.

Figure 11 demonstrates our proposed method with an example. Considering the CtC traffic, node 21 aims to send its packet to node 49. In the case of using the load-balancing mechanism, after detecting congestion in the processing layer, this packet should be sent to grid 6 connected to node 21 in the interposer layer. Once it reaches to grid 6, according to its destination address which is less than 63, this packet is associated to the CtC traffic. This means that it is supposed to go back to the upper layer to node 49, its destination, after crossing grids 5, 4, 8, and 12, respectively, on the light blue path.

At the same time, node 26 wants to send its packet to a memory block located in the west-south corner. Indeed, it should be gone to node 82 which is an interface node connected to the memory block. As it may be seen, its destination address is higher than 63 so it comes from a CtM traffic and uses the interposer layer anyway. The packet belongs to node 26 which goes through grids 5, 4, 8, and 12, respectively, on the dark blue path. Using VCs in the interposer layer provides different types of traffics not to meddle with each other. This will improve the network latency when packets are less when waiting to reach their destination either through the processing core or memory block.

After dedicating a specific virtual channel to the CtM traffic, it is proposed that memory-related packets may be separated as well according to the memory block numbers they want to interact with. We propose multiple links in the interposer layer for the virtual channel allocated to the CtM traffic. For deciding on what link is the best to choose for CtM packets, we number memory blocks as shown in Figure 12.

The reason for this numbering is related to our proposed way shown in Figure 13. This is actually a demultiplexing method. In Table 1, grids connected to memory interface have been presented alongside the number of memory blocks

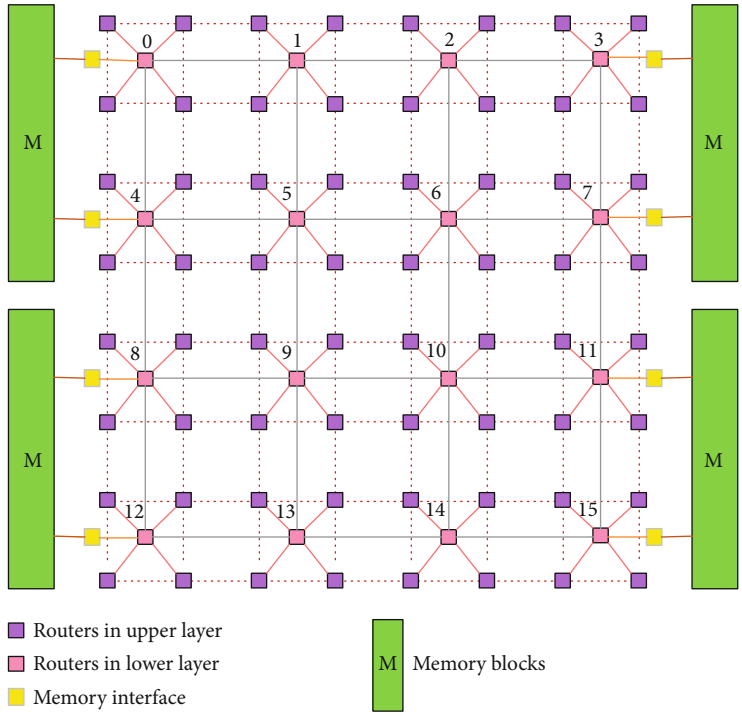


Figure 8: Illustration of grid numbers.

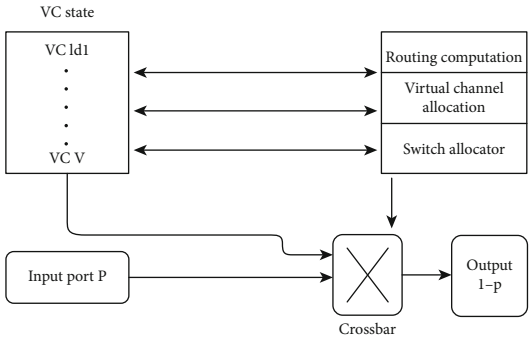


Figure 9: Proposed routing mechanism.

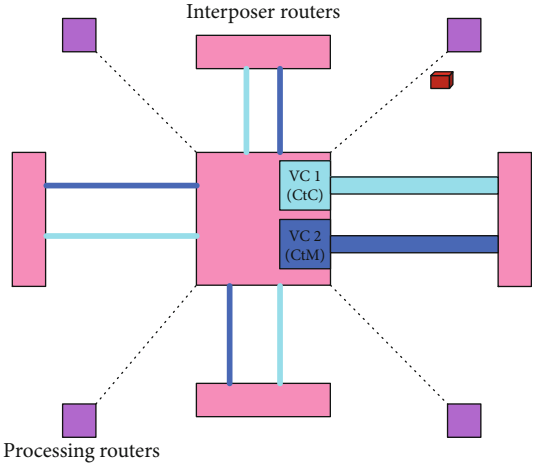


Figure 10: Using specific virtual channels in order to separate different types of traffic.

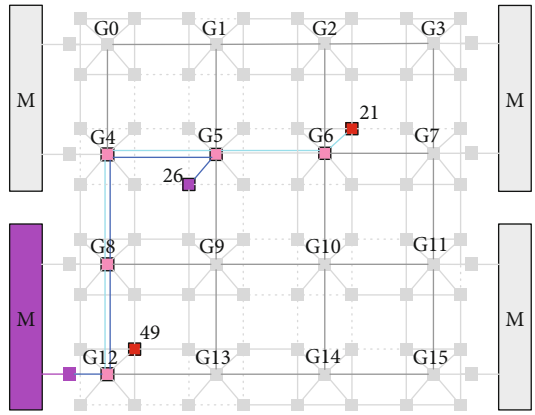


Figure 11: An example of how specific virtual channels may lead to the traffic segregation.

related to each of them both in a decimal format (Table 1(a)) and a binary format (Table 1(b)). In the binary format, there is a relationship between grid numbers and memory block numbers which we numbered accordingly. As it may be seen in Table 1(b), if we put the first and the last bit of grid numbers together, this will build the memory block numbers. That is, when a packet reaches to an interposer router and go along with its specific virtual channel, it may be decided on a specific link to go through to a particular memory block. This parallel transferring enables CtM packets to reach their memory block without any conflict with other packets either CtC or other CtM ones.

This is implemented by a demultiplexer whose selection links are the first and the last bits of destination grid address as shown in Figure 13. For example, consider a CtM-related packet with the destination grid ID of 11, which is supposed

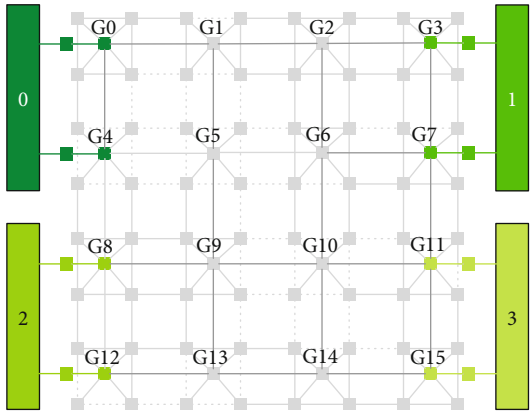


Figure 12: Allocating specific numbers to each group of memory blocks.

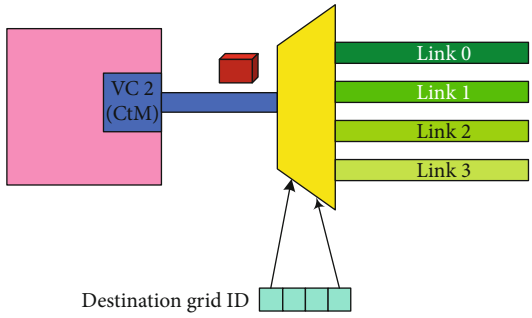


Figure 13: The demultiplexing of memory-related packets.

to transfer through the interposer layer, anyway, reaches to the VC2 (allocated to CtM traffic). In this step, the first bit and the last bit of 11 in the binary format shown in Table 1 determine which link is the best to move. In this example, it is link 3. The function of demultiplexer may be implemented in two different ways: (1) put a physical demultiplexer beside CtM virtual channel or (2) manage the function of division by controlling the time throughout each packet that wants to move to reach its specific memory block. In this paper, in order to diminish overhead, we consider the latter strategy.

We describe our proposed method in a pseudocode form as shown in Figure 14.

3. Experiments

To perform our evaluation, we used a cycle accurate interconnection network simulator, BookSim [31]. In order to simulate our proposed structure, we altered some characteristics of this simulator like workload trace and the file related to the network. A broad range of workloads have been utilized so as to size up our strategy and other methods such as the traditional load-balancing method and a scenario in which there is no load-balancing strategy. Table 2 contains the basic parameters related to the simulation.

3.1. Coherence and Memory Traffic Analysis. In order to evaluate the influence of different types of traffic, we simulated various proportions of memory traffic and processing layer ones.

Table 1: The connection between grids and memory blocks in (a) decimal and (b) binary format. Figure (b) presents the relationship between grids' numbers and related memory block numbers.

(a)

Grid (Connected to Memory Intereface) Numbers	Related Memory Block Numbers
0	0
3	1
4	0
7	1
8	2
11	3
12	2
15	3

(b)

Link	Edge Grid numbers in Bin				Memory block numbers in Bin
Link 0	0	0	0	0	00
	0	1	0	0	00
Link 1	0	0	1	1	01
	0	1	1	1	01
Link 2	1	0	0	0	10
	1	1	0	0	10
Link 3	1	0	1	1	11
	1	1	1	1	11

$Total\_Latency = Packets\ numbers \times Latency\ for\ one\ Packet$
$If\ (Total\_Latency \geq Threshold)$
{
Phase 1: *Packet ejected from upper router => connected grid*
In connected grid:
1) *Arbitration*: *evaluation of the header flit including the final destination*
$If\ (destination\ address \leq 63)\ then\ i = 0$
else $i = 1$;
2) *Switching*: *choose the* VC_i
3) *Routing function*: *XY-based routing*
Phase 2: *case* VC_2 *with*
Link 0 *if Grid ID* = 0,4;
Link 1 *if Grid ID* = 3,7;
Link 2 *if Grid ID* = 8,12;
Link 3 *if Grid ID* = 11,15;

Figure 14: The pseudocode related to our method.

For this, disparate injection rates have been used so as to yield better results. The results are shown in Figure 15;

We evaluated five various distributions among memory and processing cores. As it may be seen in Figure 15, when the percentage of the memory traffic increases, the latency of the network rises accordingly. There is also an increase

Table 2: Simulation parameters.

Primary factors	
Injection rate	0.01, 0.5, and 0.1
Processing die topology	Mesh 8 × 8
Interposer topology	Mesh 4 × 6
Coherence traffic	Uniform, neighbor, randperm, and hotspot
The size of packet	5 flits
Memory traffic percentages	25%, 30%, 40%, and 50%
Percentages of traffics	1: 25% memory traffic, 75% CPU traffic 2: 35% memory traffic, 65% CPU traffic 3: 45% memory traffic, 55% CPU traffic 4: 55% memory traffic, 45% CPU traffic
Virtual channels for the interposer layer	2 VCs

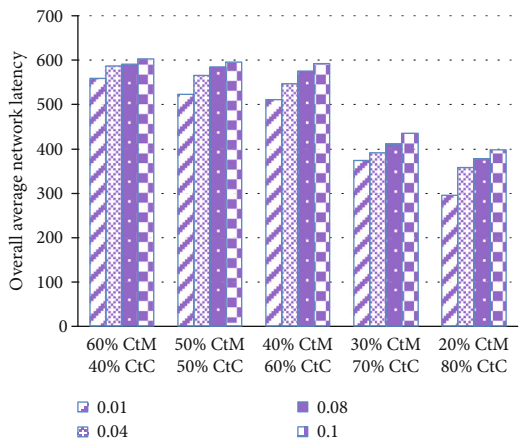


Figure 15: Various percentages of memory traffic and CPU traffic. Each bar represents latencies of different injection rates.

in the amount of latency when the rate of injection increases. The highest overall average latency is associated with the highest CtM traffic (60%), and the highest injection rate which is 0.1.

3.2. Performance Comparison. In this paper, we try to simulate various percentages of both the coherence traffic and the memory traffic. Four different states have been considered as shown in Figure 15: (a) 25% CtM and 75% CtC traffic, (b) 35% CtM and 65% CtC traffic, (c) 45% CtM and 55% CtC traffic, and (d) 55% CtM and 45% CtC.

First, the results bring about that the more proportion of traffic dedicated to memory, the more network latency is expected.

Second, in a scenario that there is no load-balancing mechanism, the overall average latency is the highest in four states. This is rational because coherence traffic is only transmitted on the processing layer and not allowed to use another alternative layer to move (the interposer layer); likewise, the lower layer is merely devoted to memory-related traffic. This is not flexible, and once the coherence traffic exceeds from a certain threshold, this is susceptible to be a bottleneck due to the potential congestion. Finally, as the packets injected to the network tend to increase, either CtM or CtC, the overall average network latency rises.

According to what was explained above and considering Figure 15, our proposed method, which uses virtual channels and multiple links to hinder the interference brought by the load-balancing mechanism, improves the latency compared to the unbalancing and conventional load-balancing mechanisms. Figure 16 demonstrates how the aforementioned scenarios affect the overall network latency with three different cases of memory traffic. We also analyzed our proposed strategy on the other types of traffic such as uniform, randperm, hotspot, and neighbor.

As it can be observed in Figure 17, the hotspot traffic pattern has far more latency in comparison with other types of traffic. The rationale behind this is that when this traffic pattern is distributed upon the processing layer, some nodes will be susceptible to face bottleneck. This, therefore, makes these nodes to be saturated for some instances which affects the overall latency and culminates in the load-balancing strategy as well. Primarily, due to the fact that specific virtual channels provide packets with disparate links as soon as they reach their connected grid at the interposer layer, they are supposed to go through their unique path which is different from the paths of other types of traffic use. This is synonymous with no conflict and no interference although accompanied by a little overhead. Using virtual channels and multiple links at the lower layer, the interposer routers face the physical overhead which may be downplayed by a considerable latency improvement. Supplementary to this, as virtual channels are the internal feature of routers, using them are not even considered an overhead.

3.3. Network Utilization and Efficiency. When using the load-balancing mechanism in 2.5D-based network-on-chip, the underutilized resources, namely, routers at the interposer layer are effectively utilized. Simply put, we have two network layers: the processing layer and the interposer layer. If congestion occurs at the top layer, most of the upper routers involve in the process of data stream are exploited, whereas the interposer routers have little function to do. This, therefore, culminates in network utilization in the light of the load-balancing method. Likewise, our proposed method enhances efficiency and network utilization at the interposer layer throughout using the load-balancing mechanism. Packets sent from the processing layer to the interposer layer are managed and controlled in our method, which enables the network not to be overwhelmed with different types of traffic. We have specified two kinds of virtual channels to avoid data conflict arising from various traffic flows, which are CtC and CtM. As well as this, multiple links at the interposer layer facilitate the way of distributing CtM packets to their particular memory which focuses mainly on the interposer network utilization and its efficiency.

Integrated multicore microprocessor with 2.5D silicon interposer, memory, and accelerator, which offers the advantages of flexibility of system integration and energy efficiency in terms of network utilization.

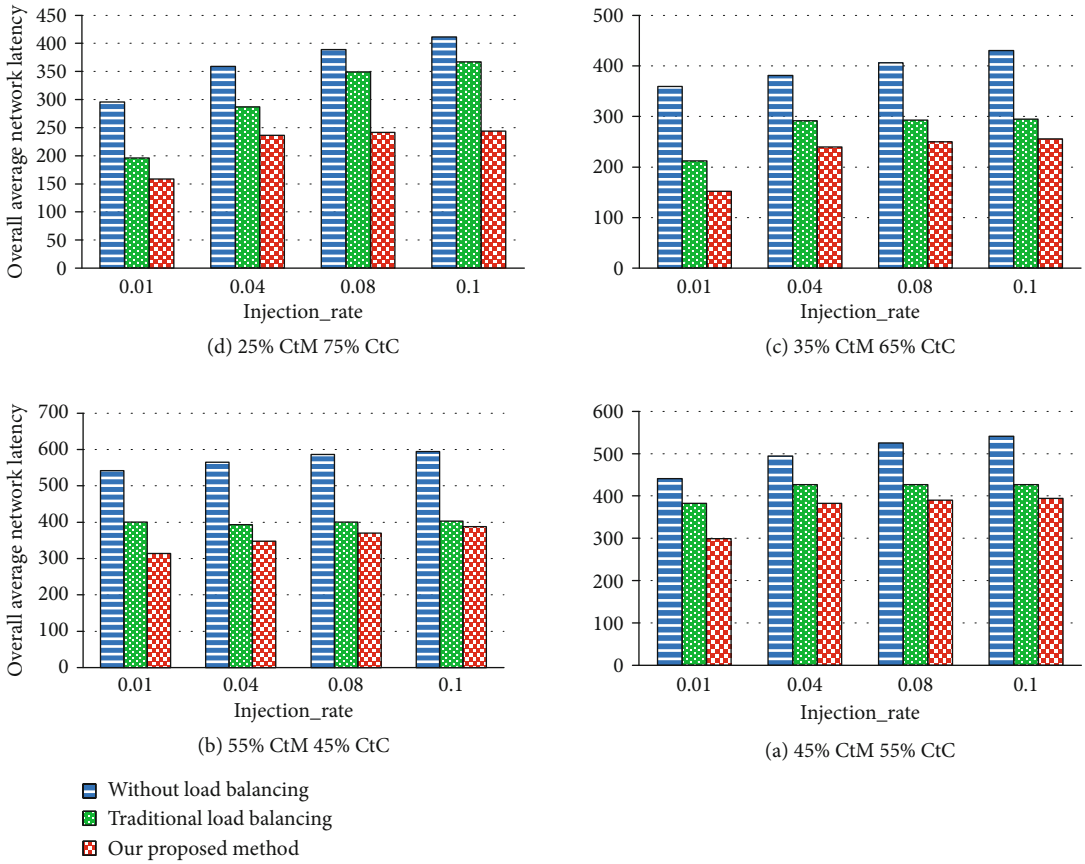


Figure 16: Overall average network latency for three methods: the lack of load balancing, conventional load balancing, and proposed strategy.

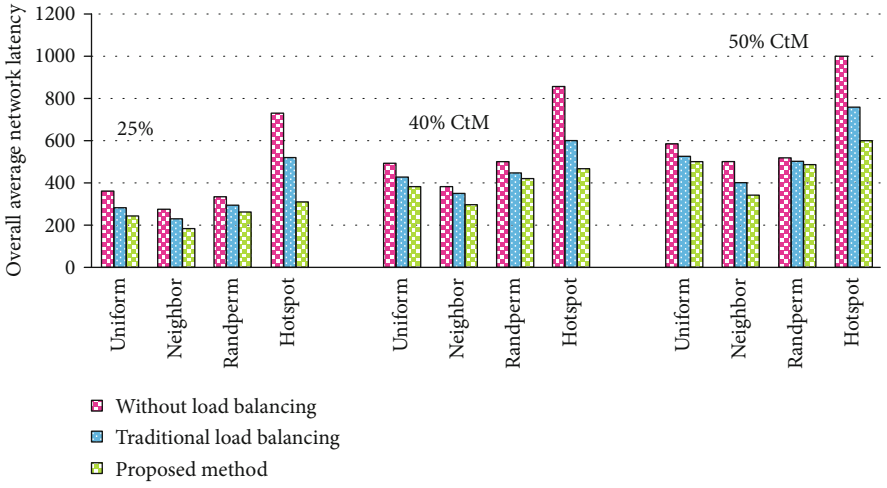


Figure 17: Overall average network latency for three methods: the lack of load balancing, conventional load balancing, and proposed strategy.

4. Conclusions

In order to manage and control interference and conflict caused by the load-balancing strategy in the interposer-based network-on-chip, we proposed a method. When coherence traffic packets are moved from up to down, based on the load-balancing method, a conflict with CtM traffic is expected. The virtual channel-based mechanism proposed in this paper separates the core-to-core traffic from the core-to-memory traffic in the interposer routers and results in no collision between CtM and CtC traffic. The CtM-related virtual channel is also proposed to use multiple links with the aim of classification of memory blocks in parallel. Our proposed strategy has improved the overall average network latency by 32% and 14% in comparison with the case that there are no load-balancing method and conventional

load-balancing mechanism, respectively. This method has been used in smart systems efficiently, and in the future, it will improve real-time applications functionality as well.

Data Availability

Data are available on request through contacting with mona.soleymani@gmail.com.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] N. E. Jerger, A. Kannan, Z. Li, and G. H. Loh, "NoC architectures for silicon interposer systems: why pay for more wires when you may get them (from your interposer) for free? in: Microarchitecture (MICRO)," in *2014 47th Annual IEEE/ACM International Symposium on*, pp. 458–470, Cambridge, UK, December 2014.

[2] S. Killge, N. Neumann, D. Plettemeier, and J. W. Bartha, "Optical through-silicon vias," *3D Stacked Chips*, pp. 221–234, 2016.

[3] G. Y. Tang, S. P. Tan, N. Khan et al., "Integrated liquid cooling systems for 3-D stacked TSV modules," *IEEE Transactions on Components and Packaging Technologies*, vol. 33, no. 1, pp. 184–195, 2010.

[4] J. H. Lau, Y. S. Chan, and S. W. R. Lee, "Thermal-enhanced and cost-effective 3D IC integration with TSV (through-silicon via) interposers for high-performance applications," in *Proceedings of the ASME 2010 International Mechanical Engineering Congress and Exposition. Volume 4: Electronics and Photonics*, pp. 137–144, Vancouver, British Columbia, Canada, January 2010.

[5] G. H. Loh, N. E. Jerger, A. Kannan, and Y. Eckert, "Interconnect memory challenges for multi-chip, silicon interposer systems," *Proceedings of the 2015 International Symposium on Memory Systems - MEMSYS '15*, pp. 3–10, 2015.

[6] S. Dadashi, M. Reshadi, A. Reza, and A. Khademzadeh, "An expandable topology with low wiring congestion for silicon interposer-based network-on-chip systems," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 12, 2019.

[7] S. Dadashi, M. Reshadi, A. Reza, and A. Khademzadeh, "Decreasing latency considering power consumption issue in silicon interposer-based network-on-chip," *The Journal of Supercomputing*, vol. 75, no. 11, pp. 7646–7664, 2019.

[8] C. Li, S. Ma, L. Wang, Z. Wang, X. Zhao, and Y. Guo, "DLL: a dynamic latency-aware load-balancing strategy in 2.5D NoC architecture," in *2016 IEEE 34th International Conference on Computer Design (ICCD)*, pp. 646–653, Scottsdale, AZ, USA, October 2016.

[9] G. Kumar, T. Bandyopadhyay, V. Sukumaran, V. Sundaram, S. K. Lim, and R. Tummala, "Ultra-high I/O density glass/silicon interposers for high bandwidth smart mobile applications," in *2011 IEEE 61st Electronic Components and Technology Conference (ECTC)*, pp. 217–223, Lake Buena Vista, FL, USA, May 2011.

[10] T. G. Lenihan and E. J. Vardaman, "Challenges to consider in organic interposer hvm," *TechSearch International for iNEMI Substrate & Packaging Workshop, Toyama*, 2014.

[11] M. Odeh, B. Voort, A. Anjum, B. Paredes, C. Dimas, and M. S. Dahlem, "Gradient-index optofluidic waveguide in polydimethylsiloxane," *Applied Optics*, vol. 56, no. 4, pp. 1202–1206, 2017.

[12] P. Gratz, B. Grot, and S. W. Keckler, "Regional congestion awareness for load balance in network-on-chip," in *2008 IEEE 14th International Symposium on High Performance Computer Architecture*, pp. 203–214, Salt Lake City, UT, USA, February 2008.

[13] J. Kim, C. Nicopoulos, D. Park et al., "A novel dimensionally-decomposed router for on-chip communication in 3D architectures," *Proceedings of the 34th annual international symposium on Computer architecture - ISCA '07*, 2007.

[14] F. Li, C. Nicopoulos, T. Richardson, Y. Xie, V. Narayanan, and M. Kandemir, "Design and management of 3D chip multiprocessors using network-in-memory," in *33rd International Symposium on Computer Architecture (ISCA'06)*, pp. 130–141, Boston, MA, June 2006.

[15] D. Park, S. Eachempati, R. Das et al., "MIRA: a multi-layered on-chip interconnect router architecture," in *2008 International Symposium on Computer Architecture*, pp. 251–261, Beijing, China, June 2008.

[16] A. Zia, S. Kannan, G. Rose, and H. J. Chao, "Highly-scalable 3D CLOS NOC for many-core CMPs," in *NEWCAS Conference*, pp. 229–232, Montreal, Mayada, June 2010.

[17] Y. Xu, Y. Du, B. Zhoo, X. Zhou, Y. Zhang, and J. Yang, "A low-radix and low-diameter 3D interconnection network design," in *2009 IEEE 15th International Symposium on High Performance Computer Architecture*, pp. 30–42, Raleigh, NC, February 2009.

[18] S. Volos, C. Seiculescu, B. Grot, N. K. Pour, B. Falsafi, and G. De Micheli, "CCNoC: specializing on-chip interconnects for energy efficiency in cache coherent servers," in *2012 IEEE/ACM Sixth International Symposium on Networks-on-Chip*, pp. 67–74, Lyngby, Denmark, May 2012.

[19] J. Xie, J. Zhao, X. Dong, and Y. Xie, "Architectural benefits and design challenges for three-dimensional integrated circuits," in *2010 IEEE Asia Pacific Conference on Circuits and Systems*, pp. 540–543, Kuala Lumpur, Malaysia, December 2010.

[20] C. Li, S. Ma, L. Wang, Z. Wang, X. Zhao, and Y. Guo, "Overcoming and analyzing the bottleneck of interposer network in 2.5D NoC architecture," in *Advanced computer Archi- tecture: 11th Conference, ACA 2016*, pp. 40–47, Weihai, China, August 2016.

[21] M. Jackson, "A silicon interposer-based 2.5D-IC design flow, going 3D by evolution rather than by revolution," in *2011 IEEE International 3D Systems Integration Conference (3DIC), 2011 IEEE International*, Osaka, Japan, January 2012.

[22] J.-H. Huang, *NVidia GPU technology conference: keynote*, 2013.

[23] K. Saban, *Xilinx stacked silicon interconnect technology delivers breakthrough FPGA capacity, bandwidth, and power efficiency*, Xilinx, White Paper, United States, 2011.

[24] M. J. Santarini, "Stacked and loaded: Xilinx SSI, 28-Gbps I/O yield amazing FPGAs," *Xcell Journal*, vol. 74, pp. 8–13, 2011.

[25] H. Shabani and X. Guo, "Clus Cross: a new topology for silicon interposer-based network-on-chip," *Proceedings of*

13th IEEE/ACM international symposium on network-on-chip, 2019.

[26] D. Stow, I. Akgun, and Y. Xie, "Investigation of cost-optimal network-on-chip for passive and active interposer systems," in *2019 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)*, Las Vegas, NV, USA, USA, June 2019.

[27] Y. Li and L. Chen, "Equi Nox: Equivalent Noc injection routers for silicon interposer based throughput processors," in *2020 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, San Diego, CA, USA, USA, February 2020.

[28] M. Soleymani, M. Reshadi, N. Bagherzadeh, and A. Khademzadeh, "CLBM: controlled load-balancing mechanism for congestion management in silicon interposer NoC architecture," *Journal of System Architecture*, vol. 98, pp. 102–113, 2019.

[29] C. Liu, L. Zhang, Y. Han, and X. Li, "Vertical interconnects squeezing in symmetric 3d mesh network-on-chip," in *16th Asia and South Pacific Design Automation Conference (ASP-DAC 2011)*, pp. 357–362, Yokohama, Japan, January 2011.

[30] R. Das, S. Narayanasamy, S. K. Satpathy, and R. Dreslinski, "Catnap," *ACM SIGARCH Computer Architecture News*, vol. 41, no. 3, pp. 320–331, 2013.

[31] N. Jiang, D. U. Becker, G. Michelogiannakis et al., "A detailed and flexible cycle-accurate network-on-chip simulator," in *2013 IEEE international symposium on performance analysis of systems and software (ISPASS)*, pp. 86–96, Austin, TX, USA, April 2013.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8887285, 12 pages
https://doi.org/10.1155/2020/8887285

Research Article

Presenting an Effective Method to Detect and Track the Broken Path in VANET Using UAVs

Zohreh Bakhtiari,[1] **Rozita Jamili Oskouei**,[1] **Mona Soleymani**,[2] **and Akhtar Hussain Jalbani**[3]

[1]*Department of Computer Science and Information Technology, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran*
[2]*Department of Computer Engineering, Parand Branch, Islamic Azad University, Tehran, Iran*
[3]*IT Department, QUAID-E-AWAM University of Engineering, Science & Technology, Nawabshah, Sindh, Pakistan*

Correspondence should be addressed to Rozita Jamili Oskouei; rozita2010r@gmail.com

Received 19 April 2020; Revised 21 May 2020; Accepted 10 June 2020; Published 17 August 2020

Academic Editor: Fawad Zaman

The routing process in vehicular ad hoc networks (VANETs) is a challenging task in urban areas which is due to the high mobility of vehicles, repetitive defects of the communication path, and the various barriers that may affect the reliability of data transmission and routing. Accordingly, the connectivity in vehicular communications has received the researchers' attention, so different geographic routing protocols have been proposed in this respect. Unmanned aerial vehicles (UAVs) are useful for overcoming routing constraints. Cloud computing has also been defined as a new infrastructure for VANET which is made up of a significant number of computing nodes including stable data centers as well as a set of mobile computing devices embedded on vehicles. The aim of this research is to simulate a VANET in an urban area using cloud computing infrastructure and applying unmanned aerial vehicles (UAV) so that the negative influence of barriers in packet delivery and routing is avoided. To evaluate, the proposed method is compared with the basic protocol ClouDiV. Ns-2 simulation results show that the proposed method outperforms with different densities and variable times in terms of efficiency and performance.

1. Introduction

Vehicular ad hoc network (VANET) is a kind of Ad hoc networks that provide the communication among the adjacent vehicles, as well as the fixed equipment and vehicles that are usually installed along the roads. The analysis of connections in these networks is important in a way that in VANET, due to the high mobility over other ad hoc networks and frequent changes in the network topology, the path break may occur frequently. The previous researches are more about the relationship between the vehicles' connectivity and distance which have been studied in terms of the speed effect, vehicle density, different mobility patterns, radio broadcast range, node rank, connection duration, etc. All of these analyses are effective in the selection of the next hop in routing in order to prevent the path breakage [1]. The routing protocols for VANETs are the main issue of this technology due to the high dynamic nature of the nodes. Since the effective and reliable dissemination is a major challenge in VANET, the geographical routing protocols are flexible for most ITS applications. The characteristics of urban traffic environment differentiate it from the highway traffic area [2]. The weather conditions also have a significant impact on this traffic environment. In urban spaces, the buildings have blocked the direct communication between two vehicles as barriers. This prevents vehicles from exchanging messages even when they are within the same transmission range. The ClouDiV protocol has been proposed as a geographic protocol for message dissemination in vehicular networks based on the cloud computing infrastructure which has recently been implemented in the field of cloud computing to deal with vehicular network issues [3]. The broken paths in VANET will be caused due to the loss of connection or the presence of network gaps and by the use of UAVs or unmanned aerial vehicles (also referred to as remotely controlled bird objects) in the communication systems can

detect these paths and share the connection information among vehicles.

In this paper, the proposed scheme is based on the ClouDiV protocol. In the proposed method, by making changes in the structure of the ClouDiV algorithm, we hope that this modified protocol accompanied with the UAVs in the sky improves the routing performance even with the presence of barriers. The proposed method is evaluated and compared to detect and trace the broken communication paths in VANETs based on the ClouDiV protocol. The evaluations are performed in terms of some criteria such as packet delivery rate, end-to-end latency, throughput, and the number of lost packets.

The organization of this paper is as follows: In Section 2, the background and related works are given. In Section 3, the details of the proposed method are described with the given goals. Section 4 presents the evaluation of the proposed method by analyzing the information and giving the simulation results. Finally, Section 5 gives the conclusion and future works.

2. Related Works

By introducing VANETs and increasing the capabilities of VANET, the consumer applications have been enabled, and the subject of connection analysis has attracted more attention from research communities [4]–[9]. In the following sections, while studying the connection analysis in the VANETs, we are going to discuss the recent works on cloud computing and UAV performance in communications and interconnections of these networks.

2.1. Analysis of Connections and Communication in VANET. Recently, connectivity has attracted the attention of researchers in vehicular communications, and various routing protocols have been proposed in this field. To reflect the high dynamics of the connection due to different patterns of mobility, the effects of speed [4, 6, 9], vehicle density [4, 10], radio broadcasting range [6, 9, 10], node rank [11], and connection duration [12] have been studied.

In [13], Wu has modeled the probability of accurately connecting two nodes with distance L over a linear network by the Poisson point process. In [14], the authors have studied the data dissemination approaches for highway scenarios in vehicular networks using roadside units in order to achieve a communicative coverage among vehicles that can be extended at some levels of their connection where this method is also useful for critical time data dissemination. In [15], a modeling method has been studied in uplink v2v communications for highways and rural, urban, and suburban environments [15]. To analyze the interconnection of wireless vehicular networks, the authors in [16] have examined the maximum number of hops in a multiple-hop path from the source to the destination and the minimum transmission power used to ensure the connection of the vehicles from the physical layer's viewpoint. The test modeling of the vehicle-to-vehicle lost path in urban, suburban, and highways at 5.8 GHz [17] has been carried out by the definition of free space, where the results are similar to the previous values. In [1], Zarei et al. has presented a scheme to analyze the connection for the dynamic motion of wireless vehicular networks that accurately calculated the connection distance for a one-way highway path and proposed a new formula for the time-dependent probability density function (pdf). Also, an analysis of the v2v connection for wireless vehicular networks has been done using mathematical models [18]. In [19], the authors have presented a combined HRAR route-aware routing protocol in urban and highway environments based on a two-level hierarchical routing in terms of road communications in VANETs. An iCAR-ll infrastructure-based and connection aware routing protocol has been proposed in [20]. This protocol consists of a number of algorithms to be run by vehicles, and it has been applied to predict the local network connection and update local servers with real-time network information in order to build a global network topology. In [21], the authors have proposed the Fast MF to extend the Internet access for vehicular nodes and get access to these nodes from any remote server for the connected vehicles under information-based architectures. A cluster-based file transfer scheme for Highway VANETs (CFT) has been presented in [22]. With CFT, when the requested file cannot be successfully transmitted from the source to the destination, a cluster is created under a direct v2v connection and the file is transmitted in several hops. In these researches, three models have been confirmed to study the problem of file transfer in VANETs: (1) high mobility of vehicles, (2) connection time prediction, and (3) v2v communications. In [23], in order to disseminate the connection aware information by estimating the ability of node transmission in nearly connected VANETs (CADDs), first, the connection specifications are examined theoretically and then CADD is presented to improve data dissemination. A connection aware routing protocol of the road network has been presented for wireless vehicular networks in [24]. The protocol tries to avoid the clustering phenomenon under the influence of traffic lights and to select the road section by disseminating information among more vehicles. Li et al. [25] have presented multihop transmission latency and connection probability in the analysis of multihop links from the 5G enabled vehicular networks. In [26], the authors have proposed a reliable IP-based routing in VAENT with network gaps based on road segments and virtual nodes that can form routing paths without network gaps (the network gap occurs when the distance between two nonadjacent vehicles is more than the transmission radius).

For vehicular communications in 700 MHz and 5.9 GHz, under the LOS and NLOS conditions, a simple dissemination model suitable for VAENT simulations has been proposed to evaluate the protocol and system setting in [27], where the lost path has been detected by narrow channel measurement. Liu et al. [28] have performed the path-loss modeling for v2v communications on an uplink path using low-height antennas at downhill and uphill in different scenarios. In [29], the authors have presented the next hop selection scheme based on the remaining lifetime of the link for wireless vehicular networks and proposed an algorithm to predict the remaining link lifetime in VANETs using a Kalman filter based on the prediction technique. [30] have presented a stable path and decision-making scheme to send a data packet

in a highway which allows the prediction of the path's lifetime and remaining time for data packet transmission and removes the path's failure message, since the source vehicle knows the time when the path would be broken. In this scheme, the packet delivery rate is reduced by increasing the network density which is due to this fact that yet no method has been presented to restore the lost data packets.

All these studies and analyses are effective in selecting the next hop to prevent the breakage of the communication path. Also, the strategies of path repair [26], packet broadcast verification [30], and using guards have been proposed to help the process of restoring the lost paths when the network failed [18].

2.2. Cloud Computing in VANET Communication. Recently, some researches have proposed cloud computing to overcome the vehicular network problems [3]. Among them, a new routing scheme, called VehiCloud [31], has been built to overcome the unreliable v2v communications and it develops the constrained computing capabilities of vehicular devices by using cloud computing architecture. This method has suggested that each vehicle predicts its future successes by generating point messages which is the description of vehicle's trajectory. In [32], the integrated Internet access is another issue that has been solved by the cloud computing infrastructure. The vehicular network literature has mentioned another vehicular scheme which is based on cloud, called vehicular cloud model (VC) [33]. It has been defined as a new mobile cloud computing model. The VC simulation study shows that the increased network density leads to lower latency in dissemination. In [34], three main cloud entities are proposed: vehicular cloud (VC), vehicles using cloud (VuC), and hybrid cloud (HC). A new approach of using RSU as a cloud server has been proposed in [35]. To provide safe and unsafe services in the vehicular applications, the vehicular cloud (VCR) has been proposed for roadside scenario in [36].

Bitam and Mellouk [3] have presented cloud computing-based message dissemination protocol for wireless vehicular networks to ensure reliable connections. ClouDiV has been considered a geographical protocol which will disseminate messages through a connection process based on cloud data centers. To evaluate the performance of this protocol, an extensive experimental study has been done in terms of average point to point latency, packet delivery rate, and normalized overhead, and the obtained results have been compared with RIVER protocol.

2.3. UAVs on VANET Connections. Recently, extensive researches have been performed in the field of routing and tracing the communication path between two vehicles in a VANET to disseminate data packets; however, most of them have not considered UAVs to solve the problem of disconnection, frequently occurring in urban areas, and only a number of them have used UAV to overcome the obstacle-based disconnection in urban areas. The authors in [37] have proposed a new mechanism in mobile ad hoc networks that is the usage of UAVs to increase the connection. The major drawback of this proposed protocol is that UAVs have not used GPS information and trajectory (path) measurement during path detection and data transmission. The traffic congestion aware routing based on connection using UAVs for VANETs (CRUV) [38] is a new routing method for VANETs that is able to detect the shortest connected path at any time to effectively send packets to their destinations by combining the real-time traffic congestion in terms of periodic exchange of Hello Messages and Dijkstra algorithm in order to determine the shortest path from the source to the destination. UVAR (VANET routing protocol by UAV) [39] is a new routing protocol in collaboration with the unmanned vehicles in sky to give a general view on the connected segments and improve the routing process. To prevent the previous limitations, UVAR protocol considers the real distribution of the vehicles in the selected parts of the road which not only provides the accurate calculation of the vehicle connection but also solves the current barriers that manipulate the calculations. Despite the use of the general rule as in [40] to send data packets in earth, UVAR has completely abused UAVs, so that the barriers in the sky would not occur anymore.

Data delivery has been optimized by using this combination of communications between the vehicles and UAVs which would help to reduce the loss of packets and delivery latency.

3. An Effective Way to Detect and Trace the Broken Paths in VANET Using UAVs

The main goal of this paper is to prevent the negative impact of barriers in packet delivery and routing in urban areas by using cloud computing infrastructure and UAVs and to detect these paths by using ClouDiV protocol in those parts of the road that the communicative paths are broken and to share the connection information among vehicles by using UAVs (unmanned aerial vehicles) as well as to find the broken paths so that at last the reliable and stable paths can be created. The ClouDiV protocol performs rebroadcast action for data dissemination in some cases, such as disconnection of the communication path that imposes an overhead on the network. The use of unmanned aerial vehicles for routing with ClouDiV will cause the disconnected communication paths to be identified and repaired prior to the data dissemination, and there is no need for a broadcasting operation, so the overhead of this protocol will decrease and the throughput will increase.

In this proposed scheme in some parts and structure of the ClouDiV protocol will be modified so that this protocol would have a general view on the connected parts with the UAVs in the sky and even with the presence of barriers (buildings, high trees, etc.), the routing would be improved. In this method, UAVs share the connection information for all VANET nodes and data centers and help updating the routing tables of data centers, and they are applied in the parts of the road where there is a problem of barriers to send data packets as other VANET nodes. In the proposed scheme, the packet delivery will be improved by decreasing the delivery latency and packet loss.

Nodes in VANET make traffic and motion and have high dynamicity; for this, the connection path between the source and destination nodes is frequently disconnected. In a

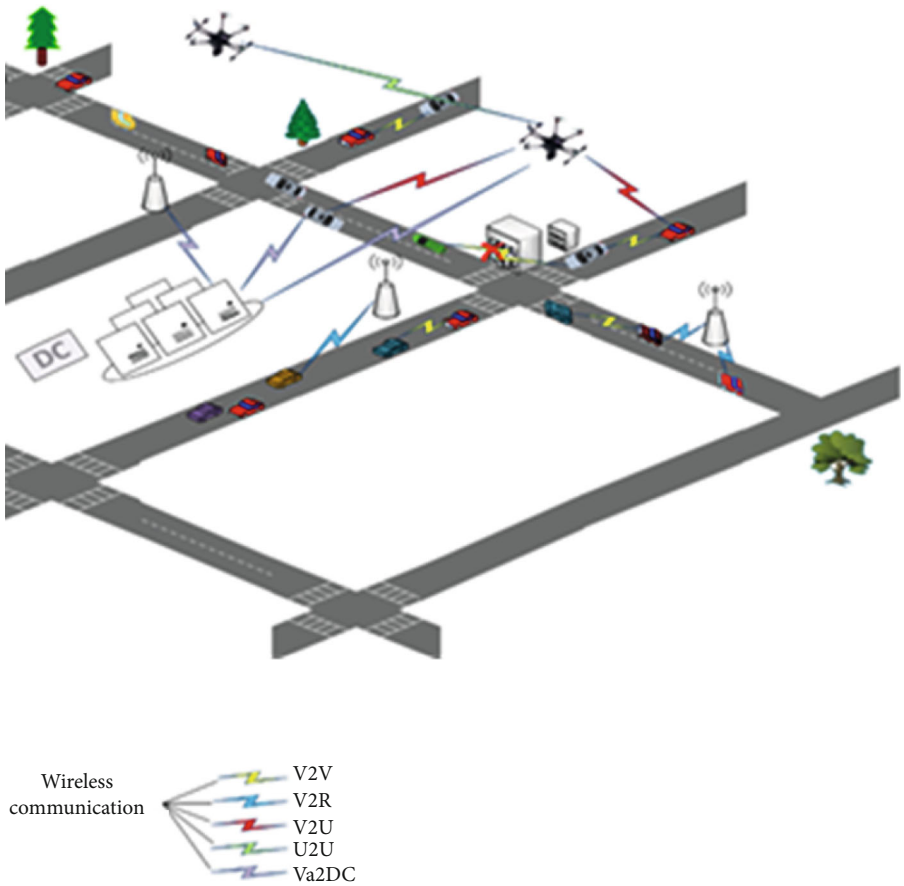


Figure 1: Vehicle and UAV communication.

VANET, the mere existence of a multihop path from source to destination cannot always ensure that linkage from source to destination is successful, since the wireless vehicular challenges are failure prone. To ensure the reliable information dissemination in a critical time, the network connection must be guaranteed. Due to the fact that in VANETs, the mobility with high speed will lead to the frequent changes in the network topology, these changes will increase the rate of frequent disconnection and packet loss during dissemination and furthermore; the packet loss is inevitable due to the network gaps. Therefore, the detection of the disconnected path in VANETs could be used to reduce the rate of packet loss and transmission delay in the geographical routing.

Due to the variable speed of moving vehicles and vehicle movements that are limited by roads and traffic lights, the main goals to be considered are maintaining the created connection, preventing the failure of the link, which happens frequently and reducing the process of disconnection and reconnection of the node through a new path to increase the network performance, reduce packet loss rates, and improve connectivity as well as reduce the loss of communication path in VANET by detecting the broken communication path in the geographical routing process. Also, the consideration of the characteristics of the urban traffic environment is aimed at reducing the negative impact of urban barriers on routing protocols. Finally, the ultimate goals for this study include reducing the road accidents and damages caused by them, improving roads and streets' safety, and improving the environmental conditions. We also hope that this project will be used and exploited by the ministry of roads and urban development and automobile companies accompanied with road transportation authority.

In order to outline the proposed method, we will first provide the architecture, and then, along with a review of the ClouDiV protocol, the changes that made to this protocol will be explained.

3.1. The Proposed System Architecture. As shown in Figure 1, this system consists of a set of vehicles, UAVs, roadside units, and data centers distributed over the network. In this hybrid communicative system, the IEEE 802.11P MAC protocol has been confirmed for v2v communications, UAV-to-vehicle communications, and VANET node communications (vehicle, UAV, and RSU) with data center. Five types of wireless communications are used in this system:

Communications include vehicle to vehicle (V2V), vehicle to roadside units (V2R), vehicle to UAV (V2U), unmanned aerial vehicles to unmanned aerial vehicles (U2U) [39], and VANET node with data center (Vn2DC). Data centers are considered the fixed nodes of cloud computing, as cloud servers, and have wireless communication with VANET nodes. However, the data center communication is wired.

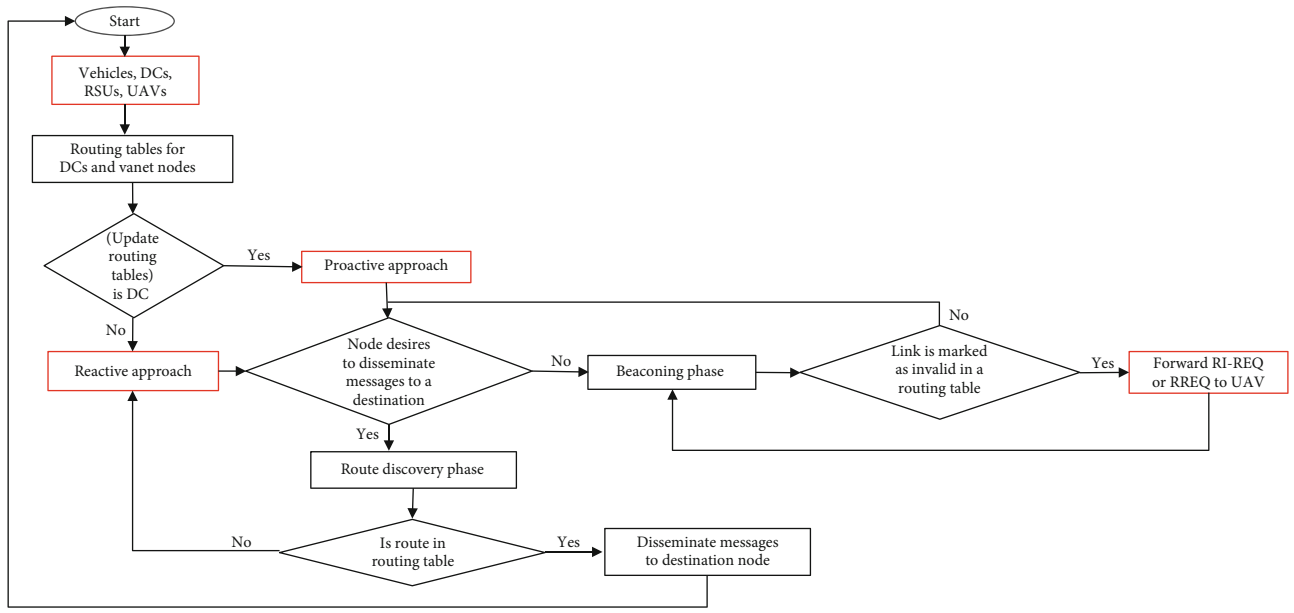


Figure 2: The proposed method flowchart.

3.2. Overview on ClouDiV Protocol Performance. ClouDiV [3] is a message dissemination protocol for vehicular networks based on cloud computing architecture which is considered a geographical protocol. ClouDiV has combined a fixed cloud computing infrastructure (e.g., data centers) and a flexible cloud structure based on computers embedded on the vehicles.

In this protocol, it is assumed that the vehicular network consists of a number of different VANET nodes (e.g., vehicles and roadside units); also, different cloud servers such as data centers are placed in the traffic environment. ClouDiV is considered a hybrid message dissemination protocol where a proactive approach is applied by each data center in order to detect new and updated paths of each node. In addition, a reactive approach is defined to be performed by each vehicle with the aim of finding the nearest data center as an intermediate node. Also, two types of routing tables are proposed: data center routing table and VANET node routing table. Each points out to using a particular type of a routing process. Beaconing phase with the aim of ensuring the routing tables' updating by Beacon packets transmitted periodically. In path detection phase, to detect a path, two complement processes are applied: proactive routing and reactive routing.

3.3. The Proposed Method. Some parts of the ClouDiV protocol have been modified to reach the given goals (Figure 2). The data center can send its request to the UAVs at the transfer range to update its routing table in the proactive routing when the control packet of requesting the path information (RI-REQ) with the aim of collecting different information of routing reaches to a node that could not respond to its requested information; in other words, it cannot find a node in its neighboring area over the earth. Since in this protocol, the routing information among VANET nodes can be stored and updated without cloud computing structure (beaconing phase); in this method, the VANET node can collect information from the neighbor nodes to prevent recording invalid connections and to cause packet failure by using UAVs. In other words, if it cannot receive the information from its neighbor node in a given time period, it can perform this action by UAVs. Also, in this scheme, in the phase of detecting a path in the reactive routing, the RREQ packet can be sent to the UAV placed within the given range instead of broadcasting the route request control packet (RREQ) by the source node which helps to significantly reduce the routing overhead.

The routing tables of the proposed method are the same as the ClouDiV protocol as well as the data center routing table and VANET node routing table. The ClouDiV routing process will be performed in the path detection phase which will be performed with slight changes in the proposed method.

3.3.1. Proactive Approach. The proactive routing process in this approach is performed by each data center in the network with the aim of tracing all paths from this data center toward each node in the network (known as a future destination) [3]. If the node does not receive the information from the neighbor nodes (VANET nodes on the earth) in a given time interval, a process as beaconing phase (Figure 3) is performed in the broken path detection and broadcasting it by UAV, and then, each node receiving RI-REQ generates a RI-REP so that to make communication, it sends its updated routing information to DC1. After receiving all RI-REPs, DC updates it routing table and directs the input paths toward different destinations.

3.3.2. Reactive Approach. The reactive approach is performed by each vehicle and UAV with the aim of finding the nearest data center as an intermediate node; each one tries to provide a new path for the given destination. In the reactive routing process, a path request packet (RREQ) is generated based

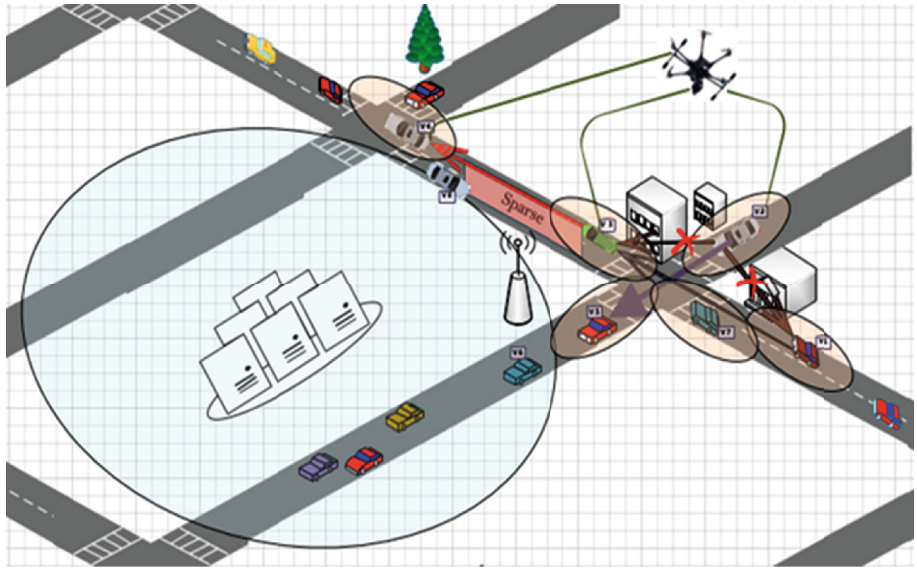


Figure 3: Disconnection of vehicles.

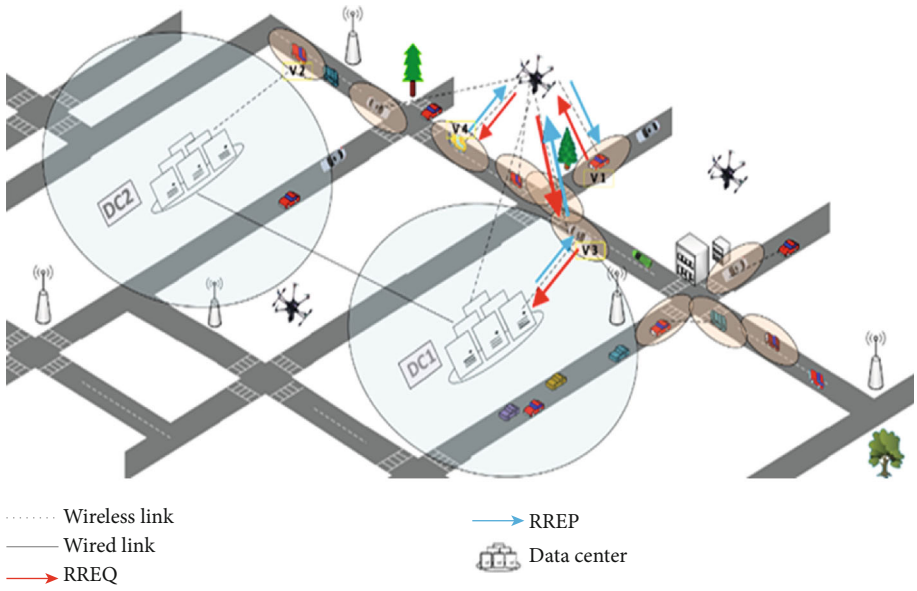


Figure 4: The reactive routing process of the proposed scheme.

on the request of the source node V1 to disseminate messages to a destination node V2 (Figure 4). In the proposed method, the RREQ is sent to the UAV within the transmission range by the source node instead of randomly broadcasted by the source node among the neighbors at that moment. The UAV node checks its routing table, and if there is a data center in its routing table, then it generates a routing response packet (RREP) and sends it to the source node; otherwise, the RREQ packet is broadcasted by UAV to the neighbor nodes within its range randomly. Here, it is assumed that the vehicles V3 and V4 have received the RREQ packet. The first data center, like DC1, has been detected by V3. As the packet has received to DC1, this data center generates a RREP and sends it to V1 across the opposite direction; therefore, the response packet will be sent to the source node along the same path by the generated node. Similarly, the RREP packet records the next hop's node in the routing table of each visited node so that the sender represents it until it reaches the source node. As a result, the source node can begin to disseminate data packets to detect the data center through the next hop's node, which has been found by the reactive routing process. Then, the DC1 data center will send these data packets to the final destination, the V2 node from (DC1-DC2-V2) path detected by the last proactive routing process.

3.3.3. Beaconing Phase. In this step, the active connections between each VANET node and data center will be updated

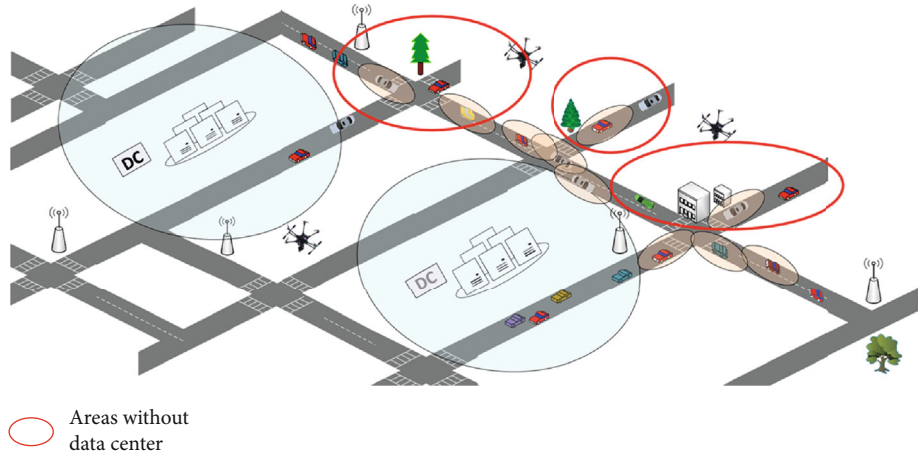


Figure 5: Updating routing information among VANET nodes.

by loading and updating different transmission parameters, such as measured available bandwidth and end-to-end latency. Also, the routing information among VANET nodes will also be stored and updated to perform a common routing (e.g., without the help of the cloud computing structure) in the absence of data centers [3] (Figure 5).

The main function of beaconing phase is described below.

(1) Detecting the Broken Paths. During an update, if an invalid connection is marked on a routing table, that is, one node does not receive information from neighboring nodes (VANET nodes on Earth) in a given time internal, instead of broadcasting an unavailable connection by other nodes using the error package, in the proposed scheme, according to the scenario shown in Figure 3, there is less vehicular density and the network is sparse (V3 and V4); and also, there are high buildings, which break the connection between V2 and V1, as well as V2 and V3; in such cases the earth nodes (e.g., V2) will send the path delivery information to the UAV node within the given range. The UAV node first sends the connection information to the given node in the packet response format according to its routing table, and then, the UAV will send the path error packet to the nearest data center (DC) in order to inform the data center about disconnection, so that finally, the data center will be able to replace the new paths in that range by updating its routing table. Therefore, if a request is sent to the UAV, it indicates a disconnection that should be solved by replacing the new route.

4. Evaluation

To evaluate the proposed method, its overall structure is presented to simulate and examine the parameters and related criteria as follows:

This scheme consists of a cloud computing infrastructure for the VANET, including a set of fixed computing nodes such as data centers and flexible computing nodes, computers embedded on vehicles, and UAVs. In general, in this method, it is assumed that the vehicular networks are located in the traffic environment by different VANET nodes (such as vehicles, roadside units, and UAVs), as well as different cloud servers (e.g., data centers). To evaluate this scheme, network simulators such as ns2 and greencloud are used on urban scenarios.

To simulate this system and parameter settings, using a modified ClouDiV protocol, a moving pattern represents the region of an assumptive $1 * 1\,km^2$ city with several intersections (Figure 6). In the studied network, two scenarios are considered; in both scenarios, there are a number of nodes moving between 20 and 160 where the vehicles are moving at speeds of 1 to 50 km/h. Also, a number of RSUs are distributed uniformly and a data center is used in the center of the studied area with a 1000 m transmission range. According to IEEE 802.11p standard, wireless accessibility in the vehicular space has been proposed for VANETs. Each vehicle has a long Wi-Fi interface of about 250 m to communicate with other vehicles over the same street. However, there cannot be any communication among the vehicles located on different streets due to the presence of barriers. These scenarios are simulated at variable times of 200 to 1000 ms. The traffic flow in this simulation is CBR (UDP), and only in one of these two scenarios, a number of UAVs with 1000 m communication range and 40 to 90 km/h rate are considered. The first scenario is simulated without UAVs and with ClouDiV protocol, and the second scenario that is the proposed scheme of this article is simulated with UAVs and the modified ClouDiV protocol, in which despite the disconnection caused by barriers and sparsity in some parts of the road, in both scenarios, to evaluate the performance of the proposed scheme, the parameters of the package delivery rate, end-to-end latency, throughput, and the number of the lost packets can be compared.

4.1. Evaluation Criteria. Four criteria are considered in the evaluation process:

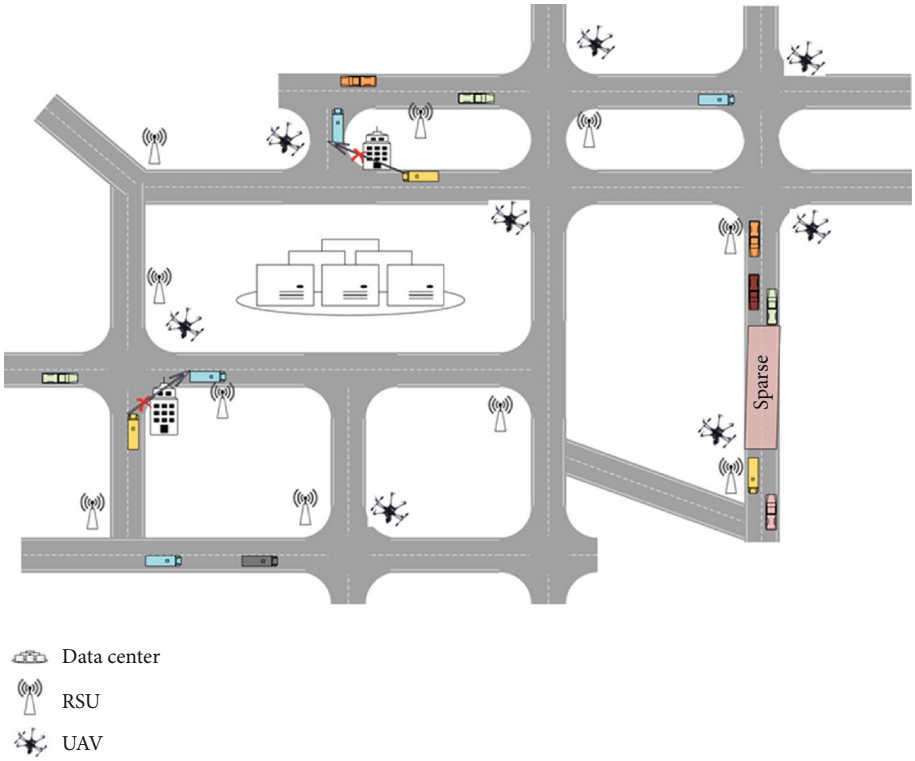


Figure 6: District map of the assumptive city.

Package delivery ratio: in VANETs, packets are usually sent hop-by-hop or via multiple hops. The delivery ratio of such network is calculated by

$$\text{PDR} = \left(\frac{\text{Number of sent packet}}{\text{Number of received packet}}\right) * 100. \quad (1)$$

Lost packets: packet loss in VANET networks will occur for various reasons; for example, the sent packets in the network fail to reach the destination node, or they will be lost due to a bit error or hardware failures. The presence of noise in the network can also cause the loss of packets. The loss of a number of packets in the mobile ad hoc networks can be caused by the attacks existing in such networks. This parameter usually is represented by percentage:

$$\text{Lost packets} = \frac{(\text{Sent packet} - \text{Received packet})}{\text{Sent packet}}. \quad (2)$$

According to (2), the number of the lost packets can be obtained in the VANET network.

End-to-end latency: in VANETs, a time interval in which information packets are transmitted over the network from the source node to the destination node is called an end-to-end latency. This latency ends up as long as packets are transmitted from the source node to the destination node throughout the network. This parameter is represented in milliseconds.

Throughput: the reason for measuring throughput in networks is that people often tend to know the maximum data throughput in a communication link or network access. The common method for measuring this parameter is to transfer a large file from one system to another and calculate the required time to complete file transmission or copy. Then, the throughput will be achieved by dividing the file size at that time, in Mbit/s, Kbit/s, or bytes per second. The following equation calculates the throughput in a network:

$$X = \frac{C}{T}. \quad (3)$$

In (3), X represents the throughput, C represents the number of requests completed by the system, and T is the total time in which the system has been monitored.

4.2. Simulation Result Analysis. In this section, we will study and compare the data and information obtained from the proposed scheme and the ClouDiV protocol, which are based on the cloud computing structure. These data have been obtained based on the simulation by the NS-2 simulator, which is the best simulator for ad hoc networks for its flexibility and software efficiency, and it is considered a very powerful simulator for wireless vehicular networks.

At each simulation, one can change a number of parameters, such as simulation time, and evaluate the simulation results. Initially, evaluations are performed under various vehicle densities, with vehicle densities between 20 and 160 where UAV has also been taken into account as a VANET

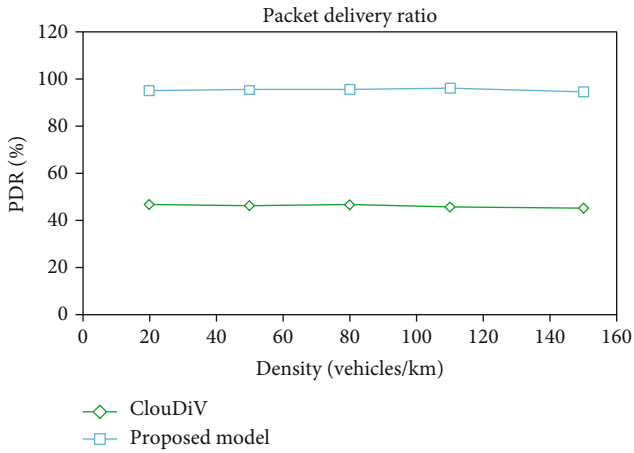


Figure 7: Packet delivery ratio with different densities.

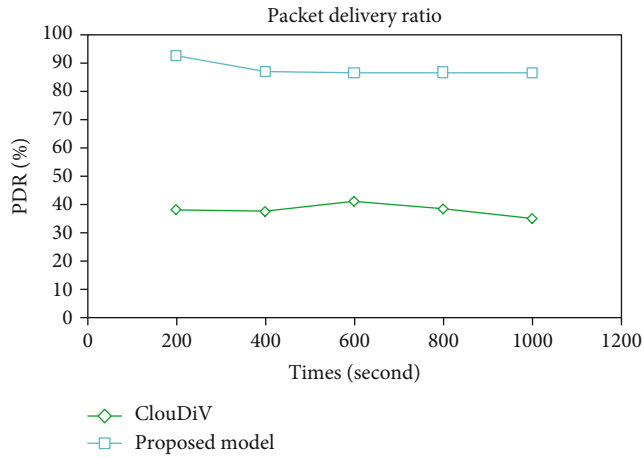


Figure 8: Packet delivery ratio at variable times.

node but with three-dimensional coordinates due to its height. Then, we have performed the tests and evaluations through simulations at different times, where we have considered the time intervals between 200 and 1000 milliseconds. The simulation results are presented as follows:

(i) Packet delivery ratio: in this section, we compare the packet delivery ratio under different vehicle densities and at variable times for the proposed method and the evaluated ClouDiV protocol. As shown in Figure 7, the simulation results show that the delivery ratio in the proposed method is higher than ClouDiV which is due to the usage of paths with ultraband-width provided by cloud-based infrastructures.

Also, according to Figure 8, it can be seen that the proposed model has a better performance in terms of a packet delivery ratio at different times. However, ClouDiV only uses a broadcasting method to find the route to the destination node, which will be time-consuming, and the packet delivery rate will be lower than the proposed method.

(ii) End-to-end latency: Figure 9 shows that the end-to-end latency of different vehicle densities is lower for the proposed method compared to the evaluated protocol. This is due to the use of UAVs in the range, which has led to the selection of reliable paths. Therefore, short distances and optimal paths through which data packets are transmitted to the destination significantly reduce the end-to-end latency while ClouDiV adds additional time when buffering the packet according to the path detection process before delivering the packet to the destination.

And, according to Figure 10, the end-to-end latency at variable times for the proposed model is better than the ClouDiV protocol. In the proposed scheme, the packets are delivered more quickly, since UAVs occasionally choose a shorter distance to deliver the packet to the destination, while in ClouDiV, it is possible that the requested path cannot be detected by the source node in the path detection process.

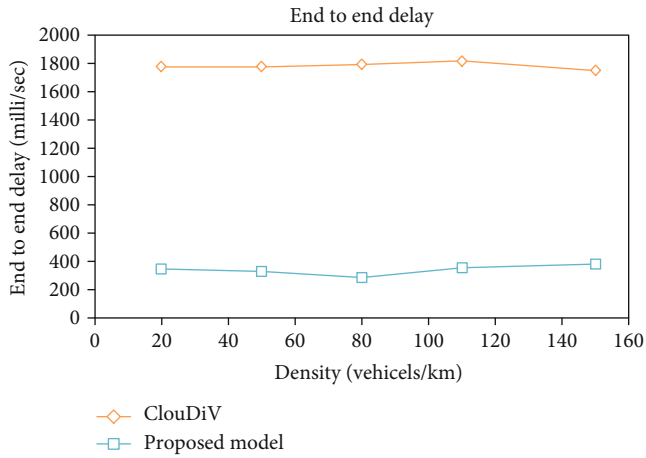


Figure 9: End-to-end latency with different densities.

(iii) Throughput: Figures 11 and 12 show that the proposed method has a better performance. We have obtained the throughput with different number of nodes at different times. However, the proposed model represents that at different times and vehicle densities, the proposed method has a better performance in comparison with the ClouDiV protocol. According to this fact that in the proposed model, the path detection processes, especially in the Reactive routing process, are carried out through a number of UAVs, so more requests are completed, which is better than ClouDiV protocol that needs to produce routing packets while delivering data packets, and requests require more time to be completed.

(iv) The number of lost packets: Figure 13 shows that the number of lost packets is reduced in the proposed method. Since in the proposed method the packets have a greater chance of reaching destinations, as sometimes, the unmanned aerial vehicles (UAVs) select a shorter distance and a reliable path to deliver

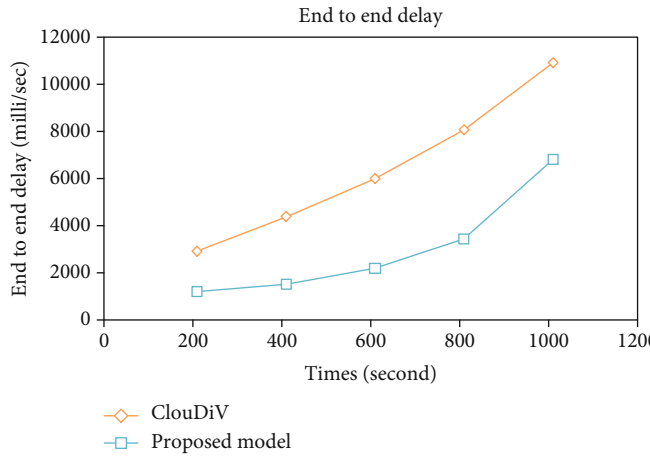


Figure 10: End-to-end latency at variable times.

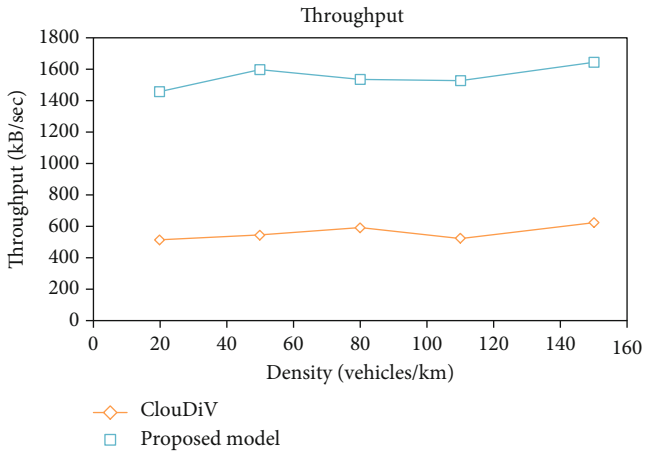


Figure 11: Throughput with different densities.

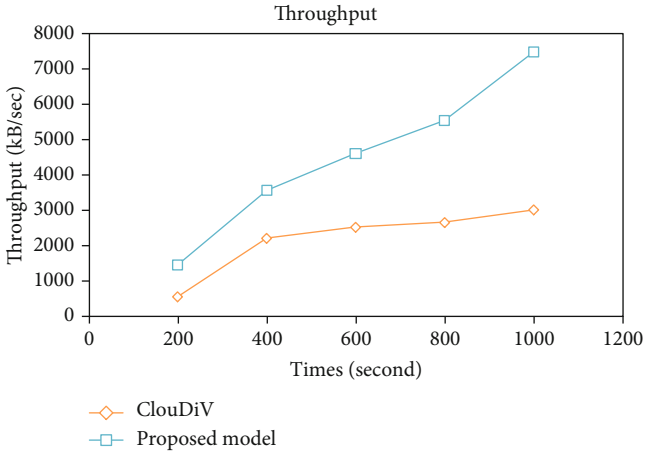


Figure 12: Throughput at variable times.

the packet to the destination, even when the network is dispersed or there is not any node in the neighborhood of the source, while in ClouDiV, it is possible that the requested path by the source node has not yet been found in the path detection process, or the route to the destination has more hops, so it is more likely that the packet is lost

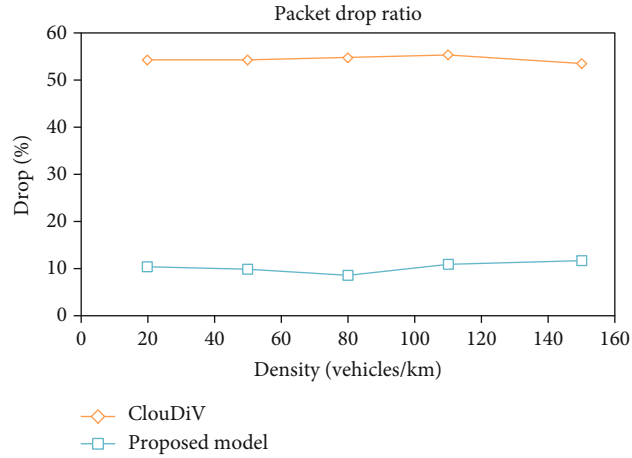


Figure 13: The number of lost packets with different densities.

However, according to Figure 14, the proposed model has acceptable results at different times which is due to this fact that in ClouDiV protocol, when the communication path is broken at any time for any reason from the source to the destination, it is likely that the data packets are lost, while in the proposed method, by detecting the disconnected paths before data dissemination, the probability of losing data packets is reduced significantly and this implication reflects the better performance of the proposed model.

5. Conclusion and Future Works

In this paper, first, the connections in VANETs have been presented, and the use of cloud computing and UAVs in VANET communications has been given to deal with the problems of the vehicular network. Then, the proposed method has been proposed by modifying the ClouDiV protocol and, in fact, using the cloud computing infrastructure with utilization unmanned aerial vehicle (UAV) in urban

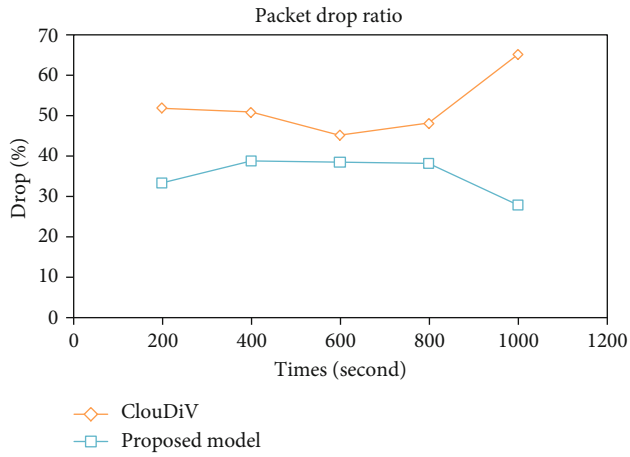


FIGURE 14: The number of lost packets at variable times.

areas. The simulation results show that the proposed method has better performance and efficiency in terms of improving packet delivery ratio, throughput, end-to-end latency, and the number of lost packets. However, UAVs can prevent the generation of additional routing packets, thus reducing network overhead. Also, in the proposed method, the UAVs will share connection information for VANET nodes on the earth and on data centers when there is no node in their neighborhood or for any reason they cannot communicate with each other. Therefore, the unmanned aerial vehicles (UAVs) will be able to be applied in the detection of broken paths.

The proposed method in this paper can be used to improve work in future researches on other geographic routing protocols in VANET networks or to be adapted to other environments such as highways and rural areas. The proposed method can also be improved on the security issue in VANET networks so that the broken connections can be better detected by examining the most important attacks on such network during routing.

Data Availability

This work is simulation.

Conflicts of Interest

We declare that we do not have any conflict of interest.

References

[1] M. Zarei, A. M. Rahmani, and H. Samimi, "Connectivity analysis for dynamic movement of vehicular ad hoc networks," *Wireless Networks*, vol. 23, no. 3, pp. 843–858, 2017.

[2] H. Akhtar and S. C. Sharma, "Performance Evaluation of Location-Based Geocast Routing using Directed Flooding Rectangular Forwarding Zone in City VANET," *International Journal of Engineering and Technology Innovation*, vol. 5, no. 4, pp. 264–278, 2015.

[3] S. Bitam and A. Mellouk, "Cloud Computing-Based Message Dissemination Protocol for Vehicular Ad Hoc Networks," in *International Conference on Wired/Wireless Internet Communication*, M. Aguayo-Torres, J. Gómez, and A. Poncela, Eds., vol. 9071, pp. 32–45, Springer, Cham, 2015.

[4] Z. Zhang, G. Mao, and B. D. O. Anderson, "On the Information Propagation Process in Mobile Vehicular Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 5, pp. 2314–2325, 2011.

[5] A. Agarwal, D. Starobinski, and T. D. C. Little, "Phase transition of message propagation speed in delay-tolerant vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 249–263, 2012.

[6] Z. Zhang, G. Mao, and B. D. O. Anderson, "Stochastic characterization of information propagation process in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 122–135, 2014.

[7] E. Baccelli, P. Jacquet, B. Mans, and G. Rodolakis, "Highway vehicular delay tolerant networks: information propagation speed properties," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1743–1756, 2012.

[8] L. Cheng and S. Panichpapiboon, "Effects of intervehicle spacing distributions on connectivity of VANET: a case study from measured highway traffic," *IEEE Communications Magazine*, vol. 50, no. 10, pp. 90–97, 2012.

[9] S. Durrani, X. Zhou, and A. Chandra, "Effect of Vehicle Mobility on Connectivity of Vehicular Ad Hoc Networks," in *2010 IEEE 72nd Vehicular Technology Conference - Fall*, pp. 1–5, Ottawa, ON, 2010.

[10] M. Khabazian and M. Ali, "A Performance Modeling of Connectivity in Vehicular Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2440–2450, 2008.

[11] A. Cardote, S. Sargento, and P. Steenkiste, "On the connection availability between relay nodes in a VANET," in *2010 IEEE Globecom Workshops*, pp. 181–185, Miami, FL, 2010.

[12] R. Nagel, "The effect of vehicular distance distributions and mobility on VANET communications," in *2010 IEEE Intelligent Vehicles Symposium*, pp. 1190–1194, San Diego, CA, 2010.

[13] J. Wu, "Connectivity of mobile linear networks with dynamic node population and delay constraint," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1218–1225, 2009.

[14] P. Tomar, B. K. Chaurasia, and G. S. Tomar, "State of the art of data dissemination in VANETs," *International Journal of Computer Theory an d Engineering*, vol. 2, no. 6, pp. 957–962, 2010.

[15] J. Karedal, N. Czink, A. Paier, F. Tufvesson, and A. F. Molisch, "Path loss modeling for vehicle-to-vehicle communications," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 323–328, 2011.

[16] P. C. Neelakantan and A. V. Babu, "Connectivity Analysis of Vehicular Ad Hoc Networks from a Physical Layer Perspective," *Wireless Personal Communications*, vol. 71, no. 1, pp. 45–70, 2013.

[17] O. Onubogu, K. Ziri-Castro, D. Jayalath, K. Ansari, and H. Suzuki, "Empirical vehicle-to-vehicle pathloss modeling in highway, suburban and urban environments at 5.8 GHz," in *2014 8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pp. 1–6, Gold Coast, QLD, 2014.

[18] G. Yan and D. B. Rawat, "Vehicle-to-vehicle connectivity analysis for vehicular ad-hoc networks," *Ad Hoc Networks*, vol. 58, pp. 25–35, 2017.

[19] Z. Amjad, W. Song, and K. Ahn, "Two-Level Hierarchical Routing Based on Road Connectivity in VANETs," in *2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA)*, pp. 1–5, Jeju, 2016.

[20] N. Alsharif and X. Shen, "iCAR-II: Infrastructure-Based Connectivity Aware Routing in Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4231–4244, 2017.

[21] M. Tavan, R. D. Yates, and D. Raychaudhuri, "Connected vehicles under information-centric architectures," in *2016 IEEE Vehicular Networking Conference (VNC)*, pp. 1–8, Columbus, OH, 2016.

[22] Q. Luo, C. Li, Q. Ye, T. H. Luan, L. Zhu, and X. Han, "CFT: A Cluster-based File Transfer Scheme for highway VANETs," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, 2017.

[23] Z. Li, Y. Song, and J. Bi, "CADD: connectivity-aware data dissemination using node forwarding capability estimation in partially connected VANETs," *Wireless Networks*, vol. 25, no. 1, pp. 379–398, 2019.

[24] H. Qin and C. Yu, "A road network connectivity aware routing protocol for Vehicular Ad Hoc Networks," in *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pp. 57–62, Vienna, 2017.

[25] S. Li, Z. Li, X. Ge, J. Zhang, and M. Jo, "Multi-hop links quality analysis of 5G enabled vehicular networks," in *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Nanjing, 2017.

[26] X. Wang, D. Wang, and Q. Sun, "Reliable routing in IP-based VANET with network gaps," *Computer Standards & Interfaces*, vol. 55, pp. 80–94, 2018.

[27] H. Fernandez, L. Rubio, V. M. Rodrigo-Penarrocha, and J. Reig, "Path Loss Characterization for Vehicular Communications at 700 MHz and 5.9 GHz Under LOS and NLOS Conditions," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 931–934, 2014.

[28] P. Liu, D. W. Matolak, B. Ai, and R. Sun, "Path loss modeling for Vehicle-to-Vehicle communication on a slope," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2954–2958, 2014.

[29] S. Shelly and A. V. Babu, "Link residual lifetime-based next hop selection scheme for vehicular ad hoc networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, 2017.

[30] M. Nabil, A. Hajami, and A. Haqiq, "A stable route prediction and the decision taking at sending a data packet in a highway environment," in *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications*, pp. 1–6, Tetouan, Morocco, 2017.

[31] Y. Qin, D. Huang, and X. Zhang, "VehiCloud: Cloud Computing Facilitating Routing in Vehicular Networks," in *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, pp. 1438–1445, Liverpool, 2012.

[32] Y.-W. Lin, J.-M. Shen, and H.-C. Weng, "Cloud-Supported Seamless Internet Access in Intelligent Transportation Systems," *Wireless Personal Communications*, vol. 72, no. 4, pp. 2081–2106, 2013.

[33] S. Olariu, T. Hristov, and G. Yan, "The Next Paradigm Shift: From Vehicular Networks to Vehicular Clouds," in *Mobile Ad Hoc Networking: Cutting Edge Directions*, S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, Eds., John Wiley & Sons, Inc, Hoboken, NJ, USA, Second edition, 2013.

[34] R. Hussain, J. Son, H. Eun, S. Kim, and H. Oh, "Rethinking Vehicular Communications: Merging VANET with cloud computing," in *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pp. 606–609, Taipei, 2012.

[35] K. Mershad and H. Artail, "Finding a STAR in a Vehicular Cloud," *IEEE Intelligent Transportation Systems Magazine*, vol. 5, no. 2, pp. 55–68, 2013.

[36] D. Baby, R. D. Sabareesh, R. A. K. Saravanaguru, and A. Thangavelu, "VCR: Vehicular Cloud for Road Side Scenarios," in *Advances in Computing and Information Technology*, N. Meghanathan, D. Nagamalai, and N. Chaki, Eds., vol. 178 of Advances in Intelligent Systems and Computing, pp. 541–552, Springer, Berlin, Heidelberg, 2013.

[37] M. Le, J. Park, and M. Gerla, "UAV Assisted Disruption Tolerant Routing," in *MILCOM 2006 - 2006 IEEE Military Communications conference*, pp. 1–5, Washington, DC, 2006.

[38] O. S. Oubbati, A. Lakas, N. Lagraa, and M. B. Yagoubi, "CRUV: Connectivity-based traffic density aware routing using UAVs for VANets," in *2015 International Conference on Connected Vehicles and Expo (ICCVE)*, pp. 68–73, Shenzhen, 2015.

[39] O. S. Oubbati, A. Lakas, F. Zhou, M. Güneş, N. Lagraa, and M. B. Yagoubi, "Intelligent UAV-assisted routing protocol for urban VANETs," *Computer Communications*, vol. 107, pp. 93–111, 2017.

[40] O. S. Oubbati, A. Lakas, N. Lagraa, and M. B. Yagoubi, "UVAR: An intersection UAV-assisted VANET routing protocol," in *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–6, Doha, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8846948, 11 pages
https://doi.org/10.1155/2020/8846948

Research Article

Support Vector Machine-Based Classification of Malicious Users in Cognitive Radio Networks

Muhammad Sajjad Khan,[1,2] Liaqat Khan,[1] Noor Gul,[1] Muhammad Amir,[1] Junsu Kim,[2] and Su Min Kim[2]

[1]*Department of Electrical Engineering, Faculty of Engineering and Technology, International Islamic University, Islamabad 44000, Pakistan*
[2]*Department of Electronics Engineering, Korea Polytechnic University, 237 Sangidaehak-ro, Siheung-si, Gyeonggi-do 15073, Republic of Korea*

Correspondence should be addressed to Su Min Kim; suminkim@kpu.ac.kr

Received 30 April 2020; Revised 23 June 2020; Accepted 7 July 2020; Published 18 July 2020

Academic Editor: Farman Ullah

Cognitive radio is an intelligent radio network that has advancement over the traditional radio. The difference between the traditional and cognitive radio is that all the unused frequency spectrum is utilized to the best of available resources in the cognitive setup unlike the traditional radio. The main role of cognitive radio is spectrum sensing, in which the secondary users (SUs) opportunistically access the spectrum while avoiding interference to the primary user (PU) channel. Various aspect of the spectrum sensing problem are studied from cognitive radio perspective. Cooperative spectrum sensing in cognitive radio has a promising performance compared to the individual sensing. However, the existence of the malicious users (MUs) highly degrades the performance of the cognitive radio network (CRN) by sending falsified results to the fusion center (FC). In this paper, we proposed a machine learning algorithm called support vector machine (SVM) to classify normal SUs and MUs in the network. SVM is used for both classification and regression, but mostly it is used for classification problems. SVM clearly classify both normal and MUs by drawing hyper plane on the base of maximal margin. The results of the legitimate SUs are combined at the FC by utilizing Dempster-Shafer (DS) evidence theory. The effectiveness of the proposed scheme is demonstrated through simulation by comparing with the other existing schemes.

Cognitive radio is an intelligent radio network that has advancement over traditional radio. The difference between the traditional radio and the cognitive radio is that all the unused frequency spectrum can be utilized to the best of available resources in the cognitive radio unlike the traditional radio. The core technology of cognitive radio is spectrum sensing, in which secondary users (SUs) opportunistically access the spectrum while avoiding interference to primary user (PU) channels. Various aspects of the spectrum sensing have been studied from the perspective of cognitive radio. Cooperative spectrum sensing (CSS) technique provides a promising performance, compared with individual sensing techniques. However, the existence of malicious users (MUs) highly degrades the performance of cognitive radio network (CRN) by sending falsified results to a fusion center (FC). In this paper, we propose a machine learning algorithm based on support vector machine (SVM) to classify legitimate SUs and MUs in the CRN. The proposed SVM-based algorithm is used for both classification and regression. It clearly classifies legitimate SUs and MUs by drawing a hyperplane on the base of maximal margin. After successful classification, the sensing results from the legitimate SUs are combined at the FC by utilizing Dempster-Shafer (DS) evidence theory. The effectiveness of the proposed SVM-based classification algorithm is demonstrated through simulations, compared with existing schemes.

1. Introduction

With burgeoning wireless technologies, the demand of spectrum is increasing consistently, which yields scarcity in spectrum resource. Previous assumptions on crisis of spectrum availability result in misconception. By the federal communication commission (FCC), it has been resolved that underutilization of licensed spectrum bands in either temporal or spatial is a principal reason of the spectrum scarcity [1]. To efficiently utilize the spectrum resource, cognitive radio (CR) with adaptive intelligence is fascinating researchers and developers to break through a spectrum congestion bottleneck [2]. CR is an intelligent wireless communication technology with efficient spectrum utilization, trying to learn environments and adjust its parameters properly [2]. Licensed primary users (PUs) can transmit at any time with no restrictions, while secondary users (SUs) in CR networks (CRNs) obtain the benefit of spectrum access when the PUs do not use the corresponding spectrum [3].

Spectrum sensing is one of the important parts of CRN. Such far, various sensing techniques such as cyclo-stationary-based sensing, waveform-based sensing, and energy detection-based sensing were proposed and utilized for spectrum sensing [4]. Among these techniques, the energy detection is the most efficient technique when no prior information of the PU is available. Individual sensing at each SU is often inaccurate due to multipath fading, shadowing, and hidden terminal problems in wireless environments [5]. These are able to cause incorrect detections in PU activity, which result in false alarm and thus reduce the SUs' opportunities to access the spectrum. Similarly, any misdetection of the occupied PU channel in CRN can produce interference to the licensed PUs. To overcome these issues, cooperative spectrum sensing (CSS) was proposed. It significantly improves the accuracy of detection of PU activity and helps to increase the performance of secondary communication systems [6–8]. The CSS is able to be implemented in either distributed or centralized manner. In distributed spectrum sensing, each SU individually senses the spectrum and decides whether the spectrum is available or not. In centralized spectrum sensing, a number of SUs form a network and send their local sensing results (either 0 or 1) to the fusion center (FC) in order to decide the existence of PUs. The final decision regarding the existence of PUs is made based on the information received from all the SUs by using AND, OR, and majority rules [9].

However, CSS is also vulnerable to security threats. Security for CRN is an important part to ensure secure operations of underlying network infrastructure [10]. Various attacks, which highly degrade the performance of network, have been studied in the literature. Two most common attacks in CRN are primary user emulation attack (PUEA) and spectrum sensing data falsification (SSDF) attacks [11, 12]. In PUEA, some outliers try to mimic data transmission of the PU to disturb sensing operations of SUs. The presence of PUEA makes the FC decide that the spectrum band under consideration is unavailable, and SUs hold their processes for opportunistic spectrum access. In SSDF attacks, false information is sent to the FC that leads an incorrect global decision on the PU channel activity. In [13], six types of SSDF attacks are elaborated. In always yes MU (AYMU) attack, the SU always sends "1" to the FC whatever a local result is determined. Hence, this attack denies the SU to access the spectrum. On the contrary, in always no MU (ANMU) attack, the SU always sends "0" to the FC. Thus, it causes interference to the PU channel. In always opposite MU (AOMU) attack, the MU sends the inverse of the local sensing result. It is the most dangerous attack. In random yes MU (RYMU) attack, the MU randomly sends "1" to the FC, regardless of the local sensing results. In random no MU (RNMU) attack, the MU randomly sends "0" to the FC, regardless of the local sensing results. In random opposite MU (ROMU), the MU randomly sends the inverse of the local sensing result to the FC. To mitigate the effect of these attacks, several different schemes were proposed [14–16].

Some heuristic approaches in CSS can lead to an optimal global decision. Among them, a genetic algorithm (GA), a class of computational algorithm motivated by evolution, is a good candidate to find the optimal solution by applying bioinspired approaches to given problems [17, 18]. On the other hand, a machine learning (ML) technique is another good candidate by learning surrounding environments. The heuristic nature of ML technique encourages employing in CRN as well. Moreover, these techniques can provide sufficiently good performance in spectrum sensing classification [19].

As representative ML-based classification and regression algorithms, there are k-nearest neighbor, decision tree, naive Bayes, logistic regression, support vector machine (SVM), k-means clustering neural networks, and so on [20, 21]. In general, the SVM-based classifier outperforms the other techniques in practical problems due to kernel function trick [22–24]. In the SVM-based classifier, the problems that are not properly classifiable in a feature space are transformed to a high-dimensional space where classification is possible using a linear hyperplane.

In this paper, we employ an SVM-based classifier in order to classify the spectrum sensing results into legitimate SU and MU categories. In addition, an energy detection technique is utilized for sensing environments. Once the sensing is performed, the proposed SVM-based algorithm is employed on the data set and, it finds the maximal margin between the legitimate SUs and MUs. After the classification of legitimate SUs and MUs at the FC, the FC employs the DS evidence theory to measure the performance of the proposed SVM-based algorithm. The proposed scheme is verified in terms of false sensing probability when there exists either AYMUs, ANMUs, or random MUs (RMUs) in CSS environments. The AYMU sends higher energy statistics of the channel than actual status, and thus, it increases false alarm probability. The ANMU forwards lower energy statistics than actual status. Therefore, it results in misdetection and induces interference to the PUs. The RMUs randomly behaves in between both classes with probability 1-p. The effectiveness of the proposed scheme is evaluated through simulations in comparison with other existing schemes.

The rest of this paper is organized as follows. In Section 2, the system model considered through this paper is presented. In Section 3, the proposed SVM-based algorithm to classify

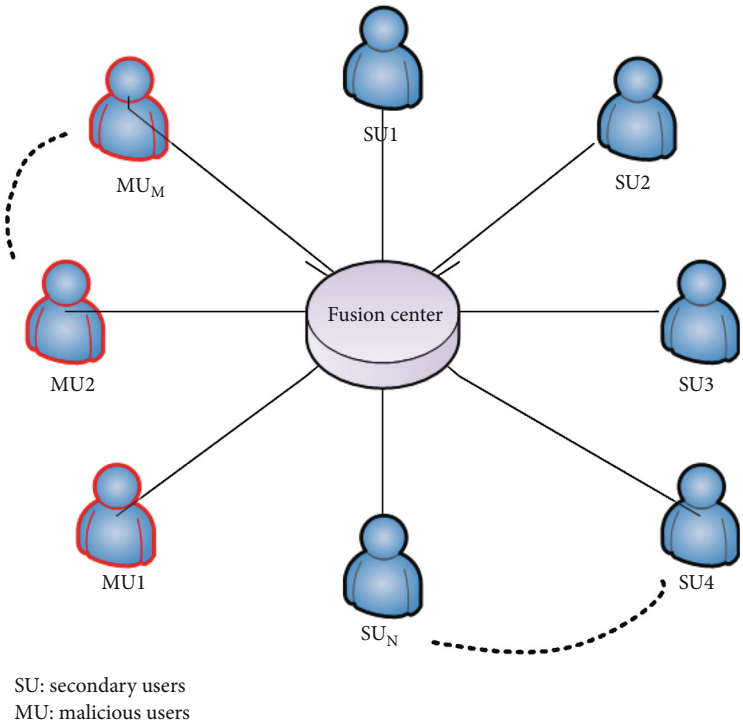


Figure 1: CSS system model.

legitimate SUs and MUs in CRN is presented. Numerical results are shown through simulations in Section 4. Finally, the paper is concluded in Section 5.

2. System Model

In this paper, we consider a CRN with N SUs and M MUs, where $M < N$ as shown in Figure 1. Initially, all the SUs including MUs perform spectrum sensing to determine the status of the PU in the network.

As in Figure 1, the SUs cooperate to sense the activity of the PU channel and inform the FC on their sensing information. The received information from the AYMU gives a higher energy signal which implies busy status of the PU channel. Similarly, the ANMU provides a low energy signal to the FC. The FC makes a global decision on the existence of the PUs in the network.

Each SU performs local sensing and sends its local result, either H_0 or H_1 for the absence or presence of PUs, respectively. The binary hypotheses test at the j^{th} SU is expressed as follows:

$$x_j(t) = \begin{cases} H_0 & n_j(t), \\ H_1 & h_j s(t) + n_j(t), \end{cases} \tag{1}$$

where H_0 corresponds to the absence of the PU, H_1 corresponds to the presence of the PU, $x_j(t)$ is the received signal at the j^{th} SU, $n_j(t)$ is the additive white Gaussian noise (AWGN), h_j is the channel gain between the PU and the j^{th} SU, and $s(t)$ is the signal transmitted by the PU.

Energy detection technique is very popular in CSS due to its ease of implementation and no requirement of prior information for the PU signal [25]. In this paper, we consider the energy detection for sensing the PU signals in the network.

The received signal test statistics of the PU channel by the j^{th} SU is given by

$$E_j(i) = \begin{cases} \sum_{t=t_i}^{t_i+K-1} \left| n_j(t) \right|^2, & H_0, \\ \sum_{t=t_i}^{t_i+K-1} \left| h_j s(t) + n_j(t) \right|^2, & H_1, \end{cases} \tag{2}$$

where K is the number of samples in the i^{th} sensing interval. According to the central limit theorem (CLT), the number of samples needs to be large enough so that the energy reported by each SU becomes similar to a Gaussian random variable under both H_0 and H_1 as in [25]:

$$\begin{cases} N\left(\mu_0 = K, \sigma_0^2 = 2K\right), H_0, \\ N\left(\mu_1 = K\left(\gamma_j + 1\right), \sigma_1^2 = 2K\left(\gamma_j + 1\right)\right), H_1, \end{cases} \tag{3}$$

where γ_j is the signal to noise ratio (SNR) between the PU and the j^{th} SU. Similarly, (μ_0, σ_0^2) and (μ_1, σ_1^2) are the mean and variance values of the reported signals under H_0 and H_1 hypotheses, respectively.

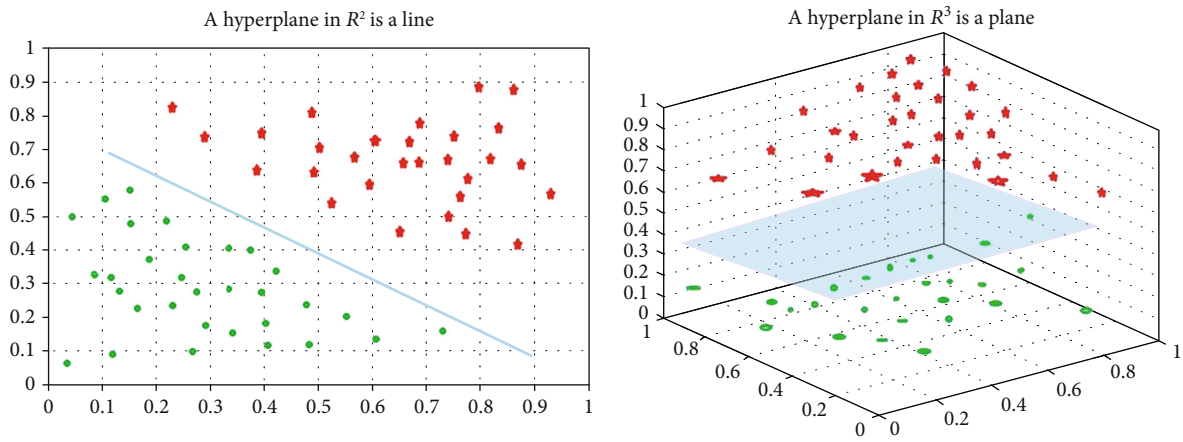


FIGURE 2: A hyperplane in R^n is an n-1 dimensional space [29].

3. Proposed Support Vector Machine-Based MU Classification Algorithm

ML provides a computational machine with the ability to learn without being explicitly programmed. ML methods are very effective when the data set is large, diverse, and fast changing. These algorithms give deep and predictive analysis of data, and they are classified into two big groups: supervised learning (classification and regression) and unsupervised learning (clustering techniques) [26, 27]. Our objective is to classify legitimate SUs and MUs among all the SUs available in the environment. SVM is a highly competitive learning method which is popular in many fields based on statistical learning theory [28].

In this paper, we proposed an SVM-based classification algorithm in CRN to classify legitimate SUs and MUs. The SVM can be used both for classification and regression problems. However, it is mostly used for classification [26]. It works on the basis of a hyperplane, which divides the different classes of data well. A hyperplane is a classifier that may be a dot, a line, or a plane depending upon the data scattered. The dimension of the hyperplane is less than the dimension of the data, i.e., if data is three dimensional, the hyperplane has to be a plane of two dimensions, and if the data is two dimensional, the hyperplane can be a line as shown in Figure 2.

The hyperplane is a point for one dimensional data. The concept of maximal margin decides the optimal hyperplane, whereas the margin is the distance between two support vectors. The support vectors are the data points nearest to the hyperplane, and these points are called critical points. If these points are moved, the position of hyperplane can be altered. Whenever test data is added, its position is decided as one of the classes.

As an application of SVM classification in CRN, we consider an environment in which N legitimate SUs and M MUs exist.

The notation of the data set is expressed as

$$D = \left\{ \left(\mathbf{x}_j, \mathbf{y}_j \right) \mid \mathbf{x}_j \in \mathbb{R}^n, \mathbf{y}_j \in \{-1, 1\} \right\}_{j=1}^{n}, \tag{4}$$

where $\mathbf{x}_j$ is the energy vector of N SUs, $\mathbf{y}_j$ is the class vector, and class "1" and "-1" represent legitimate SUs and MUs, respectively.

Once the sensing is done, the sensing results are fed into the SVM. The main objective of the proposed SVM-based scheme is to precisely classify legitimate SUs and MUs. First of all, we trace the support vectors and then draw two support hyperplanes. The optimal forms that define support hyperplanes to classify the legitimate SUs and MUs are given by

$$w.x + b = \delta, \tag{5}$$

$$w.x + b = -\delta, \tag{6}$$

where w is the weight vector obtained in training phase, b is a threshold value, and δ is an arbitrary constant. In the training phase, the distance between these two hyperplanes is called margin, and there are no data points in the margin. The margin can be clearly visualized by the following formula:

$$Y_i(w.x_i + b) \leq \delta. \tag{7}$$

The overlap region formed by (7) is the margin. At last step, we classify the given data by the hyperplane. The linearly nonseparable patterns are mapped onto a higher dimensional space such that the classification is possible using a linear hyperplane.

The overall flow chart of the proposed SVM-based scheme is shown in Figure 3.

The proposed SVM-based MU classification algorithm are shown in Algorithm 1. The algorithm consists of three phases of data generation, sensing, and classification.

Once the classification is done through the proposed SVM-based algorithm, the FC utilizes the DS evidence theory to combine the evidence values of H_0 and H_1 to make a global decision for the existence of the PU in the network.

In the DS evidence theory, the frame of discernment can be defined as $F_r = \{H_1, H_0, \Omega\}$ where Ω is the ignorance hypothesis, which describes whether hypotheses are true or

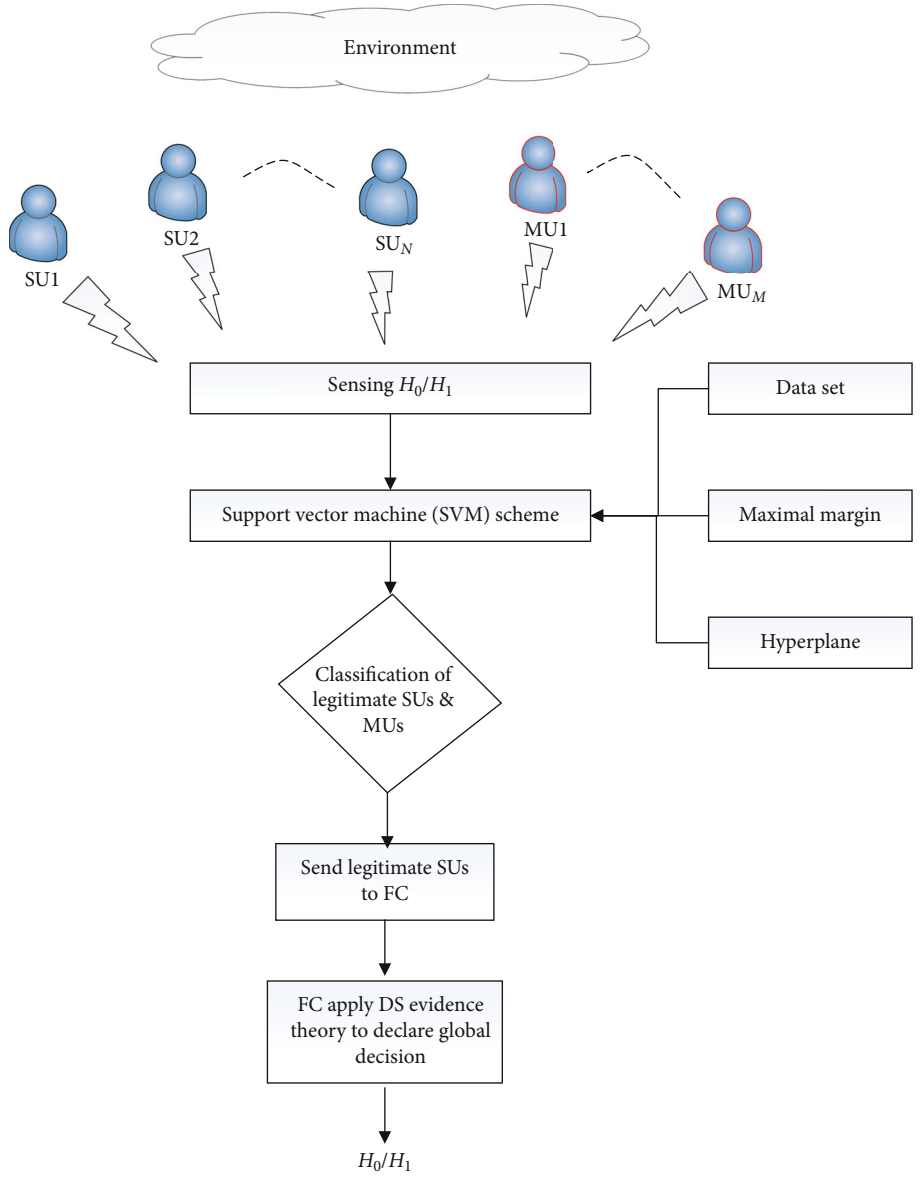


Figure 3: Proposed flowchart.

not. After each SU measures the basic probability assignment (BPA), $m(H_0)$ and $m(H_1)$, under hypotheses H_0 and H_1, respectively. The BPA measures are defined in the form of cumulative distribution function as follows [30]:

$$m_j(H_0) = \int_{E_j}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_{0j}} \exp\left(-\frac{\left(\mathbf{X}_j - \mu_{0j}\right)^2}{2\sigma^2{}_{0j}}\right) dx, \quad (8)$$

$$m_j(H_1) = \int_{-\infty}^{E_j} \frac{1}{\sqrt{2\pi}\sigma_{1j}} \exp\left(-\frac{\left(\mathbf{X}_j - \mu_{1j}\right)^2}{2\sigma^2{}_{1j}}\right) dx, \quad (9)$$

where $m_j(H_1)$, $m_j(H_0)$, and $m_j(\Omega)$ are the BPA hypotheses of j^{th} SU, respectively. These values are sent to the FC by SUs, and the FC makes a global decision on the existence of the PU by using these measures.

According to the DS evidence theory, the BPA can be combined based on the following equations [30]:

$$m_j(H_0) = \sum_{Fr_1 \cap Fr_2 \cdots Fr_n = H_0} \prod_{j=1}^{n} m_j(\mathrm{Fr}_j)/(1-k), \quad (10)$$

$$m_j(H_1) = \sum_{Fr_1 \cap Fr_2 \cdots Fr_n = H_1} \prod_{j=1}^{n} m_j(\mathrm{Fr}_j)/(1-k), \quad (11)$$

where $k = \sum_{Fr_1 \cap Fr_2 \cdots Fr_n = \varnothing} \prod_{j=1}^{n} m_j(\mathrm{Fr}_j)$, and Fr_j is an element of the set $\{H_1, H_0, \Omega\}$.

A. Generation of Data
Initialization of parameters such as number of iteration, number of Sus.
Generate random MUs with Gaussian distribution.
Generate normal SUs.
Generate the indices on which MUs attack.
Generate indices for position of normal SUs.
B. Sensing the data
For $n = 1$ to *Sensing Interval*
For $i = 1$ to N
Energy reported by the j^{th} SUs.
End
For $i = 1$ to M
Energy reported by the MUs.
End
End sensing interval
C. Support Vector Machine Algorithm
1. Data Processing
i. Combining the data
ii. Input the data
iii. Train the data
iv. Find the number of examples and attributes used in the data.
v. Extract the attribute matrix X and the label vect Y.
2. Support Vector
i. Finding the support vectors (Corner points)
ii. Draw the upper and lower hyperplanes.
iii. Finding the maximal margin by the upper and lower hyperplane.
3. Classification
i. Linear
Finding weights
$Y_1 = -(w_1.x_1 + b)/w_2$
ii. Drawing the hyperplane to classify the data.
D. Plotting
i. Normal data plotting.
ii. Malicious data plotting.
iii. Plotting of hyperplane.

Algorithm 1. Proposed SVM-based MU classfication algorithm.

Table 1: Simulation parameters.

Parameters	Values
Number of SUs	10
Probability of PU	0.5
Number of MUs	4
Number of iterations	100

Finally, a simple decision strategy is chosen at the FC to declare the global decision as

$$f_d = \begin{cases} H_1; & \dfrac{m(H_1)}{m(H_0)} > \lambda, \\ H_0; & \dfrac{m(H_1)}{m(H_0)} \leq \lambda. \end{cases} \quad (12)$$

4. Numerical Results and Evaluation

To evaluate the effectiveness of the proposed SVM-based scheme, we conduct extensive simulations by using MATLAB tool. In our simulation, we consider a CRN with $M = 10$ SUs. Among the total number of SUs, six SUs are selected as legitimate SUs, and four are randomly selected as MUs. According to IEEE 802.22 standards, it is assumed that the used bandwidth is 6 MHz, and the PU activity is 0.5. The detailed simulation parameters are listed in Table 1.

We perform the simulation in two parts. In the first part, we simulate the proposed SVM-based scheme when no MUs exist in the network, AYMUs exist in the network, ANMUs exist in the network, and RMUs exist in the network. In the second part, we compare the performance of the proposed SVM-based scheme with those of the other existing schemes.

First of all, we show the results when only legitimate SUs exist in the network, ANMUs exist in the network, AYMUs exist in the network, and when RMUs exist in the network. In Figures 4–7, the random generation of legitimate SUs and MUs is shown, and the classification of legitimate SUs and MUs is clearly presented by employing the proposed SVM-based scheme. The AYMU is the one which always feeds the local sensing result as the PU absence. The ANMU is the one which always feeds it as the PU presence. The RMU is the most difficult attack to classify, since in this attack, the

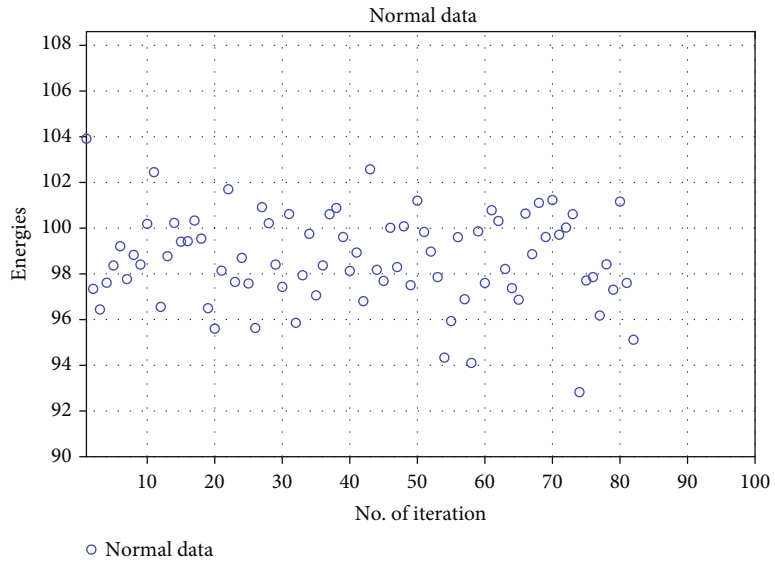


Figure 4: Normal data generation.

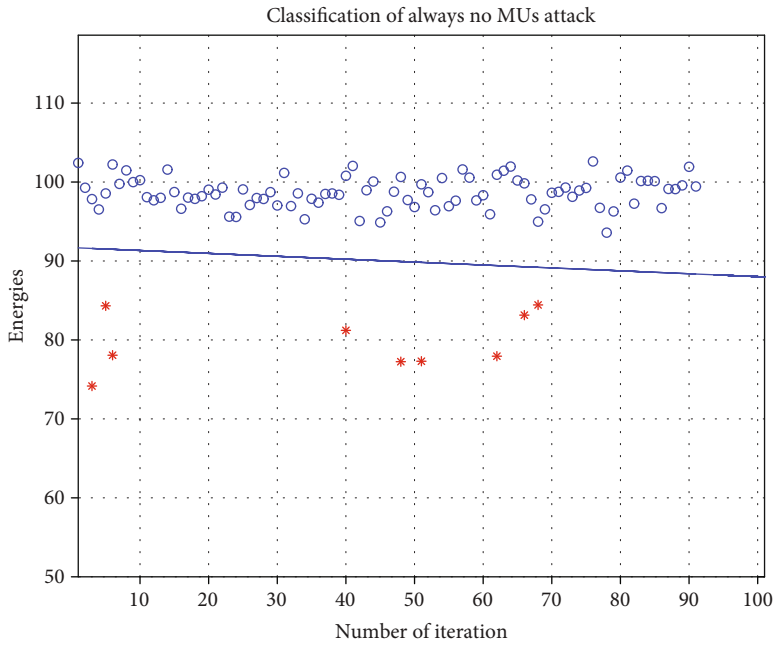


Figure 5: Classification of legitimate SUs and MUs, when ANMUs exist.

MUs sometime behave like AYMU and sometime like ANMU with probability 1-p.

Figure 4 shows the normal data generation by legitimate SUs. The range of sensing energies at the legitimate SUs lies on the range of 90-108. From Figure 4, it can be observed that the legitimate SUs send different energies with different probabilities and different number of iterations. None of the number of iterations is out of the range of the defined energy range (i.e., 90-108). Thus, for all the number of iterations, the data of the legitimate SUs are represented.

Figure 5 shows the classification result of the legitimate SUs and ANMUs by employing the proposed SVM-based scheme. The SVM works on the concept of hyperplane. A hyperplane is an n-dimensional line used to classify different classes of the data by maximum margin. In Figure 5, the legitimate SUs are denoted by blue circles, while MUs are denoted by red circles. The result shows that when the ANMU attack where MUs always produce lower energies than actual status exist in the network, it is well-classified by the proposed SVM-based algorithm.

Figure 6 shows the classification of legitimate SUs and MU, when AYMUs exist in the networks. The AYMU always sends higher energies than actual status to the FC, which results in the existence of the PU in the network. The AYMU degrades the system performance in terms of the opportunity of channel access by the SUs. It is shown that the proposed

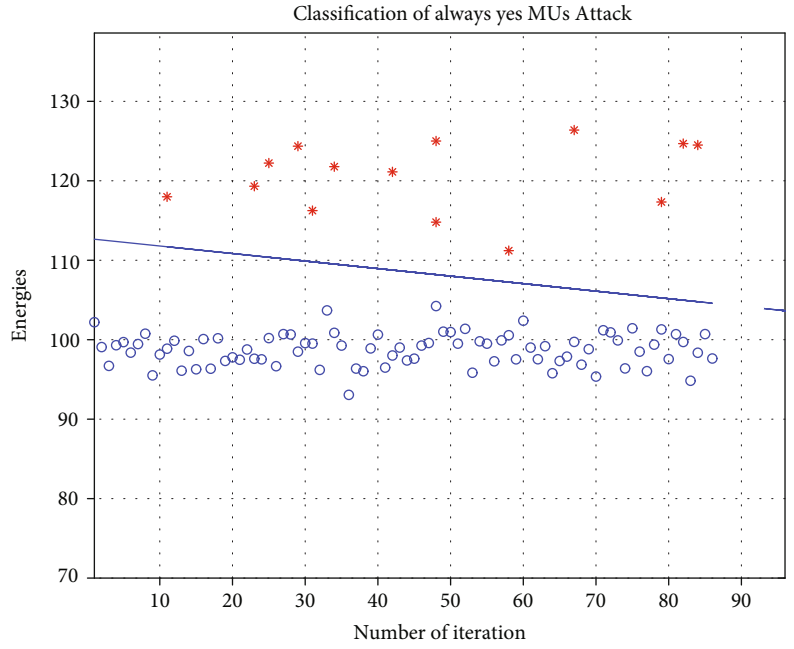


Figure 6: Classification of legitimate SUs and MUs, when AYMUs exist.

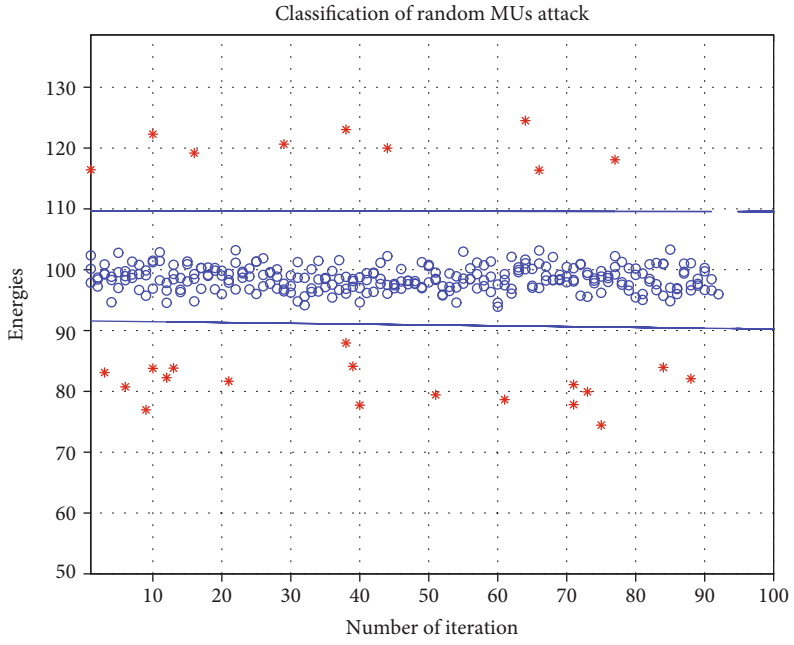


Figure 7: Classification of legitimate SUs and MUs, when RMUs exist.

SVM-based algorithm efficiently classifies the legitimate SUs from the MUs in the network.

Figure 7 shows the classification of legitimate SUs and MUs, when RMUs, who sometimes behave like AYMUs and sometimes like ANMUs, exist in the networks. The legitimate SUs are in the range of 90-108 energy level. The AYMUs have energy level higher than 108, and the ANMUs have the energy level less than 90. The RMUs are the most difficult attack to classify, since the SUs behave randomly with probability 1-p. Through the proposed SVM scheme, the legitimate SUs and the RMUs can be efficiently classified.

In this part of the simulation, we compare the performance of the proposed SVM-based scheme with the other existing schemes. Through Figures 8–10, we show the performance of the proposed SVM-based scheme when ANMUs, AYMUs, and RMUs attackers are in the network, compared to the existing schemes.

Figure 8 shows the region of convergence (ROC) curve of the proposed SVM-based scheme in comparison with other

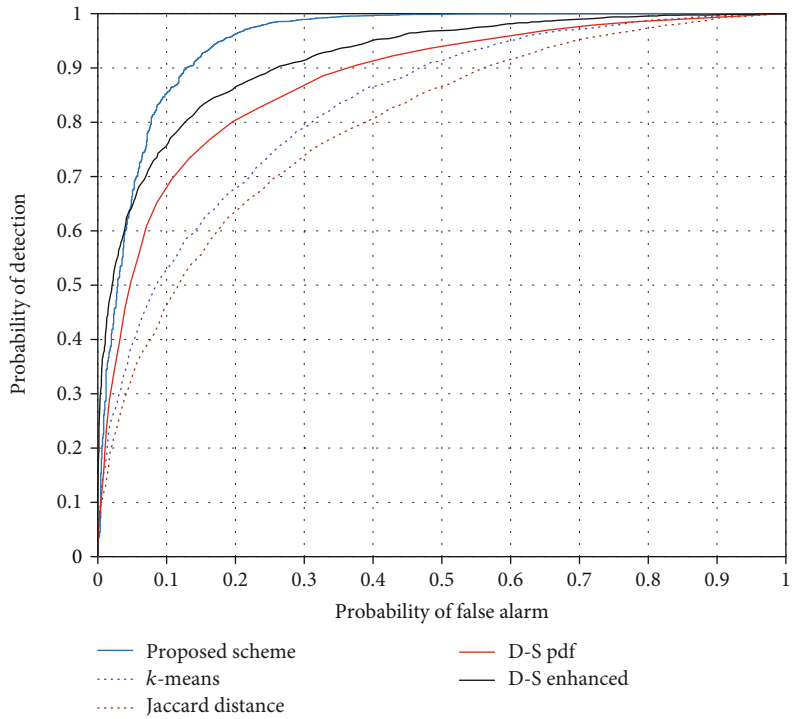


Figure 8: ROC curve of proposed scheme with other schemes, when ANMUs exist.

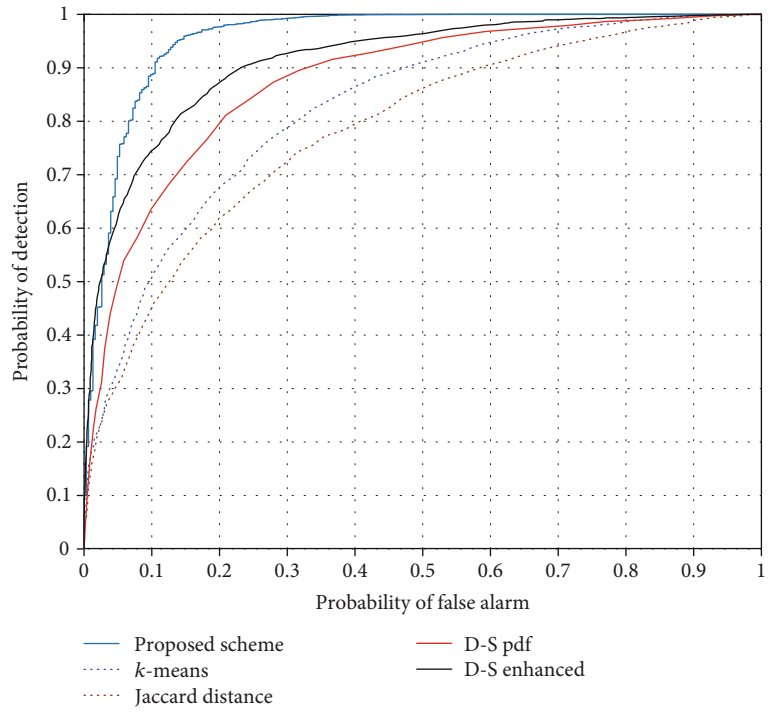


Figure 9: ROC curve of proposed scheme with other schemes, when AYMUs exist.

existing schemes, when ANMUs exist in the network. It is shown through simulation results that the proposed SVM-based scheme efficiently classifies the legitimate SUs and ANMUs. Once the legitimate SUs and ANMUs classified, the ROC of the proposed SVM-based scheme is plotted in comparison with those of the existing schemes. It is observed that the proposed SVM-based scheme outperforms the other existing schemes.

Figure 9 shows the ROC curve of the proposed SVM-based scheme in comparison with other existing schemes,

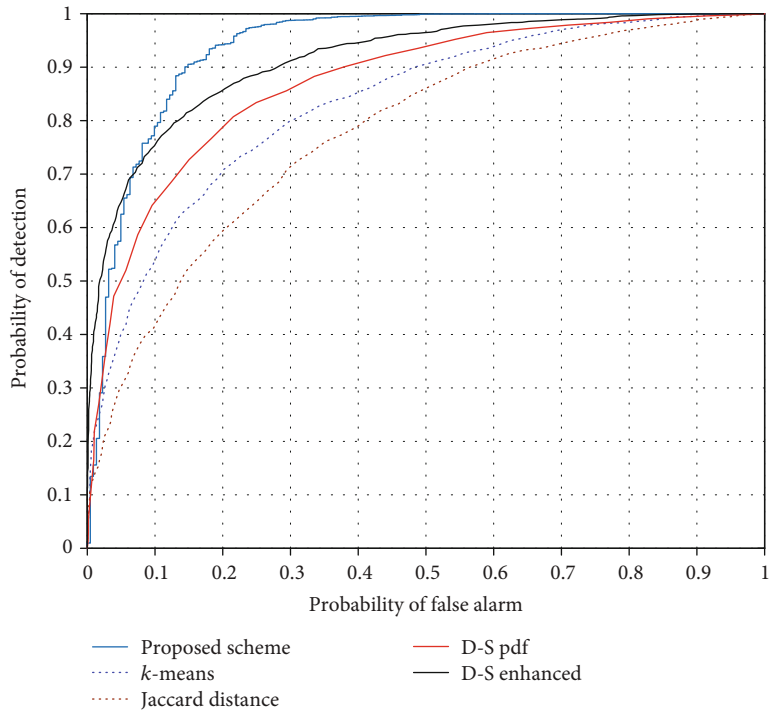


Figure 10: ROC curve of proposed scheme with other schemes, when RMUs exist.

when AYMU exist in the network. It is shown through simulation that the proposed SVM-based scheme efficiently classifies the legitimate SUs and AYMUs. The ROC curve shows that the proposed SVM-based scheme has better performance than the other existing schemes.

Figure 10 shows the ROC curve of the proposed SVM scheme in comparison with other schemes, when RMUs exist in the network. It is shown that proposed SVM-based scheme efficiently classifies the legitimate SUs and RMUs. The ROC shows that the proposed SVM-based scheme also outperforms the other existing schemes even when RMUs exist.

It is clear that based on the SVM-based classification, the proposed SVM-based scheme can optimize to classify legitimate SUs from MUs efficiently. The risk of considering the MUs in CSS is significantly removed with the proposed SVM-based scheme. Consequently, the proposed SVM-based scheme is able to identify and classify the MUs and provide the reliable sensing results in CSS-based CRNs.

5. Conclusions

Recently, machine learning has attracted attentions in spectrum sensing. The main reason of the attraction is that it is a heuristic approach without requiring the prior information about surrounding environments. Cooperative spectrum sensing (CSS) improves the performance of cognitive radio networks (CRNs). However, the performance of CSS severely degrades by attacks from malicious users (MUs). In this paper, we proposed a support vector machine- (SVM-) based algorithm to classify legitimate secondary users (SUs) and MUs. Once the legitimate SUs and MUs are classified through the proposed SVM-based algorithm, a fusion center (FC) combines the diversified sensing reports received from the legitimate SUs based on the DS evidence theory in order to make a global decision on the existence of primary users (PUs) in the network. The numerical results verified the superiority and the authenticity of the proposed SVM-based classification of the legitimate SUs and MUs.

Data Availability

The data used to support the finding of this study are included in the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01426) supervised by the IITP (Institute for Information and Communication Technology Planning & Evaluation) and in part by the National Research Foundation (NRF) funded by the Korea government (MSIT) (No. 2019R1F1A1059125).

References

[1] FCC, *Notice of proposed rule-making and order*, Washington, D.C., 2003ET Docket No. 03-222.

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[3] L. Zhai, H. Wang, and C. Gao, "A Spectrum Access Based on Quality of Service (QoS) in Cognitive Radio Networks," *PLoS One*, vol. 11, no. 5, p. e0155074, 2016.

[4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[5] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: state-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[6] M. S. Khan, M. Jibran, I. Koo, S. M. Kim, and J. Kim, "A Double Adaptive Approach to Tackle Malicious Users in Cognitive Radio Networks," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 2350694, 9 pages, 2019.

[7] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radio," in *2006 IEEE International Conference on Communications*, Istanbul, Turkey, 2006.

[8] Y. He, J. Xue, T. Ratnarajah, M. Sellathurai, and F. Khan, "On the Performance of Cooperative Spectrum Sensing in Random Cognitive Radio Networks," *IEEE Systems Journal*, vol. 12, no. 1, pp. 881–892, 2018.

[9] S. Chilakala and M. S. S. Ram, "Optimization of cooperative secondary users in cognitive radio networks," *Engineering Science and Technology, an International Journal*, vol. 21, no. 5, pp. 815–821, 2018.

[10] M. Jenani, "Network security, a challenge," *International Journal of Advanced Networking and Applications*, vol. 8, no. 5, pp. 120–123, 2017.

[11] J. Marinho, J. Granjal, and E. Monteiro, "A survey on security attacks and countermeasures with primary user detection in cognitive radio networks," *EURASIP Journal on Information Security*, vol. 2015, no. 1, 2015.

[12] H. Wu, X. Sun, C. Guo, and S. Ren, "Malicious user detection for wide-band cognitive radio network," in *2016 Asia-Pacific Microwave Conference (APMC)*, New Delhi, India, 2016.

[13] A. Taggu, C. Chunka, and N. Marchang, "CODES: A COllaborative DEtection Strategy for SSDF Attacks in Cognitive Radio Networks," in *Proceedings of the Third International Symposium on Women in Computing and Informatics - WCI '15*, Kochi India, 2015.

[14] B. Sarala, S. R. Devi, M. Suganthy, and S. J. Ida, "A novel authentication mechanism for cognitive radio networks," *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 713–718, 2019.

[15] R. Wan, L. Ding, N. Xiong, and X. Zhou, "Mitigation strategy against spectrum sensing data falsification attack in cognitive radio sensor networks," *International Journal of Distributed Sensor Networks*, vol. 15, no. 9, 2019.

[16] F. Farmani, M. A. Jannatabad, and R. Berangi, "Detection of SSDF Attack Using SVDD Algorithm in Cognitive Radio Networks," in *2011 Third International Conference on Computational Intelligence, Communication Systems and Networks*, Bali, Indonesia, 2011.

[17] U. Mehboob, J. Qadir, S. Ali, and A. Vasilakos, "Genetic algorithms in wireless networking: techniques, applications, and issues," *Soft Computing*, vol. 20, no. 6, pp. 2467–2501, 2016.

[18] M. S. Khan, N. Gul, J. Kim, I. M. Qureshi, and S. M. Kim, "A Genetic Algorithm-Based Soft Decision Fusion Scheme in Cognitive IoT Networks with Malicious Users," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 2509081, 10 pages, 2020.

[19] F. Azmat, Y. Chen, and N. Stocks, "Analysis of Spectrum Occupancy Using Machine Learning Algorithms," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 6853–6860, 2016.

[20] P. Harrington, *Machine Learning in Action*, Manning Publications, 2012.

[21] K. Patan, "Artificial neural networks for modelling and fault diagnosis of technical process," *Lecture Notes in Control and Information Sciences*, vol. 377, 2008.

[22] F. Wang, Z. Zhen, B. Wang, and Z. Mi, "Comparative study on KNN and SVM based weather classification Models for day ahead short term solar PV power forecasting," *Applied Science*, vol. 8, no. 1, p. 28, 2018.

[23] K. Elangovan, Y. Krishnasamy Tamilselvam, R. Mohan, M. Iwase, N. Takuma, and K. Wood, "Fault Diagnosis of a Reconfigurable Crawling–Rolling Robot Based on Support Vector Machines," *Applied Science*, vol. 7, no. 10, p. 1025, 2017.

[24] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, "Sensor Fault Classification Based on Support Vector Machine and Statistical Time-Domain Features," *IEEE Access*, vol. 5, pp. 8682–8690, 2017.

[25] M. S. Khan and I. Koo, "The Effect of Multiple Energy Detector on Evidence Theory Based Cooperative Spectrum Sensing Scheme for Cognitive Radio Networks," *Journal of Information Processing Systems*, vol. 12, no. 2, pp. 295–309, 2015.

[26] Y. Molina-Tenorio, A. Prieto-Guerrero, R. Aguilar-Gonzalez, and S. Ruiz-Boqué, "Machine Learning Techniques Applied to Multiband Spectrum Sensing in Cognitive Radios," *Sensors*, vol. 19, no. 21, p. 4715, 2019.

[27] V. Sharma and V. Bohara, "Exploiting machine learning algorithms for cognitive radio," in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, New Delhi, India, 2014.

[28] D. Zhang and X. Zhai, "SVM-based spectrum in cognitive radio," in *2011 7th International Conference on Wireless Communications, Networking and Mobile Computing*, Wuhan, China, 2011.

[29] N. S. Chauhan, *A Friendly Introduction to Support Vector Machine (SVM)*, Towards Data Science, 2019.

[30] M. S. Khan and I. Koo, "Mitigation of Adverse Effects of Malicious Users on Cooperative Spectrum Sensing by Using Hausdorff Distance in Cognitive Radio Networks," *Journal of information and communication convergence engineering*, vol. 13, no. 2, pp. 74–80, 2015.