

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

1/2023

Beam Pattern Optimization Via Unequal Ascending Clusters

A.J. Abdulqader, J.R. Mohammed, and Y.E.M. Ali

Paper

1

Synthesis of Reconfigurable Multiple Shaped Beams of a Concentric Circular Ring Array Antenna Using Evolutionary Algorithms

S.K. Dubey and D. Mandal

Paper

8

CG-SCMA Codebook Design Based on Maximized Euclidian Distance

S.S. Mohamed and H.N. Abdullah

Paper

18

A New Class of Fractional Cumulative Residual Entropy – Some Theoretical Results

S. Benmahmoud

Paper

25

Concept of Joint Functioning of Access Control Systems

M. Strelbitskiy, V. Mazur, E. Prokopenko, R. Rachok, and D. Mul

Paper

30

Adaptive Rider Feedback Artificial Tree Optimization-Based Deep Neuro-Fuzzy Network for Classification of Sentiment Grade

S. Jasti and G.V.S. Raj Kumar

Paper

37

Improved Association Rule Mining-Based Data Sanitization for Privacy Preservation Model in Cloud

R. Patil and G. HimaBindu

Paper

51

Single Linkage Weighted Steepest Gradient Adaboost Cluster-Based D2D in 5G Networks

S. Varadala and S.E. Roslin

Paper

60

(Contents continued on back cover)

Editor-in Chief

Adrian Kliks, Poznan University of Technology, Poland

Steering Editor

Paweł Plawiak, National Institute of Telecommunications, Poland

Editorial Advisory Board

Hovik Baghdasaryan, National Polytechnic University of Armenia, Armenia

Naveen Chilamkurti, LaTrobe University, Australia

Luis M. Correia, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Luca De Nardis, DIET Department, University of Rome La Sapienza, Italy

Nikolaos Dimitriou, NCSR “Demokritos” Athens, Greece

Ciprian Dobre, Politechnic University of Bucharest, Romania

Filip Idzikowski, Poznan University of Technology, Poland

Andrzej Jajszczyk, AGH University of Science and Technology, Poland

Albert Levi, Sabanci University, Turkey

Marian Marciniak, National Institute of Telecommunications, Poland

George Mastorakis, Technological Educational Institute of Crete, Greece

Constandinos Mavromoustakis, University of Nicosia, Cyprus

Klaus Mößner, Technische Universität Chemnitz, Germany

Imran Muhammad, King Saud University, Saudi Arabia

Mjumo Mzyece, University of the Witwatersrand, South Africa

Daniel Negru, University of Bordeaux, France

Ewa Orłowska, National Institute of Telecommunications, Poland

Jordi Perez-Romero, UPC, Spain

Michał Pióro, Warsaw University of Technology, Poland

Konstantinos Psannis, University of Macedonia, Greece

Salvatore Signorello, University of Lisboa, Portugal

Adam Wolisz, Technische Universität Berlin, Germany

Tadeusz A. Wysocki, University of Nebraska, USA

Editorial Team

Content Editor: **Robert Magdziak**

Managing Editor: **Ewa Kapuściarek**

Technical Editor: **Grażyna Woźnica**

eISSN 1899-8852

© Copyright by National Institute of Telecommunications, Poland 2023

Beam Pattern Optimization Via Unequal Ascending Clusters

Ahmed Jameel Abdulqader¹, Jafar Ramadhan Mohammed¹, and Yessar E. Mohammad Ali²

¹College of Electronics Engineering, Ninevah University, Mosul, Iraq,

²College of Engineering, Nawroz University, Duhok, Iraq

<https://doi.org/10.26636/jtit.2023.168523>

Abstract — In this paper, two different architectures based on completely and sectionally clustered arrays are proposed to improve the array patterns. In the wholly clustered arrays, all elements of the ordinary array are divided into multiple unequal ascending clusters. In the sectionally clustered arrays, two types of architectures are proposed by dividing a part of the array into clusters based on the position of specific elements. In the first architecture of sectionally clustered arrays, only those elements that are located on the sides of the array are grouped into unequal ascending clusters, and other elements located in the center are left as individual and unoptimized items (i.e. uniform excitation). In the second architecture, only some of the elements close the center are grouped into unequal ascending clusters, and the side elements were left individually and without optimization. The research proves that the sectionally clustered architecture has many advantages compared to the completely clustered structure, in terms of the complexity of the solution. Simulation results show that PSLL in the side clustered array can be reduced to more than -28 dB for an array of 40 elements. The PSLL was -17 dB in the case of a centrally clustered array, whereas the complexity percentage in the wholly clustered array method was 12.5%, while the same parameter for the partially clustered array method equaled 10%.

Keywords — beam pattern, clustered array, genetic algorithm, sidelobe.

1. Introduction

Synthesizing a large antenna array for modern communication applications, such as MIMO in 5G, is one of the most difficult challenges that designers face in practice. Therefore, it is necessary to use adapted techniques to simplify the construction process. Relying on subarrays is an important technique that is widely used in synthesizing large arrays to reduce complexity of the system by controlling the exciting points at the subarray level instead of the individual elements [1]. Such a reduction materially simplifies the manufacturing of large arrays and, thus, their cost may be reduced. The subarray technique means that the array containing N radiating elements is portioned into small classes (called clusters), with each cluster containing C radiating elements, so that $C \ll N$, and only C control points are required in the excitation of the array [2]. Each cluster has only one controller to feed the clusters in terms of amplitude and phase, or both. Although the desired beamwidth and directivity are still achieved, the use

of the cluster architecture offers limited operating bandwidth and low reconfiguration capabilities due to the quantized distribution of the excitation across the aperture array [3]. For example, if a linear array is portioned into equal clusters, the radiation pattern will suffer from the appearance of grating lobes, and the number of these lobes increases with the growing number of elements in the cluster. To bypass this problem, arrays made of irregularly arranged or polyomino-shaped clusters are used in large planar arrays [3], [4].

In the past years, several efficient methods have been proposed to address the problems of cluster synthesis. As far as we know, two types of methods based on evolutionary algorithms are used, such as genetic algorithm [2], [3] and particle swarm [5] algorithm. Mathematical analysis strategies, such as compressed sensing [6]–[8] are relied upon as well. In [2], two different structures are proposed to build linear clustered arrays based on a genetic algorithm. In the first structure, all the elements in the large array are fragmented into equal-sized clusters, meaning that unwanted side lobes may appear in the radiation pattern. This problem was solved through the use of a second structure, where the edges of the linear array were divided into clusters, while the center elements remained without clustering, i.e. by optimizing the center elements individually. This leads to a lower reduction in the system's complexity. In [3], two structures based on the genetic algorithm are proposed as well. In the first structure, the radiating elements in the main array are divided into several irregular clusters, and in the second structure, only the elements located at the sides are divided into irregular clusters. In the two methods, high degrees of freedom were obtained to address the problem of undesirable side lobes, but the complexity remained rather high.

In this paper, the methods described in [2] and [3] have been extended and simplified by proposing fully and partially clustered architectures that are arranged in ascending order in terms of the number of elements in each cluster needed to build the desired radiation pattern under the applicable constraints. In the fully clustered array, the original linear array is divided into a group of clusters, with each one of them having an unequal number of elements. The clusters are arranged in an ascending order (in terms of the number of elements), starting from the edges and heading towards the center. Such a solution is referred to as a completely clustered array (CCA). Meanwhile, in the case of the partially clustered array, two

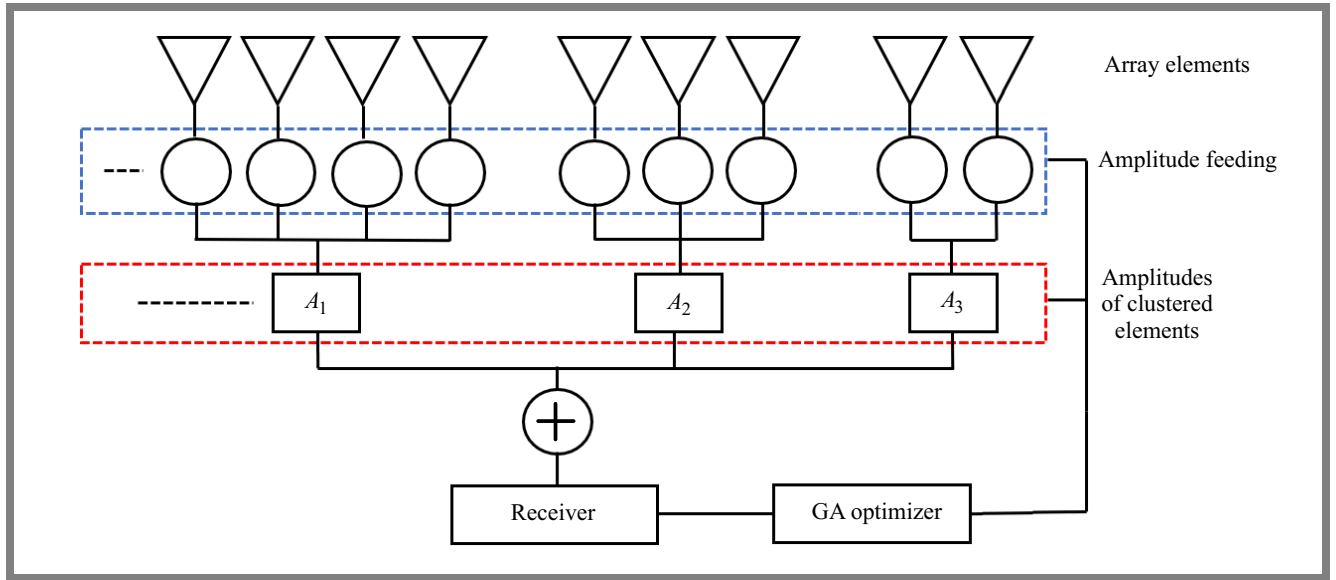


Fig. 1. CCA architecture.

architectures are proposed. In the first approach, the elements located on the sides are divided into ascending groupings, while the elements located in the middle are left without clustering and optimization. Such a solution is referred to as a side section clustered array (SSCA). In the second architecture, in turn, the elements located in the middle are divided into ascending groupings, while the elements located on the sides are left without clustering and optimization. Such a solution is referred to as a center section clustered array (CSCA). The sectionally clustered array architecture is characterized by a lower degree of complexity than the completely clustered array structure. The two proposed architectures are capable of ensuring good radiation properties with a simplified excitation network. To suppress unwanted quantized sidelobes existing in the desired radiation pattern, an upper boundary sidelobe limit constraint exempting the main lobe is proposed.

2. Formulation of the Proposed Clustered Structures

In this section, wholly and partially ascending clustered array architectures are presented and their practical advantages are demonstrated. For both architectures, the array factor and the cost function are declared with the constraint condition. The genetic algorithm is used under the following specifications: population size 60, mutation rate 0.16, and single-point crossover is used to construct the desired radiation pattern.

2.1. CCA Elements

Let us consider a linear array containing an even number of N radiating elements. The elements are arranged and fed symmetrically around the center of the array, meaning that half of the elements are dealt with instead of all of them, as is the case in the optimization process. This means that

complexity will be reduced by half as well. In CCA, all the elements in the array are portioned into several C clusters and each c cluster contains B elements. B is always lower than N and N must be chosen so that the quotient of N/B is zero (i.e. C must be an integer). Figure 1 shows the structure of the proposed CCA. It can be seen that each cluster contains a set of elements. These clusters are arranged in an ascending order, starting from the edges to the center, depending on the number of elements. This means that the first cluster contains the least number of elements (for instance two) with the number of elements increasing by one in the next cluster, and so on. Such a gradual increase in the number of elements prevents the appearance of periodic quantized lobes in the desired radiation pattern.

To represent the mathematical function of controlling the upper bounds of the side lobes constraint in the optimization process, we start by writing the array factor (AF) of the clustered array as [3]:

$$AF(\theta) = 2 \sum_{c=1}^{\frac{C}{2}} w_c \sum_{n=1}^{\frac{N}{2}} \beta_{cn} w_n \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right], \quad (1)$$

$$\beta_{cn} = \begin{cases} 1, & \text{if the } n\text{-th elements belong to the } c\text{-th cluster} \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where w_c is the complex weight for each cluster output and equals $w_c = A_c e^{jp_c}$, A_c and p_c are the amplitude and phase excitation, respectively, w_n is the complex weight of each independently exciting element in the main array and equals $w_n = A_n e^{jp_n}$, A_n and p_n are the amplitude and phase excitation, respectively, k is the wave number equal to $2\pi/\lambda$ and λ is the free space wavelength, d is the distance between any two elements in the linear array, while θ is the angle of the main beam's direction.

It is clear from Eq. (1) that the number of clusters on each side of the linear array is equal and is arranged symmetrically.

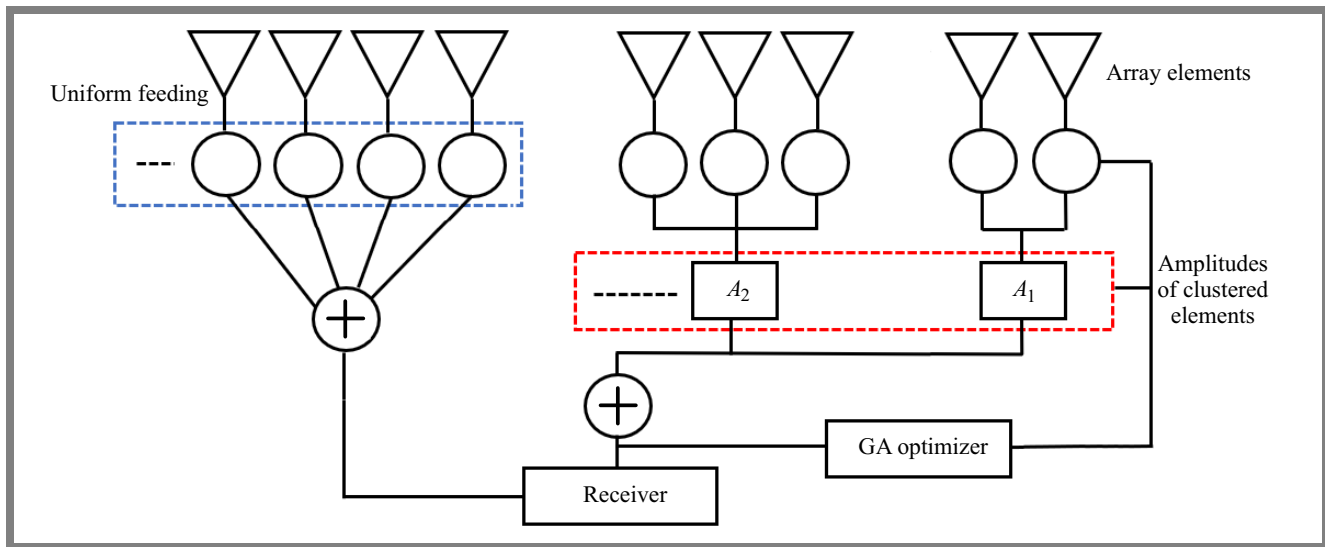


Fig. 2. SSCA architecture.

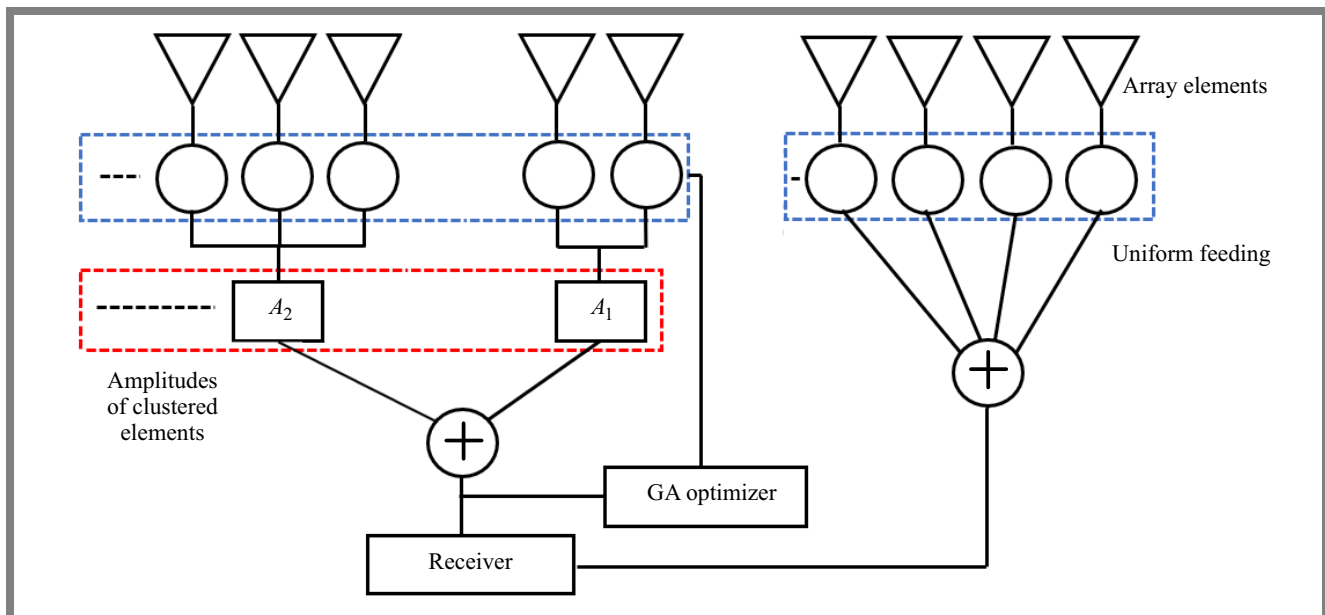


Fig. 3. CSCA architecture.

Thus, half of the clusters will be dealt with in the optimization process. To further simplify the exciting network, amplitude-only weights are used, meaning $p_c = p_n = 0$. Therefore, Eq. (1) can be simplified to:

$$AF(\theta) = 2 \sum_{c=1}^{\frac{C}{2}} A_c \sum_{n=1}^{\frac{N}{2}} \beta_{cn} A_n \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right], \quad (3)$$

Equation (3) will be used in the optimization process to determine the optimal weight values for the elements of each cluster in the array system. A_c and A_n weights are optimized at the initial stage to construct the desired radiation pattern. After that, the average value of $A_1, A_2, \dots, A_c$ weights is taken within the cluster until C is calculated. In this way, the weights of $A_1, A_2, \dots, A_n$ are quantitatively determined by the value of A_c .

2.2. SSCA and CSCA Elements

In this section, a new method based on a partial cluster is presented. Two architectures are proposed to build a low-complexity array system. In the first architecture (SSCA), some elements in the linear array located on the edges are selected to be converted into ascendingly arranged clusters, with the center elements remaining without clustering and with uniform weights. Figure 2 shows this proposed architecture. The second architecture (CSCA) is the opposite of the first approach, as a group of elements close to the center is converted into ascending clusters, while the edge elements remain without clustering and improvement (i.e. uniform exciting – see Fig. 3).

The main advantage of these architectures is that the level of complexity of large array systems is lower compared to the

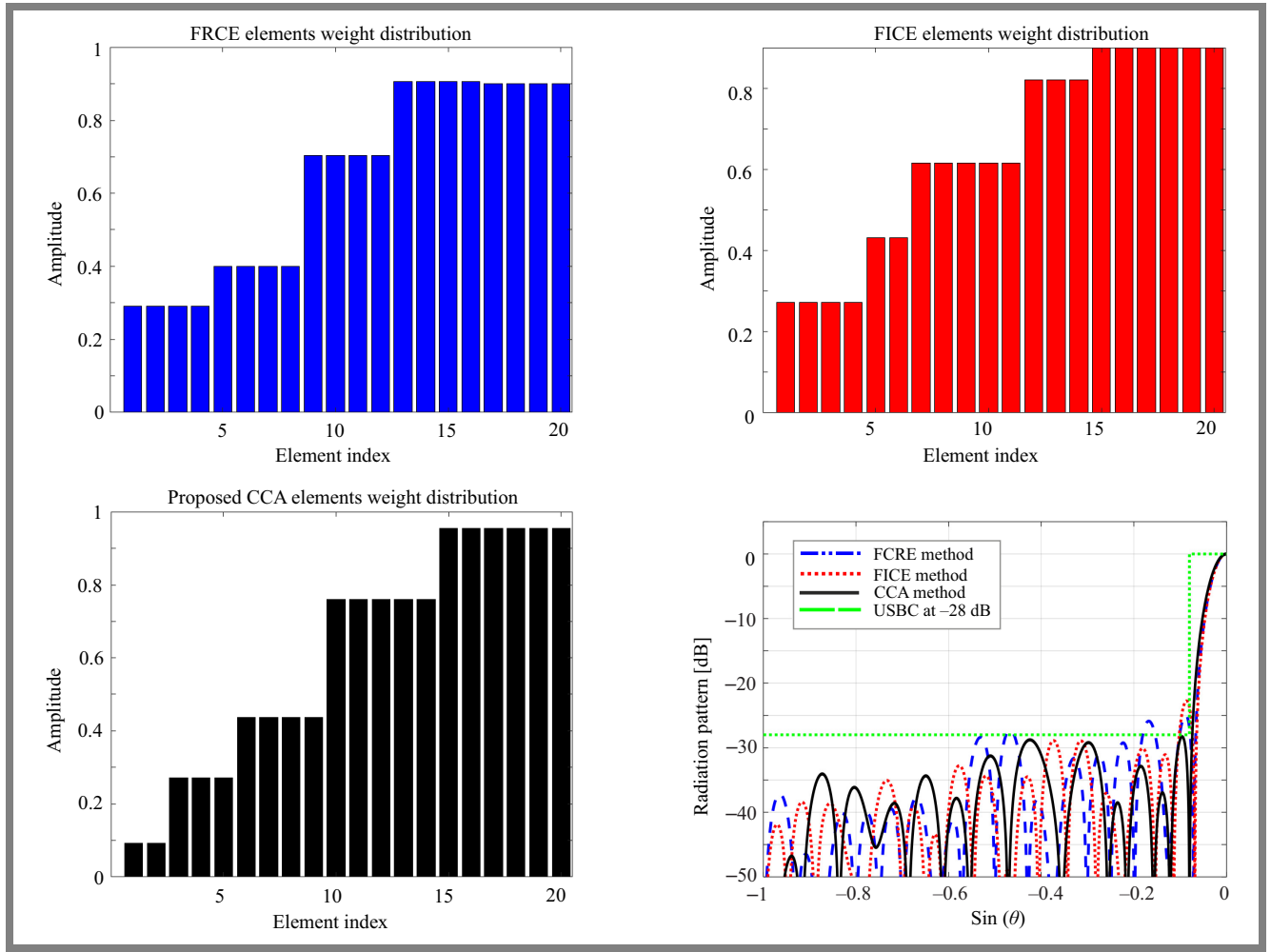


Fig. 4. Comparisons between the proposed CCA architecture and the existing structures from [2], [3].

first approach and to the structures proposed in [2], [3], while a good and desired beam pattern is guaranteed. Hence, the AF of these systems can be written as:

$$\begin{aligned}
 AF(\theta)_{SSCA} = & \underbrace{2 \sum_{u=1}^{\frac{U}{2}} b_u \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right]}_{\frac{U}{2} \text{ individually exciting elements}} \\
 & + \underbrace{2 \sum_{c=1}^{\frac{C}{2}} A_c \sum_{n=\frac{(N-U)+1}{2}}^{\frac{(N-U)+1}{2}} \beta_{cn} A_n \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right]}_{\frac{C}{2} \text{ clusters on each side of the linear array}}, \quad (4)
 \end{aligned}$$

$$\begin{aligned}
 AF(\theta)_{CSCA} = & \underbrace{2 \sum_{c=1}^{\frac{C}{2}} A_c \sum_{n=1}^{\frac{(N-U)+1}{2}} \beta_{cn} A_n \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right]}_{\frac{C}{2} \text{ individually exciting elements}} \\
 & + \underbrace{2 \sum_{u=1}^{\frac{U}{2}} b_u \cos \left[\frac{2n-1}{2} kd \sin(\theta) \right]}_{\frac{U}{2} \text{ clusters individually exciting elements}}, \quad (5)
 \end{aligned}$$

where b_u represents the uniform weights of the individual elements in the main array. For the objective functions described in Eqs. (3)–(4) to work precisely in the optimization process, an additional condition is added to these functions that draw a definite borderline to control the levels of the side lobes in order to prevent them from being higher than the required levels. This will be referred to as the upper sidelobe boundary constraint (USBC). Then, it may be written as:

$$\text{Additional cost function} = \sum |AF(\theta) - USBC|^2. \quad (6)$$

3. Computer Simulation Results

In order to check the effectiveness of the proposed systems, i.e. CCA, SSCA, and CSCA, several tests have been performed. In all simulations, phase excitation ($p_c = p_n = 0$) of the elements is set to zero, i.e. amplitude-only weights are used in the optimization process. The total number of elements used in the main array is chosen to be 40, with 20 elements on each side. The number of elements in each cluster C is specified, and then this number is increased by one in the next cluster. For example, if the number of ele-

ments in the first cluster $b = 2$, then the second cluster will contain three elements $b = 3$, and so on (when $b = 1$, it represents a non-clustering condition). Because a symmetric linear array was used in all tests, only one aspect of the radiation pattern and amplitude distribution of the elements is shown.

In the first test, the wholly clustered array methods presented in [2], [3], i.e. the fully regular clustered elements (FRCE)

method and the fully irregular clustered elements (FICE) method, are enhanced. Figure 4 shows the results of the proposed CCA method, compared with the results of FRCE and FICE methods in terms of radiation pattern and amplitude-only taper distribution. For the FRCE structure, the number of clusters of $C = 5$ on each side, and the number of elements in each cluster of $c = 4$. Here, we expect the appearance of uncontrolled periodic high-side lobes, and these lobes in-

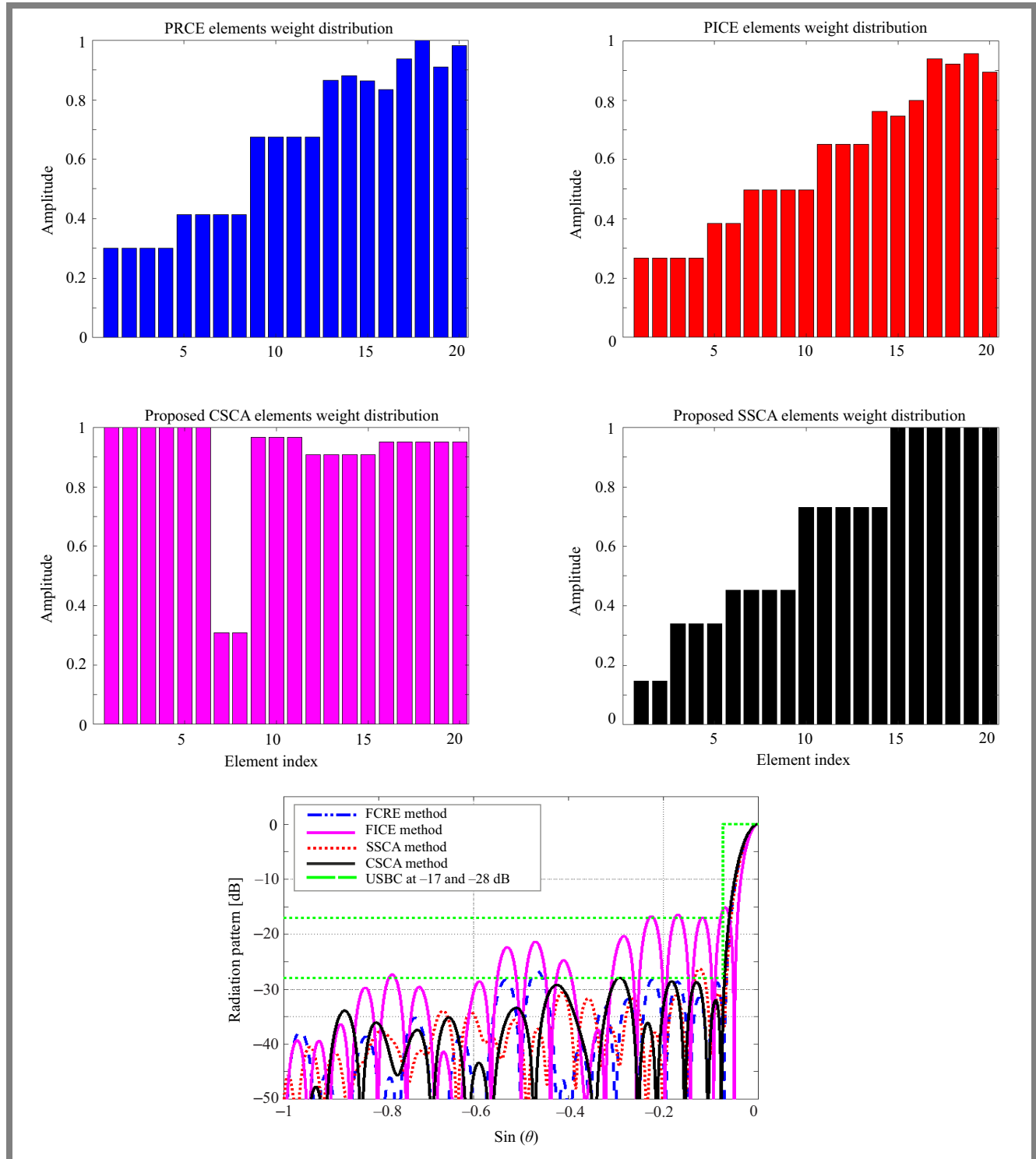


Fig. 5. Comparisons between the proposed SSCA and CSCA architectures, and the existing structures from [2], [3].

crease as the number of elements in the cluster grows. To solve this problem, the FICE method described in [3] is used. Here, the method depends on constructing irregular and unordered clusters, so there is a possibility of exceeding one of the periodic side lobes of the specified constraint border, especially if the lobes' constraint border is less than -25 dB. The proposed CCA approach solves the problem of periodic sidelobes definitively with an orderly arrangement of the complete feeding network, as such a solution facilitates its construction in practice.

In the second test, SSCA and CSCA architectures are illustrated. To show the performance of these two architectures in terms of obtaining the desired radiation pattern and simplifying the feed network significantly, these two methods are also compared with the partially regular clustered elements (PRCE) approach and the partially irregular clustered elements (PICE) scenario shown in papers [2] and [3], respectively. Through these four methods (PRCE, PICE, SSCA, and CSCA), we find that the two proposed architectures offer a significant reduction in the feeding network while reducing the level of side lobes (to a degree that is close to the one achieved by the other two methods). Figure 5 shows the results achieved by those four methods with $2N = 40$, $C = 4$, and the number of individually excited elements $b = 6$. In PRCE and PICE methods, the individual elements are optimized, while in the two proposed methods, they are fed uniformly, without improvement. From Fig. 5 it may be seen that the level of the side lobes can be reduced to -28 dB in the three methods (PRCE, PICE, and SSCA), and to -17 dB in the case of the CSCA method, because the weights of the individual elements at the sides are higher than the weights of the clusters in the middle.

The complexity ratio of the three proposed methods (CCA, SSCA, and CSCA) may be defined as the ratio of the number of ascending clusters to the total number of elements in the main array:

$$\text{Complexity of the CCA SSCA and CSCA methods} = \frac{C}{2N} \times 100 [\%]. \quad (7)$$

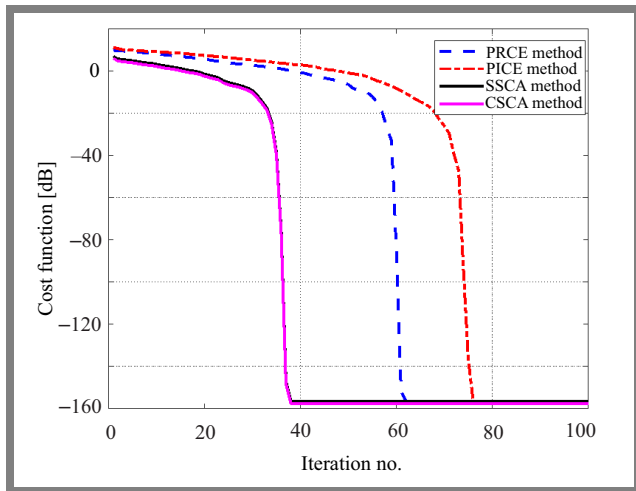


Fig. 6. Comparisons between the proposed SSCA and CSCA architectures, and the existing structures from [2], [3].

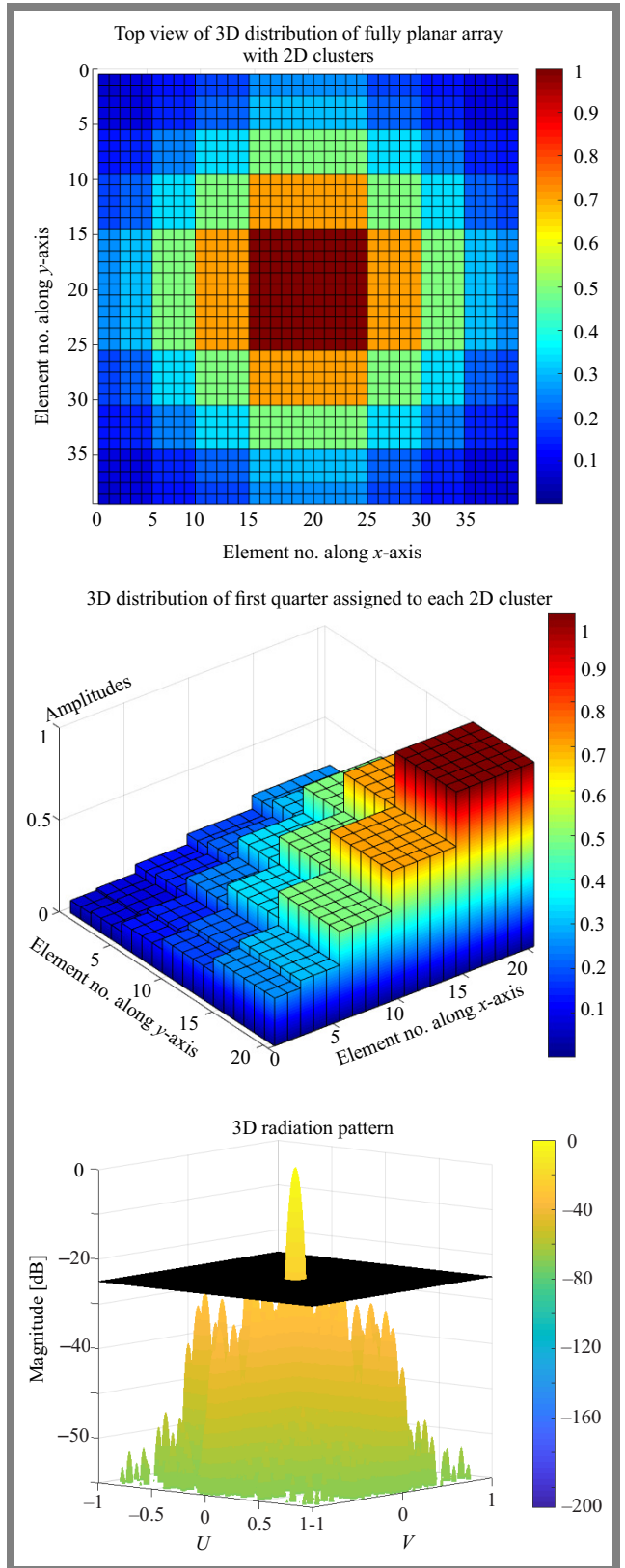


Fig. 7. Comparisons between the proposed SSCA and CSCA architectures, and the existing structures from [2], [3].

The complexity percentage in the CCA method was 12.5% (5 clusters were used symmetrically on each side), while in the SSCA and CSCA methods it equaled only 10%

(4 clusters were used symmetrically on each side with 6 uniform exciting elements), with the individual elements not taken into account in the complexity calculation. The computation of the complexity percentage in the PRCE and PICE methods will consider the individual elements, so the complexity percentage will increase. Its level amounted to 27.5%. It is noticed that there is a clear reduction in the complexity ratio in the proposed methods.

Figure 6 shows the variation of the cost function with the number of iterations for the proposed methods, comparing it with the methods described in [2] and [3]. It can be seen from this figure that the two methods (SSCA and CSCA) require the least number of iterations to converge, due to having the lowest number of degrees of freedom compared to other methods. This offers other advantages to the methods proposed in this paper.

The idea of an unequal ascending clustered array was extended to the two-dimensional (planar) array with a size of 40×40 , and the results are illustrated in Fig. 7. The number of elements in each 2D cluster is selected to build two-dimensional ascending clusters. Here, the SSCA method was adopted when building the planar array. It is observed from Fig. 7 that 11×11 elements in the center of the planar clustered array remain excited regularly, in addition to the presence of symmetry in the four quadrants of the amplitude distribution, which means that there is a significant reduction in complexity in practice. Moreover, radiation-related properties have been satisfied as well, as evidenced by the three-dimensional array pattern.

4. Conclusion

It is clear from the simulation tests that the desired beam pattern can be obtained with the required restrictions using three types of wholly and partially clustered architectures, depending on the ascending order of the clusters. The use of such architectures has resulted in a significant reduction of the exciting network and the associated costs. The complexity percentage in the proposed CCA method equaled 12.5%, amounting to 10% in the proposed CSCA and SSCA methods. The use of an ascendingly arranged cluster architecture helps completely suppress the periodic side-lobe levels that were inevitable with regular and irregular clusters.

In addition, the suggested concept was expanded and deployed in 2D planar arrays, where each cluster's components may be created as tiles. Subsequently, the array radiation layout was adjusted by means of optimal sub-arraying arrangements. As a result, the suggested concept may be effectively implemented in large solutions, such as massive MIMO arrays.

References

- [1] X. Zhao, Q. Yang, and Y. Zhang, "Synthesis of sub-arrayed linear array via l1-norm minimization compressed sensing method", *IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP)*, 2018 (<https://doi.org/10.1109/APCAP.2018.8538246>).
- [2] A.J. Abdulqader, J.R. Mohammed, and R.H. Thaher, "Antenna pattern optimization via clustered arrays", *Progress in Electromagnetics Research M*, vol. 95, pp. 177–187, 2020 (<https://doi.org/10.2528/PIERM20042307>).
- [3] A.J. Abdulqader, J.R. Mohammed, and R.H. Thaher, "Unconventional and irregular clustered arrays", *1st International Ninevah Conference on Engineering and Technology (INCET2021)*, 2021 (<https://doi.org/10.1088/1757-899X/1152/1/012003>).
- [4] R.J. Mailloux, S.G. Santarelli, T.M. Roberts, and D. Luu, "Irregular polyomino-shaped subarrays for space-based active arrays", *Int. J. of Antennas Propag.*, vol. 2009, Article ID 956524, 2009 (<https://doi.org/10.1155/2009/956524>).
- [5] X. Chen, Y. Sun, F. Xu, and X. Yang, "Sub-array partition method based on particle swarm optimisation for large aperture phased array radar", *The Journal of Engineering*, vol. 2019, no. 19, pp. 6318–6321, 2019 (<https://doi.org/10.1049/joe.2019.0278>).
- [6] J.R. Mohammed, R.H. Thaher, and A.J. Abdulqader, "Linear and planar array pattern nulling via compressed sensing", *Journal of Telecommunications and Information Technology*, no. 3, pp. 50–55, 2021 (<https://doi.org/10.26636/jtit.2021.152921>).
- [7] X. Zhao, Q. Yang, and Y. Zhang, "Synthesis of minimally sub-arrayed linear arrays via compressed sensing method", *IEEE Antennas and Wireless Propag. Lett.*, vol. 18, no. 3, pp. 487–491, 2019 (<https://doi.org/10.1109/LAWP.2019.2894826>).
- [8] G. Oliveri, M. Salucci, and A. Massa, "Synthesis of modular contiguously clustered linear arrays through a sparseness-regularized solver", *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 10, pp. 4277–4287, 2016 (<https://doi.org/10.1109/TAP.2016.2595623>).

Ahmed Jameel Abdulqader, Ph.D.

Lecturer

 <https://orcid.org/0000-0003-4938-1441>

E-mail: ahmed.abdulqader@uoninevah.edu.iq

College of Electronics Engineering, Ninevah University, Mosul, Iraq

Jafar Ramadhan Mohammed

Professor and Vice Chancellor for Scientific Affairs

 <https://orcid.org/0000-0002-8278-6013>

E-mail: jafar.mohammed@uoninevah.edu.iq

College of Electronics Engineering, Ninevah University, Mosul, Iraq

Yessar E. Mohammad Ali

Assistant Professor and Head of Computer and Communications Engineering Department

E-mail: yessar.mali@nawroz.edu.krd

College of Engineering, Nawroz University, Duhok, Iraq

Synthesis of Reconfigurable Multiple Shaped Beams of a Concentric Circular Ring Array Antenna Using Evolutionary Algorithms

Sanjay Kumar Dubey¹ and Debasis Mandal²

¹K. K. University, Bihar, India,

²Department of Electronics and Communication Engineering, Ramchandra Chandravansi University, Bishrampur, Palamu, Jharkhand, India

<https://doi.org/10.26636/jtit.2023.168022>

Abstract — The approach described in this paper uses evolutionary algorithms to create multiple-beam patterns for a concentric circular ring array (CCRA) of isotropic antennas using a common set of array excitation amplitudes. The flat top, cosec², and pencil beam patterns are examples of multiple-beam patterns. All of these designs have an upward angle of $\theta = 0^\circ$. All the patterns are further created in three azimuth planes ($\varphi = 0^\circ, 5^\circ$, and 10°). To create the necessary patterns, non-uniform excitations are used in combination with evenly spaced isotropic components. For the flat top and cosecant-squared patterns, the best combination of common components, amplitude and various phases is applied, whereas the pencil beam pattern is produced using the common amplitude only. Differential evolutionary algorithm (DE), genetic algorithm (GA), and firefly algorithm (FA) are used to generate the best 4-bit discrete magnitudes and 5-bit discrete phases. These discrete excitations aid in lowering the feed network design complexity and the dynamic range ratio (DRR). A variety of randomly selected azimuth planes are used to verify the excitations as well. With small modifications in the desired parameters, the patterns are formed using the same excitation. The results proved both the efficacy of the suggested strategy and the dominance of DE over GA as well as FA.

Keywords — cosec² beam, differential evolution algorithm (DE), firefly algorithm (FA), flat top beam, genetic algorithm (GA), multiple shaped beam patterns, pencil beam

1. Introduction

The concentric ring array antenna with 360° azimuthal symmetry is highly helpful in wireless communications, especially for satellite and radar-related applications. Furthermore, a large sidelobe with a significant ripple issue was observed while producing the structured beams, including the flat-top, cosec², as well as pencil-shaped beam, at a variety of azimuth planes with similar excitation amplitudes. A variety of methods exist for creating array patterns and multiple beam patterns that have been documented in literature. These include, inter alia, [1]–[14]. With the help of a simulated annealing optimization technique, Diaz *et al.* suggested, in

paper [5], a technique which enables the effective synthesis of multiple-patterns of linear array antennas characterized by constant amplitude dispersals that differ in phase only. An efficient iterative approach for the power synthesis of programmable array antennas is provided by Buttazzoni and Vescovo in [6]. The approach is applicable to arrays of any shape, covering situations when there are many items present. The excitation magnitude of every array element remains unchanged throughout the configuration process thanks to phase-only control, which determines how reconfigurability is accomplished.

According to a procedure suggested and created by Lei *et al.* in [7], an isotropic array of linear antenna components may produce a cosecant-squared beam pattern by including optimal amplitudes and phases. Paper [8] shows that a single array antenna may need to emit more than one pattern per area or financial constraints. Each pattern is chosen and electronically controlled in way, where only the phase may be changed. With such a scenario, a synthesis approach is provided that can simultaneously calculate the common amplitude and the individual phases. Additionally, the strategy is adaptable enough to accommodate extra restrictions and permits an effective deployment. Phase-only as well as amplitude-phase syntheses of an identical dual beam pattern of a linear array antenna employing real-valued and floating-point genetic algorithms are compared by Mahanti *et al.* in paper [9], with sum and sector array patterns serving as two examples.

To create customized beam patterns, uniform as well as Gaussian distributions of identical magnitude are used (flat top and cosec²). Additionally, various beam patterns require distinct kinds of phases. Durr *et al.* in [10] proposed the Woodward-Lawson approach to identify such excitations. With the use of evolution algorithms, Dubey and Mandal [11] describe a pattern synthesis technique for creating dual array patterns of an isotropic planar array antenna of rectangular shape. The cosec² pattern and a pencil beam pattern are two examples of dual-beam patterns that are both directed at an upper angle of twenty degrees. Chakraborty *et al.* in [12] pro-

vide a technique for figuring out the phase allocation of an array having a linear structure of discrete excitations with a given array pattern with specified aperture magnitude distributions. The infinite series, whose terms are individually expressed as integrals, is created from the specific series that represents the beam pattern. The asymptotic assessment of the integral using the stationary phase approach serves as the basis for synthesis. Sector and cosecant beam computed outcomes are displayed. Again, for shaped-beam radiation patterns and a synthesis utilizing both line sources or evenly separated arrays, Stutzman [13] introduces the iterative sampling technique. A number of corrective patterns are applied to an original pattern that approximates the intended pattern. This method is used repeatedly until the required performance is attained. A matching set of adjustments leads to the discovery of the current distribution. Numerous examples demonstrate the possibility of obtaining patterns having either low major-beam ripple (Δ), low sidelobe level, or abrupt separation from the main lobe.

Using the concentric ring array antenna's best combinations of radial magnitudes and phases, Chatterjee *et al.* [14], [15] suggested a technique for producing twin beams. The gravitational search algorithm (GSA) is used to generate such excitations. They also suggested using the firefly algorithm (FA) to create the dual configuration for an array of concentric rings of antennas. The excitations employed in the circles of the CCAA uphold the supplied components, which are isotropic. By applying particular modes or switching combinations, the intended patterns are produced. Mangoud and Elragal [16] suggested using the enhanced practical swarm optimization (EPSO) method to create patterns from linear arrays. To increase the precision of the converging of traditional PSO, updated formulae of the global optimal particle location and velocities are adjusted. To produce adaptable wide nulls steered under the limitations of peak SLL as well as lowest principal beam width, the cumulative several deep nulls technique and the straight weights perturbations technique are taken into consideration. By carefully adjusting an element's existing magnitude or complex weights, beginning with the original Chebyshev pattern, one or more broad nulls may be produced, which can then be either symmetrically or asymmetrically oriented with respect to the primary beam.

Using a rectangular array of isotropic antennas, Mandal *et al.* offer, in [17], a pattern synthesis technique focused on the differential evolution algorithm (DE) to generate twin beam patterns. These include the pencil beam pattern plus the cosec² pattern. Such patterns are created by determining the best possible combinations of identical element magnitudes and phases for the cosec² pattern alone. To simplify the construction of the feed network, 5-bit discrete phases with 4-bit discrete amplitudes are employed. In article [18], the methods for keeping lower minor lobes in concentric ring array (CRA) antennas are thoroughly reviewed. The level of peak SLL reduction across a predetermined round field-of-view is suggested for directional uniform-amplitude concentric circular ring array antennas, thus introducing an iterative convex

optimization (ICO)-driven array design synthesis methods. The ICO method optimizes the circle radii with the rotational angle for a particular size of rings and the size of equally balanced items within each ring, with potential limits on the minimum component spacing and greatest array size. Yi Jiang *et al.* [19] introduced modified integer genetic algorithm that gets around the ineffective searching and the significant processing burden of genetic algorithms. IIGA is used in [19] to create uniformly stimulated concentric ring arrays having aperture diameters ranging from tiny to larger, where the diameter of the array is nearly 50λ . All of the arrays that were produced adhere to the half-wavelength optimum component separation requirement.

In this paper, flat top, cosec², and pencil beam patterns of an antenna with sixty isotropic components are obtained by applying evolutionary algorithms to identify the best combination of common element magnitudes for all patterns as well as a band of phases for the flat top and cosec² shaped beam. Here, the geometry of the array is a concentric circle. The excitations produced through DE, FA, and GA were used to construct the patterns in three predetermined azimuth planes. With this method, instead of using a singular φ plane, it is additionally confirmed that the patterns maintain their required characteristics across a variety of azimuth planes. These have also been demonstrated by choosing a random set of φ planes, applying the identical excitations with each evolutionary algorithm, and obtaining patterns that are comparable but with a few slight differences. For a reduced dynamic range ratio (DRR), the acquired magnitude (4-bit) and phases (5-bit) are indeed digitized. Such discontinuous excitations are employed to reduce the feed network's design complexity, since they produce fewer attenuators as well as phase shifters due to the reduced DRR. These three evolutionary algorithms: differential evolution algorithms (DE), firefly algorithms (FA), and genetic algorithms are compared in terms of performance.

2. Problem Formulation

An array with a concentric circular layout, consisting consists of isotropic components is depicted in Fig. 1. The array pattern in the far field may be put into the following terms [1]–[3]:

$$AF(\theta, \varphi) = \sum_{m=1}^M \sum_{n=1}^{N_m} I_{mn} e^{j[kr_m \sin \theta \cos(\varphi - \varphi_{mn}) + \alpha_{mn}]}, \quad (1)$$

where:

- m = quantity of concentric circles,
- in the m -th ring, the number of isotropic components is N_m ,
- I_{mn} = excitation magnitudes of mn -th component,
- d_m = inter-component rotational gapping (0.5λ),
- θ, φ = vertical and horizontal angle, item $\varphi_{mn} = \frac{2n\pi}{N_m}$ the arc stand of mn -th component; the range on n mentioned as $1 \leq n \leq N_m$,
- $r_m = \frac{N_m d_m}{2\pi}$ radial distance of m -th ring,
- the wave number is denoted as $k = \frac{2\pi}{\lambda}$.

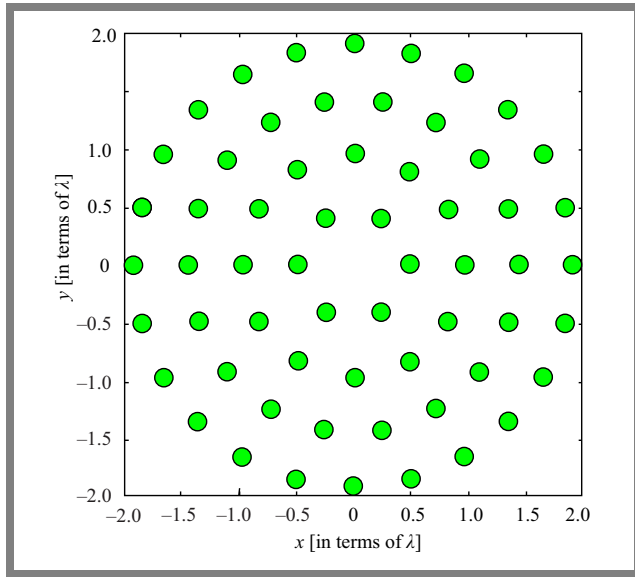


Fig. 1. Geometry of a CCRA array of 60 isotropic elements.

The objective functions of multiple shaped beam patterns are expressed as:

$$F(\rho) = k_1 \left\{ peakSLL^{d_1} - \max_{\theta \in A_1} [AF_{dB}^\rho(\theta, \varphi)] \right\}^2 + k_2 \times \Delta_1 + k_3 \left\{ peakSLL^{d_2} - \max_{\theta \in A_2} [AF_{dB}^\rho(\theta, \varphi)] \right\}^2 + k_4 \times \Delta_2 + k_5 \left\{ peakSLL^{d_3} - \max_{\theta \in A_3} [AF_{dB}^\rho(\theta, \varphi)] \right\}^2, \quad (2)$$

where Δ_1 and Δ_2 are defined as:

$$\Delta_1 = \sum_{\theta_{ripple} \in \{0^\circ \text{ to } 30^\circ\}} |AF_{dB}^\rho(\theta_{ripple}, \varphi) - D_{1dB}(\theta_{ripple}, \varphi)|, \quad (3)$$

$$\Delta_2 = \sum_{\theta_{ripple} \in \{-15^\circ \text{ to } +15^\circ\}} |AF_{dB}^\rho(\theta_{ripple}, \varphi) - D_{2dB}(\theta_{ripple}, \varphi)|. \quad (4)$$

In Eq. (2), ρ is the unspecified parameter combination for both techniques and is expressed as:

$$\rho = I_{mn}, \alpha 1_{mn}, \alpha 2_{mn}; \quad \begin{cases} 1 \leq m \leq M, \\ 1 \leq n \leq N_m, \end{cases} \quad (5)$$

$peakSLL^{d_1}$, $peakSLL^{d_2}$ and $peakSLL^{d_3}$ are the expected outcomes of max SLL, while A_1 , A_2 and A_3 are the sidelobe regions for the cosec² beam, flat top beam and pencil beam pattern, respectively. $D_{1dB}(\theta_{ripple}, \varphi)$ and $D_{2dB}(\theta_{ripple}, \varphi)$ are the desired patterns at azimuthal angles of 0°, 5° and 10° planes for cosec² beam and flat top array pattern, respectively. The estimated span of θ_{ripple} for the cosec² pattern ranges from 0° to 30° and for the flat top array pattern from -15 to 15 degrees. k_1 , k_2 , k_3 , k_4 , and k_5 are the weighting factors. With the use of several evolutionary algorithms, such as the DE, GA, and FA, these objective functions must be reduced to their minimum value.

3. Evolution Algorithm (EA)

3.1. Overview of Differential Evolution Algorithm

Evolutionary algorithms have been widely applied in a variety of problem areas over the past decade, and they have consistently found near-optimal solutions. DE is one such evolutionary method, and a full description and its features may be found in [20]–[26]. DE offers the following benefits: a straightforward construction, simplicity of use, efficiency, and durability for using in real-world situations. It has been used to find effective solutions to virtually intractable issues in a variety of scientific and engineering applications, without relying on expert knowledge or sophisticated design procedures. DE can provide the mechanism for obtaining the highest potential performance from a system that may be properly assessed. It uses mutations as a search query, with directing its searches towards the most probable locations in the available space. Using selection processes, genetic algorithms build a succession of populations. Crossover and mutation have been applied as search techniques in GAs.

The major difference between GAs and DE would be that GAs rely on the crossover approach to discover effective options, i.e. a stochastic and beneficial process for exchanging information across alternatives. So, even though evolutionary processes utilize mutation also as their primary search framework, DE seems to be an inhabitant searching technique which is capable of populating D-dimensional factor matrices with NP elements for each repetition. If there is no data about the situation, the starting crowd is chosen randomly. In a situation with an accessible preparatory answer, the initialization is often increased by combining uniformly distributed random variances to the existing preliminary solution. DE is based on such a revolutionary and experimental factor generation mechanism.

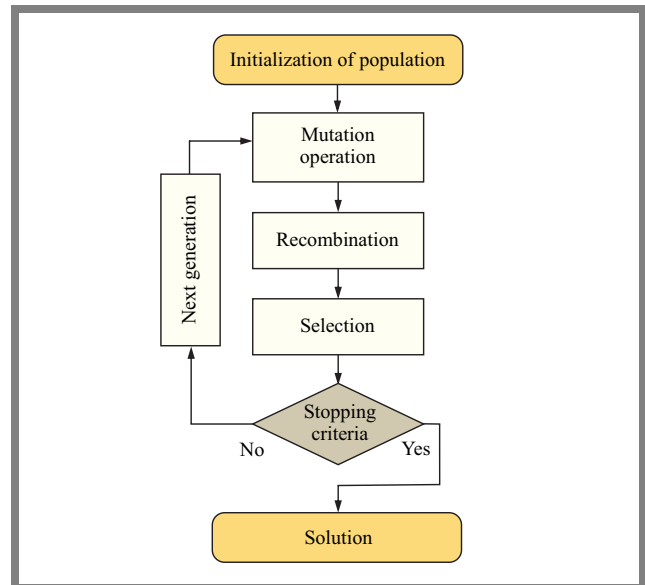


Fig. 2. Differential evolution algorithm (DE) flow chart.

Figure 2 depicts the DE process. The aim is to reduce function $f(Y)$, such as:

$$Y = [y_1, y_2, y_3, y_4, \dots, y_{D_1}], \quad (6)$$

where D_1 is the dimensionality or quantity of the search area. The population matrix as well as the assumed size N_1 of population are presented as follows:

$$y_{n,i}^g = [y_{n,1}^g, y_{n,2}^g, y_{n,3}^g, y_{n,4}^g, \dots, y_{n,D_1}^g], \quad (7)$$

where $n = 1, 2, 3, \dots, N_1$, and g is the number of iterations. The starting population is produced randomly between the higher y^H and lower bounds of y^L . The DE algorithm has three main steps: mutation, recombination, and selection which follow randomized initialization as well as creation in D_1 dimensionality searching field.

Three distinct vectors, y_{r1n}^g , y_{r2n}^g and y_{r3n}^g , were chosen from each input vector for the mutation procedure. Now it is possible to write the donor vectors s_n^{g+1} as:

$$s_n^{g+1} = y_{r1n}^g + F(y_{r2n}^g - y_{r3n}^g), \quad (8)$$

where scale factor F has a value ranging from 0 to 1.

Target vector $y_{n,i}^g$ and donor vector s_n^{g+1} are used to create the trial vector $t_{n,i}^{g+1}$ during the recombination procedure. The target function values are finally weighed against each target vector $y_{n,i}^g$ and trial vector $t_{n,i}^{g+1}$ in the selection process. For the following generation, those that provide the lowest function value, or the best fitness function, are chosen. These operations are repeated until the predetermined value of generation is reached. The outcome or the best objective function solution in the present population is written as $Y_{best,G}$.

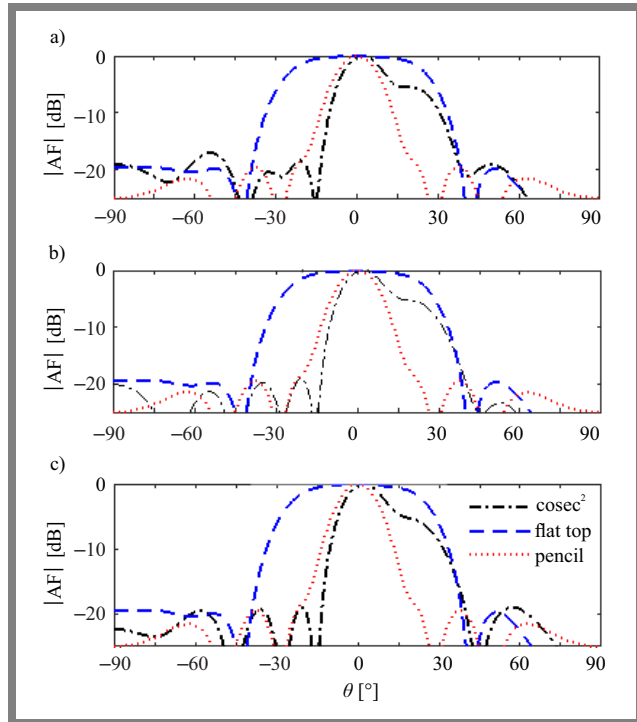


Fig. 3. DE utilized to produce various beam patterns for: a) $\varphi = 0^\circ$, b) $\varphi = 5^\circ$, and c) $\varphi = 10^\circ$ planes.

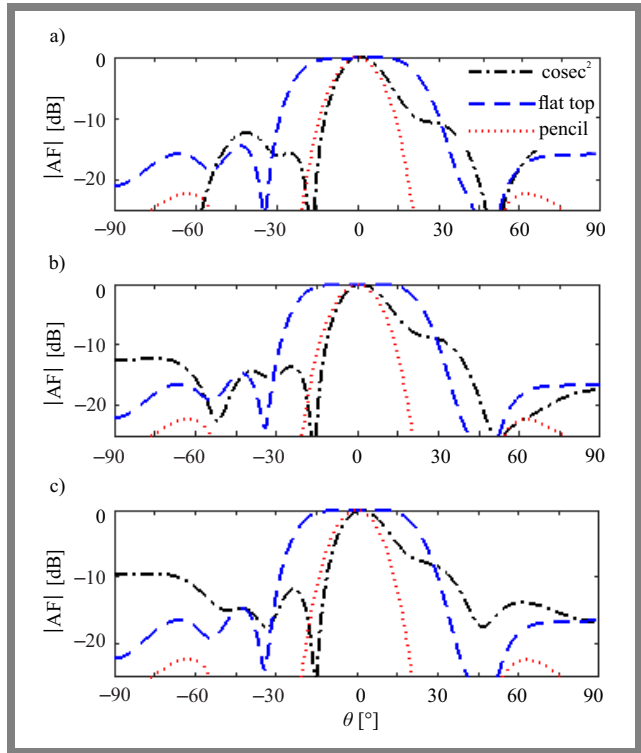


Fig. 4. FA utilized to produce various beam patterns for: a) $\varphi = 0^\circ$, b) $\varphi = 5^\circ$, and c) $\varphi = 10^\circ$ planes.

The population size for such an optimization model is set to 50, the scaling factor (F) is set at 0.8, and indeed the DE's configured crossover rate (CR) is 0.2. This non-linear challenge is solved using the $DE/best/1/bin$ technique, with iterations up to 3000 [20]–[26].

3.2. Parametric Setup of Genetic Algorithm and Firefly Algorithm

The genetic algorithm (GA) [4], [30]–[34] is a technique that is capable of solving not only particular problems. It also allows to study in a direction where practice of interconnection with natural-inspired modification and significance of these in computation. With the purpose of discovering true solutions of fitness functions, GA is a helpful searching methodology commonly applied in computing. It is an evolutionary algorithm which is stimulated by evolutionary biology. The most straightforward GA has three types of operators, including selection, crossover, and mutation. Usually, this evolution method involves a population that is generated randomly using specific individuals. Here, each possible solution is known as an individual and the group of all individuals is called the population. For the GA algorithm, a significantly higher proportion of individuals or chromosomes makes up a set of artificially adapted structures at position t . The initial batch of individuals is generated at random, so each individual is subsequently built using a selection process and a collection of operators. Step 1 creates the initial iteration, is followed by step 2 creating the next iteration, and so on. This means that GA utilities sequence t to start generating a correct number of chromosomes for another iterative pro-

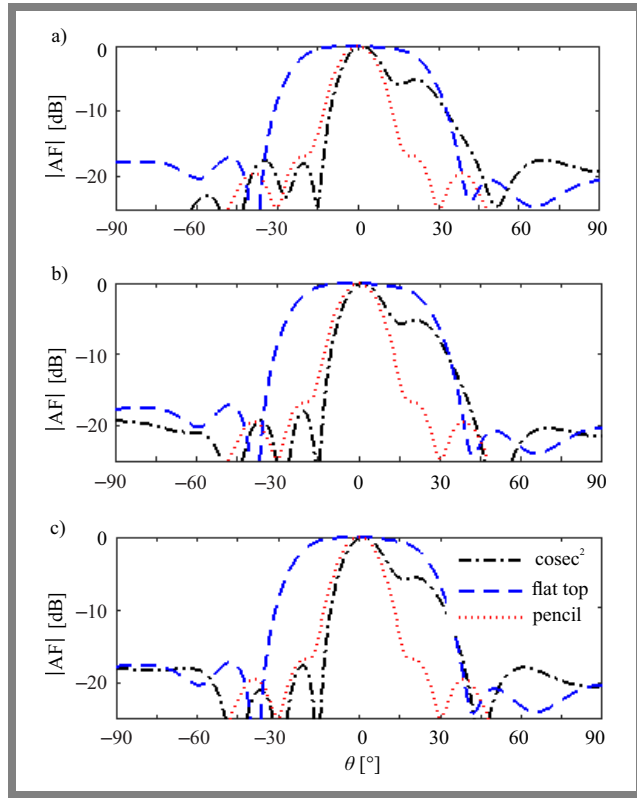


Fig. 5. GA utilized to produce various beam patterns for: a) $\varphi = 0^\circ$, b) $\varphi = 5^\circ$, and c) $\varphi = 10^\circ$ planes.

cess at time $t + 1$. The algorithm performs between 1 and T repetitions, where T is a fixed variable or the stage where the most powerful individual is found or another terminating requirement is fulfilled.

The swarm-based firefly algorithm (FA) [14], [27]–[29] is based on the assumption that the majority of fireflies of nearly two thousand species produce short and rhythmic flashes. Although these are unisexual insects, regardless of their sex, every firefly is attracting too many other fireflies. The pattern by which both males and females attract each other is created using rhythmic flashes. The rate of the flashing and the duration of each blink are monitored. Brighter fireflies are attracted to those fireflies with a lower level of brightness. The flashing light is produced by bioluminescence, and the precise functions of such a communication system are indeed up for discussion. However, these flashes serve two main purposes: to draw in potential prey and to draw in mates (communication). Flashing can also serve as a warning mechanism for safety.

The two-point crossover strategies are chosen with a population size of 50 in GA. The odds for both crossover and mutation are set at 0.08 and 0.01, respectively. For this suggested scenario, the roulette wheel method is considered throughout the selection procedure, and just a termination or ending condition of a limit of 3000 iterations is selected as the method applied. Additional GA variable configurations are chosen from the list provided in [4], [30]–[34]. In FA [14], [27]–[29], 50 fireflies were taken, with $\gamma = 0.25$, $\beta_0 = 0.20$, $\alpha = 1$. The search space dimension and 48 ran-

domly selected fireflies are used as the original population, with the search space dimensions being taken into account (3000 at the most iterations).

A four circular rings of array consist of sixty isotropic components have been estimated. The number of components in each round is a factor of six, i.e. $6m$, where m denotes the count of the circles. The radius for the first, second, third and fourth ring is 0.4775λ , 0.9549λ , 1.4324λ and 1.9099λ , respectively. The population for DE and FA for each individual case is:

$$P = [I_{1,1}, I_{1,2}, \dots, I_{1,6}, I_{2,1}, \dots, I_{2,12}, I_{3,1}, \dots, I_{3,18}, I_{4,1}, \dots, I_{4,24}, \alpha_{1,1,1}, \alpha_{1,1,2}, \dots, \alpha_{1,1,6}, \alpha_{1,2,1}, \dots, \alpha_{1,2,12}, \alpha_{1,3,1}, \dots, \alpha_{1,3,18}, \alpha_{1,4,1}, \dots, \alpha_{1,4,24}, \alpha_{2,1,1}, \alpha_{2,1,2}, \dots, \alpha_{2,1,6}, \alpha_{2,2,1}, \dots, \alpha_{2,2,12}, \alpha_{2,3,1}, \dots, \alpha_{2,3,18}, \alpha_{2,4,1}, \dots, \alpha_{2,4,24}] \quad (9)$$

The limits of the variable are:

$$0 \leq I_{mn} \leq 1; \quad -\pi \leq \alpha_{1mn} \leq \pi; \quad -\pi \leq \alpha_{2mn} \leq \pi, \quad (10)$$

where: $m = 1, 2, 3, 4$, for the first ring $n = 1, 2, \dots, 6$, for the second ring $n = 1, 2, \dots, 12$, for the third ring $n = 1, 2, \dots, 18$, and for the fourth ring $n = 1, 2, \dots, 24$.

The number of elements in this case is set to 60 and corresponds to the length of the search process which expands to 180.

4. Results

An antenna with 60 equally spaced isotropic components has been proposed as a concentric circular ring array. Here, M is the number of concentric rings. 4 rings have been considered,

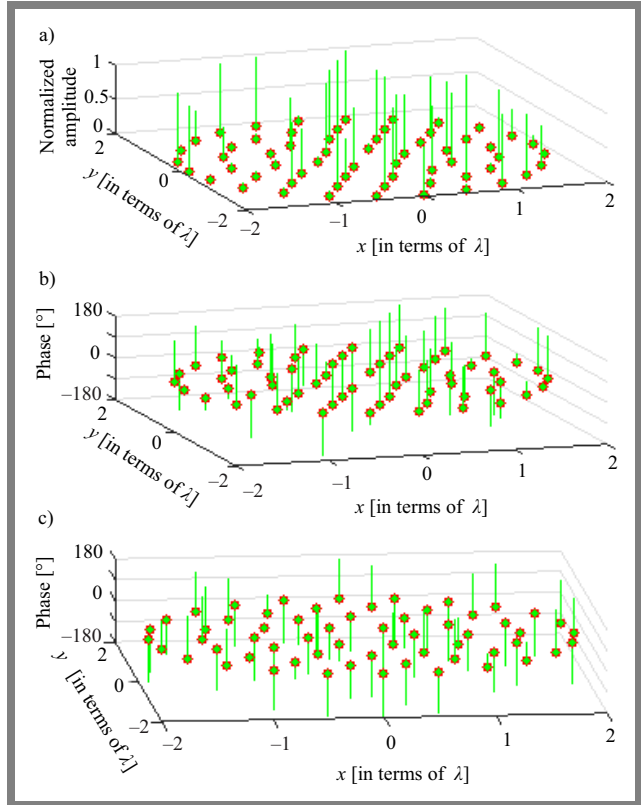


Fig. 6. Element-wise excitations using DE.

and each ring has N_m different components. This array has 6 isotropic elements in the first ring, 12 in the second, 18 in the third, and 24 in the fourth ring. The inter-element separation is taken into account and equals 0.5λ , as shown in Fig. 1. So, the ring radius becomes 0.4775λ for the first ring, 0.9549λ for the second ring, 1.4324λ for the third ring, and 1.9099λ for the fourth ring.

Tab. 1. Expected and achieved outcomes of the designed specifications.

φ cut	Beam type	Parameters [dB]	GA	FA	DE
0°	Flat top	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-17.04	-19.57
		Δ	Expected	0	0
			Achieved	20.85	21.76
	cosc ²	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-16.82	-17.05
		Δ	Expected	0	0
			Achieved	23.20	24.63
	Pencil	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-16.77	-18.08
5°	Flat top	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-17.13	-19.33
		Δ	Expected	0	0
			Achieved	21.93	22.79
	cosc ²	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-17.69	-18.17
		Δ	Expected	0	0
			Achieved	21.10	19.77
	Pencil	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-16.77	-18.08
10°	Flat top	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-17.16	-19.36
		Δ	Expected	0	0
			Achieved	21.91	22.77
	cosc ²	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-17.15	-18.05
		Δ	Expected	0	0
			Achieved	23.24	20.60
	Pencil	$SLL_{\max}$	Expected	-20.00	-20.00
			Achieved	-16.77	-18.08

Table 1 shows the design specifications for the multiple-beam patterns and the corresponding outcomes. It shows that the maximum side lobe level for the flat top beam pattern on 3 separate preset planes was achieved by applying DE and equaled -19.57 dB, -19.33 dB, and -9.36 dB, respectively, matching to its target value of -20.00 dB. The overall difference between the achieved and the intended values is calculated using the ripple Δ parameter within the angular region ($\theta = -15^\circ$ to $+15^\circ$). The value of ripple Δ is 21.76 dB,

22.79 dB, and 22.77 dB for $\varphi = 0^\circ$, $\varphi = 5^\circ$ and $\varphi = 10^\circ$, respectively.

It has also been observed from Tab. 1 that using GA, the peak side lobe levels for the flat top radiation pattern in three separate preset planes ($\varphi = 0^\circ$, $\varphi = 5^\circ$, and $\varphi = 10^\circ$) were achieved and correspond to the intended level of -20.00 dB at -17.4 dB, -17.13 dB, and -17.16 dB, with ripple Δ amounting to 20.85 dB, 21.93 dB, and 21.91 dB, respectively. The desired value of ripple Δ is 0 dB. The peak SLL as well as ripple Δ values (acquired and intended) obtained using FA are presented in Tab. 1. Maximum SLL is attained and equals -14.51 dB, -14.7 dB and -14.71 dB, with ripple Δ amounting to 21.95 dB, 18.76 dB, and 18.11 dB, respectively.

In contrast, the peak SLL for every pencil shaped beam pattern was measured to be -18.08 dB for DE, -16.77 dB for GA, and -22.30 dB for FA, when applied along the same predetermined azimuth planes.

From Tab. 1, it is additionally clear that the peak side lobe levels for the cosc² beam pattern along three separate preset planes were achieved employing the same DE method and equaled -17.05 dB, -18.17 dB, and -18.05 dB, respectively, in lieu of the desired value of -20.00 dB. Only within the angular range of ($\theta = 0^\circ$ to 30°), the quantity ripple Δ is used to evaluate the overall discrepancy between the achieved and the intended quantities. For the azimuth angles of 0, 5, and 10° , the values of ripple Δ are 24.63 dB, 19.77 dB and 20.60 dB, respectively.

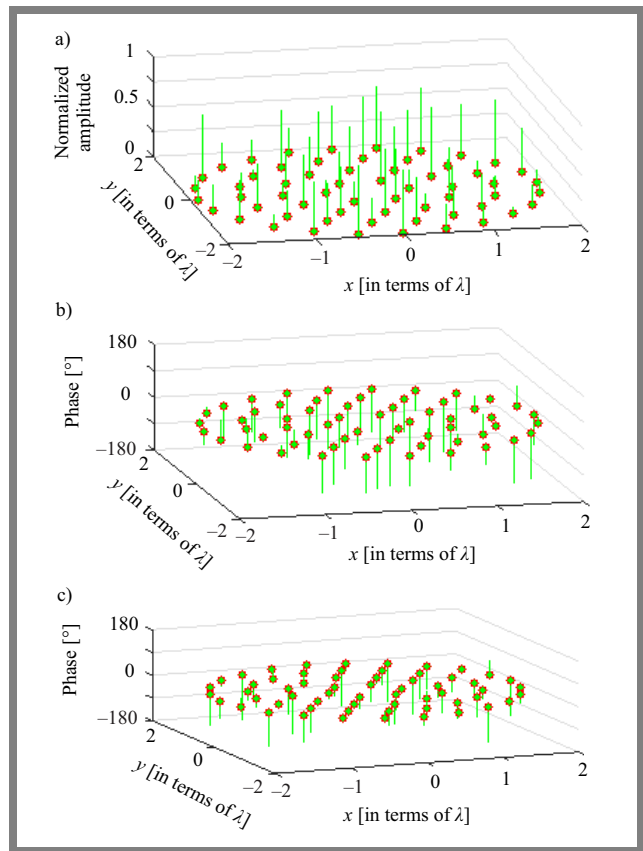


Fig. 7. Element-wise excitations using FA.

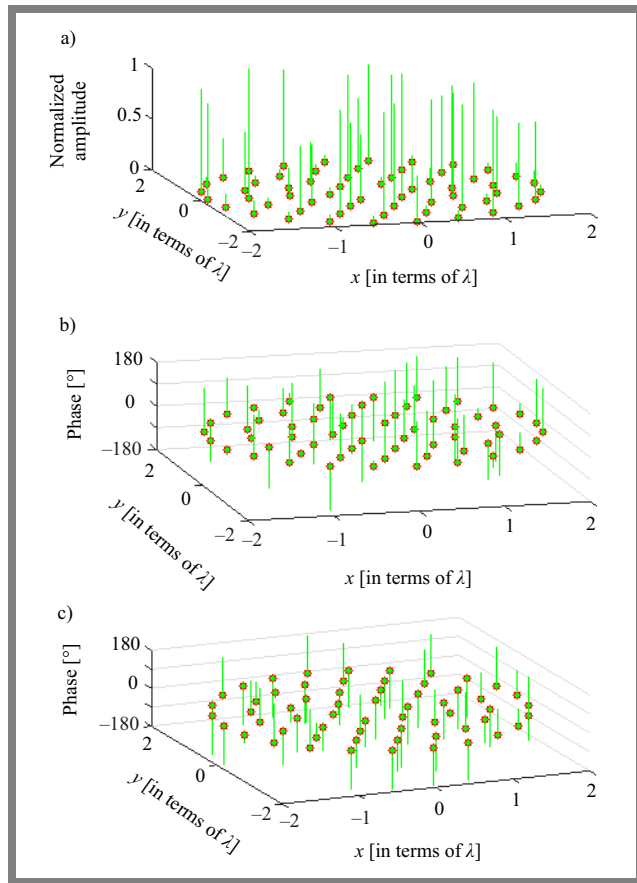


Fig. 8. Element-wise excitations using GA.

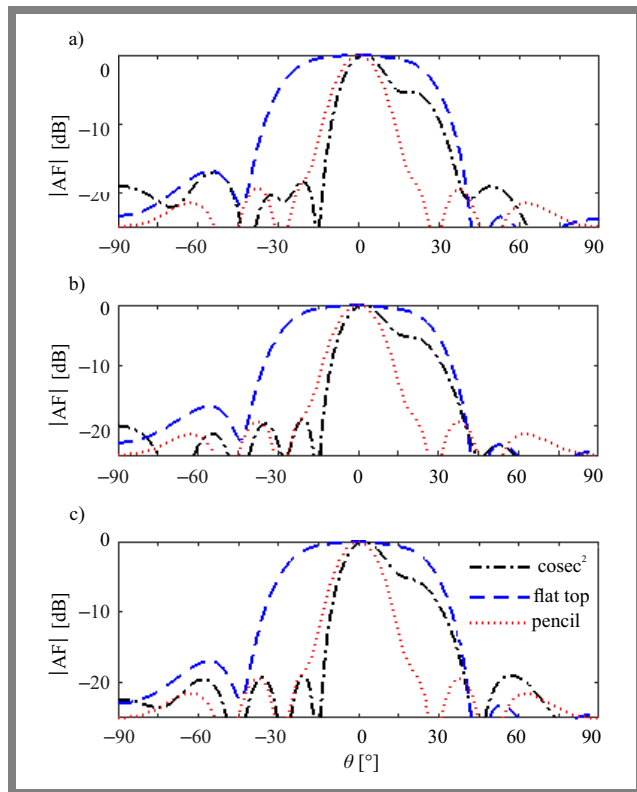


Fig. 9. Employing identical excitations acquired from DE, several shaped beam patterns have been achieved for three randomly selected planes: a) $\varphi = 2.5^\circ$, b) $\varphi = 7.5^\circ$, and c) $\varphi = 11.5^\circ$.

Additionally, it is noted that using GA, peak side lobe levels obtained for the cosec² beam pattern in three distinct pre-determined azimuth planes (0, 5, and 10°) are −16.82 dB, −17.69 dB, and −17.15 dB, respectively, compared to the desired level of −20.00 dB. Ripple Δ equals 23.20 dB, 21.10 dB, and 23.24, whereas the anticipated value is 0 dB. Similarly, Tab. 1 displays the expected and actual values of the FA's peak SLL and ripple Δ . Along with Δ of 74.36 dB, 47.60 dB, and 36.86 dB, the measured maximum SLL is −12.21 dB, −12.21 dB, and −09.46 dB.

The element-wise identical magnitudes of the array elements acquired employing DE, FA, and GA for creating multiple beam patterns are displayed in Figs. 6a, 7a, and 8a. These normalized magnitudes range from 0.625 to 1 and have 16 separate levels (2^4).

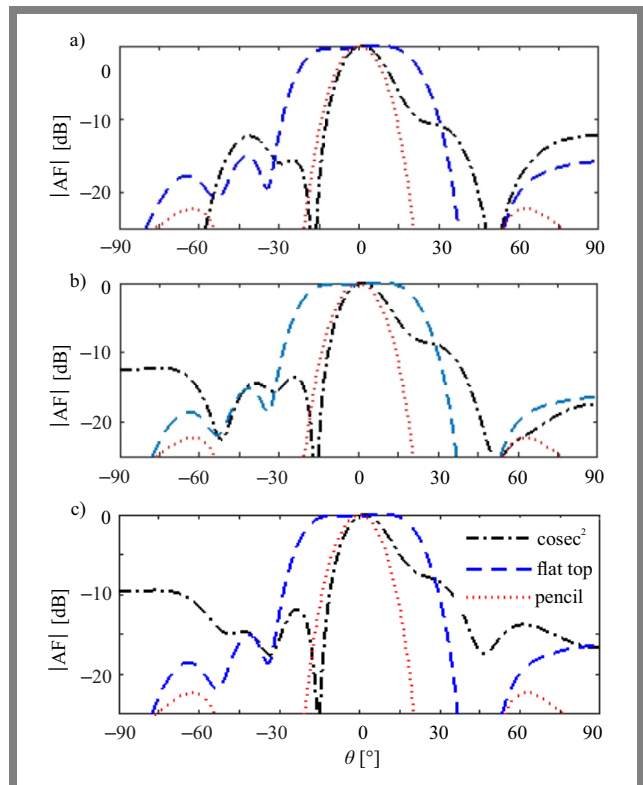


Fig. 10. Employing the identical excitations acquired from FA, several shaped beam patterns have been achieved for three randomly selected planes: a) $\varphi = 2.5^\circ$, b) $\varphi = 7.5^\circ$, and c) $\varphi = 11.5^\circ$.

The DRR is lower, as the lowest value of normalized amplitudes is 0.625 and the highest value is 1. The proposed technique uses 4-bit amplitudes of the components in the array, which guarantees that the DRR remains lower than the 16-point restriction, which really is beneficial in terms of a dependable network feed arrangement. Separate excitations as well the decreased quantity of attenuators and phase shifters, will additionally lower the price and complexity of the mechanism.

Similarly, the phases having 2^5 i.e. 32 levels within -180° to $+180^\circ$ are incorporated. Such 5-bit discrete phases are shown in Figs. 6b-c, 7b-c, and 8b-c for DE, FA, and GA, respectively.

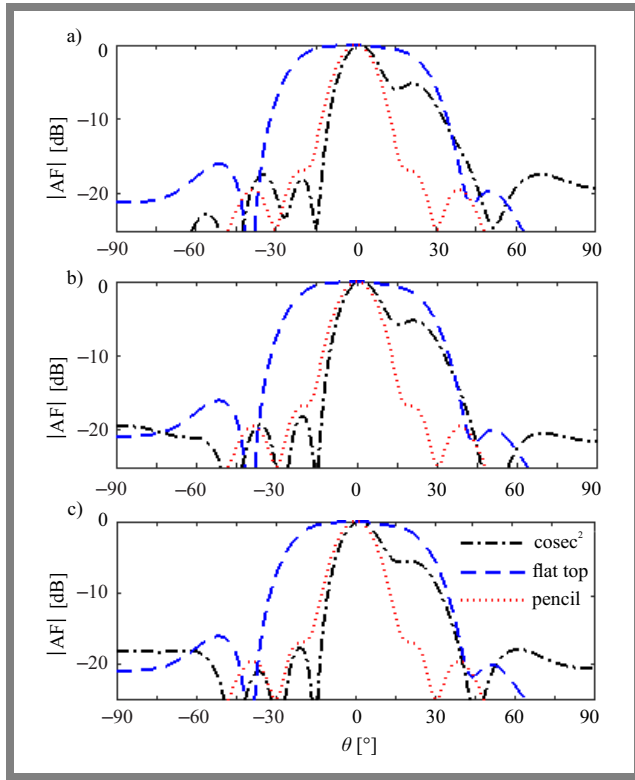


Fig. 11. Employing the identical excitations acquired from GA, several shaped beam patterns have been achieved for three randomly selected φ planes: a) $\varphi = 2.5^\circ$, b) $\varphi = 7.5^\circ$, and c) $\varphi = 11.5^\circ$.

Figure 6b shows the required phases applied for generating flat top beams, while Fig. 6c refers to the cosec² beam pattern with a common distribution of amplitudes. Figure 6a uses DE, whereas Figs. 7b and 8b show the required phases applied for generating flat top beams. Figures 7c and 8c relate to the cosec² beam pattern with a common distribution of amplitudes, while Figs. 7a and 8a are concerned with FA and GA, respectively.

In Figs. 9, 10, and 11, the same excitations result in somewhat different beam patterns in three randomly selected azimuth planes. The very first random azimuth angle in each figure is 2.5° ($0^\circ < 2.5^\circ < 5^\circ$ between the φ plane's boundaries), the second is 7.5° (inside the chosen φ plane of 5° and 10°), and the third is 11.5° ($> 10^\circ$ past the fixed azimuth plane), respectively.

The optimal excitations are produced using DE in both situations. Figure 3 depicts the multiple array pattern in pre-determined azimuth planes and Fig. 9 depicts the patterns in randomly chosen azimuth planes. Additionally, Figs. 10 and 11 depict multiple array patterns in randomly selected azimuth planes, while Figs. 4 and 5 depict the patterns in pre-determined φ planes. In both situations, the best excitations are produced using FA and GA in tandem.

The obtained flat top and cosec² beam patterns are following the desired beam patterns within the coverage range of -15° to $+15^\circ$ and 0° to 30° , respectively.

With all three EAs taken into consideration, the results of the design requirements are presented in Tab. 2 for random angles. Comparable design parameter values were found for φ planes

that were chosen at random and those that were predetermined. Therefore, it is not necessary to take into account all azimuth planes for such an array synthesis technique. However, there are some preset azimuth planes that provide a span where the pattern retains the required feature. The converging curve for DE, FA, and GA algorithms is shown in Fig. 12. Such convergence curves make it obvious that DE is superior to FA and GA, since its fitness value is lower.

Tab. 2. Expected and achieved outcomes of the designed specifications.

φ cut	Beam type	Parameters [dB]	DE	FA	GA
2.5°	Flat top	SLL_{max}	-16.91	-15.07	-15.99
		Δ	22.35	22.75	21.59
	consec ²	SLL_{max}	-17.05	-12.21	-16.82
		Δ	24.63	74.36	23.20
	Pencil	SLL_{max}	-18.08	-22.30	-16.77
7.5°	Flat top	SLL_{max}	-16.90	-15.11	-15.99
		Δ	23.38	20.47	22.65
	consec ²	SLL_{max}	-18.17	-12.21	-17.69
		Δ	19.77	47.60	21.10
	Pencil	SLL_{max}	-18.08	-22.30	-16.77
11.5°	Flat top	SLL_{max}	-16.93	-15.05	-16.02
		Δ	23.36	20.40	22.63
	consec ²	SLL_{max}	-18.05	-09.46	-17.15
		Δ	20.60	36.86	23.24
	Pencil	SLL_{max}	-18.08	-22.30	-16.77

Tab. 3. Comparison of DE, GA, and FA performance.

Algorithm	Best fitness (out of 15)	Worse	Mean	Standard deviation
DE	356.983	369.960	359.1682	3.1523
GA	390.444	408.513	397.5156	4.2589
FA	398.555	410.987	402.6547	4.4783

The accuracy of DE, FA, and GA is compared to show design-related difficulties (Tab. 3). The algorithm that turns out to be the best and most consistent in terms of this issue is DE, as it has a lower average fitness value than GA, PSO, and FA. As far as this kind of design requirements is concerned, Tab. 4 shows the results of the Wilcoxon rank sum test for

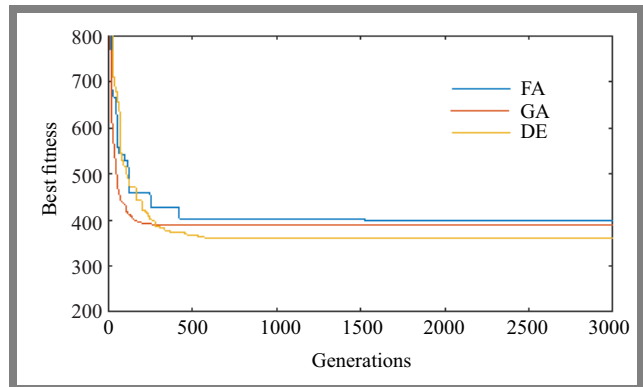


Fig. 12. Convergence curve for DE, GA, and FA.

Tab. 4. P-value for Wilcoxon's two-sided rank sum test.

Contrasting pair	P-value
DE-GA	$3.3928 \cdot 10^{-6}$
DE-FA	$4.4938 \cdot 10^{-6}$

DE-FA and DE-GA pairings. The greatest support for the null hypothesis stating that the best final fitness value achieved by the best algorithm is statistically meaningful comes from any generated values that are lower than 0.05 (5% importance threshold).

5. Conclusions

In three separate azimuth planes, a multiple-beam concentric circular ring array antenna has been created employing three well-known evolutionary algorithms. Three separate beam patterns (flat top, cosec², and pencil beam) are created in three distinct azimuth planes. To achieve the necessary pattern characteristics, every evolutionary algorithm creates the best 4-bit discrete amplitudes as well as 5-bit discrete phases. The approach adopted demonstrates its capability to generate the very same beam pattern not only in the pre-determined azimuth planes, but also across a span of azimuth planes, upon choosing and obtaining the identical pattern in random planes. By determining the ideal combination of array excitations using the DE, GA, plus the FA algorithm, such design requirements as peak side lobe level (peak SLL) and ripple Δ are decreased. By employing these digitized excitations, the dynamic range ratio is significantly decreased, thus requiring fewer attenuators and phase shifters in the circuit configuration.

References

- [1] C.A. Balanis, *Antenna Theory: Analysis and Design*, 2nd edition. New York: Wiley, 1997 (ISBN: 9780471592686).
- [2] R.S. Elliott, *Antenna Theory and Design, Revised Edition*. Wiley – IEEE Press, 2002 (ISBN: 9780471449966).
- [3] J.R. Mailloux, *Phased Array Antenna Handbook*, 2nd edition. Boston: Artech House, 2009 (ISBN: 9781580536899).
- [4] Z. Michalewicz, *Genetic Algorithms + Data Structures*. Berlin, Heidelberg, New York: Springer, 1999 (ISBN: 9783540606765).
- [5] X. Diaz, J.A. Rodriguez, F. Ares, and E. Moreno, "Design of phase-differentiated multiple pattern antenna arrays", *Microwave and Optical Technology Letters*, vol. 26, no. 1, pp. 52–53, 2000 ([https://doi.org/10.1002/\(SICI\)1098-2760\(20000705\)26:1<52::AID-MOP16>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1098-2760(20000705)26:1<52::AID-MOP16>3.0.CO;2-0)).
- [6] G. Buttazzoni and I.R. Vescovo, "Reconfigurable antenna arrays with phase-only control in the presence of near-field nulls", *Journal of Telecommunications and Information Technology*, no. 3, pp. 88–93, 2017 (<https://doi.org/10.26636/jtit.2017.118817>).
- [7] J. Lei, G. Fu, L. Yang, and D.M. Fu, "Wide band linear printed antenna array with low sidelobe cosecant square-shaped beam pattern", *Progress in Electromagnetics Research C*, vol. 15, pp. 233–241, 2010 (<https://doi.org/10.2528/PIERC10072506>).
- [8] O.M. Bucci, G. Mazzarella, and G. Panariello, "Reconfigurable arrays by phase-only control", *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 7, pp. 919–925, 1991 (<https://doi.org/10.1109/8.86910>).
- [9] G.K. Mahanti, A. Chakraborty, and S. Das, "Phase-only and amplitude-phase only synthesis of dual-beam pattern linear antenna arrays using floating-point genetic algorithms", *Progress in Electromagnetics Research*, vol. 68, pp. 247–259, 2007 (<https://doi.org/10.2528/PIERO6072301>).
- [10] M. Durr, A. Trastoy, and F. Ares, "Multiple-pattern linear antenna arrays with single prefixed amplitude distributions: modified Woodward-Lawson synthesis", *Electronics Letters*, vol. 36, no. 16, pp. 1345–1346, 2000 (<https://doi.org/10.1049/el:20000980>).
- [11] S.K. Dubey and D. Mandal, "Digitally controlled steered dual beam pattern synthesis of a rectangular planar array antenna in a range of azimuth plane using evolutionary algorithms", *Progress in Electromagnetics Research C*, vol. 114, pp. 185–202, 2021 (<https://doi.org/10.2528/PIERC21062303>).
- [12] A. Chakraborty, B. Das, and G. Sanyal, "Beam shaping using non-linear phase distribution in a uniformly spaced array", *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 5, pp. 1031–1034, 1982.
- [13] W. Stutzman, "Synthesis of shaped-beam radiation patterns using the iterative sampling method", *IEEE Transactions on Antennas and Propagation*, vol. 19, no. 1, pp. 36–41, 1971 (<https://doi.org/10.1109/TAP.1971.1139892>).
- [14] A. Chatterjee, G.K. Mahanti, and A. Chatterjee, "Design of a fully digital controlled reconfigurable switched beam concentric ring array antenna using firefly and particle swarm optimization algorithm", *Progress in Electromagnetics Research B*, vol. 36, pp. 113–131, 2012 (<https://doi.org/10.2528/PIERB11083005>).
- [15] A. Chatterjee, G.K. Mahanti, and R.P.S. Mahapatra, "Design of fully digital controlled reconfigurable dual-beam concentric ring array antenna using gravitational search algorithm", *Progress in Electromagnetics Research B*, vol. 18, pp. 59–72, 2011 (<https://doi.org/10.2528/PIERC10101806>).
- [16] M.A. Mangoud and H.M. Elragal, "Antenna array pattern synthesis and wide null control using enhanced particle swarm optimization", *Progress in Electromagnetics Research B*, vol. 17, pp. 1–14, 2009 (<https://doi.org/10.2528/PIERB09070205>).
- [17] D. Mandal, J. Tewary, K.S. Kola, and V.P. Roy, "Synthesis of dual beam pattern of planar array antenna in a range of azimuth plane using evolutionary algorithm", *Progress in Electromagnetics Research Letters*, vol. 62, pp. 65–70, 2016 (<https://doi.org/10.2528/PIERL16060802>).
- [18] Y. Aslan, A. Roederer, and A. Yarovoy, "Concentric ring array synthesis for low side lobes: An overview and a tool for optimizing ring radii and angle of rotation", *IEEE Access*, vol. 9, pp. 120744–120754, 2021 (<https://doi.org/10.1109/ACCESS.2021.3109171>).
- [19] Y. Jiang, S. Zhang, Q. Guo, and M. Li, "Synthesis of uniformly excited concentric ring arrays using the improved integer GA", *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1124–1127, 2016 (<https://doi.org/10.1109/LAWP.2015.2496173>).
- [20] 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 (<https://doi.org/10.1023/A:1008202821328>).
- [21] K.V. Price, R. Storn, and J.A. Lampinen, *Differential Evolution – A Practical Approach to Global Optimization*. New York: Springer, 2005 (ISBN: 9783540209508).
- [22] S. Das, A. Abraham, U.K. Chakraborty, and A. Konar, "Differential evolution using a neighborhood-based mutation operator", *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 526–553, 2009 (<https://doi.org/10.1109/TEVC.2008.2009457>).
- [23] J. Guo and J. Li, "Pattern synthesis of conformal array antenna in the presence of platform using differential evolution algorithm", *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 9, pp. 2615–2621, 2009 (<https://doi.org/10.1109/TAP.2009.2027046>).
- [24] Q. Xiaomei, H. Zhehuang, and C. Yidong, "An improved differential evolution algorithm based on adaptive parameter", *Journal of Control Science and Engin.*, vol. 2013, 2013 (<https://doi.org/10.1155/2013/462706>).

- [25] S. Yang, Y.B. Gan, and A. Qing, "Sideband suppression in time-modulated linear arrays by the differential evolution algorithm", *IEEE Transactions on Antennas and Propagations Letters*, vol. 1, no. 1, pp.173–175, 2002 (<https://doi.org/10.1109/LAWP.2002.807789>).
- [26] A. Massa, M. Pastorino, and A. Randazzo, "Optimization of the directivity of a monopulse antenna with a subarray weighting by a hybrid differential evolution method", *IEEE Transactions on Antennas and Propagations Letters*, vol. 5, no. 1, pp. 155–158, 2006 (<https://doi.org/10.1109/LAWP.2006.872435>).
- [27] X.S. Yang, "Firefly algorithms for multimodal optimization", in *Stochastic Algorithms: Foundations and Applications. LNCS*, vol. 5792, pp. 169–178, 2009 (https://doi.org/10.1007/978-3-642-04944-6_14).
- [28] S. Łukasik and S. Żak, "Firefly algorithm for continuous constrained optimization tasks", in *Computational Collective Intelligence. Semantic Web, Social Networks and Multi agent Systems. LNCS*, vol. 5796, pp. 97–106, 2009 (https://doi.org/10.1007/978-3-642-04441-0_8).
- [29] X.S. Yang, *Engineering Optimization: An Introduction with Meta-heuristic Applications*. Wiley, 2010 (ISBN: 9780470582466).
- [30] R.L. Haupt, "Introduction to genetic algorithms for electromagnetics", *IEEE Antennas and Propagation Magazine*, vol. 37, no. 2, pp. 7–15, 1995 (<https://doi.org/10.1109/74.382334>).
- [31] K.F. Man, K.S. Tang, and S. Kwong, "Genetic algorithms: Concepts and applications", *IEEE Transactions on Industrial Electronics*, vol. 43, no. 5, pp. 519–534, 1996 (<https://doi.org/10.1109/41.538609>).
- [32] J.M. Johnson and Y. Rahmat-Samii, "Genetic algorithms in engineering electromagnetics", *IEEE Antennas and Propagation Magazine*, vol. 39, no. 4, pp. 7–21, 1997 (<https://doi.org/10.1109/74.632992>).
- [33] D. Marciano, and F. Duran, "Synthesis of antenna arrays using genetic algorithm", *IEEE Antennas and Propagation Magazine*, vol. 42, no. 3, pp. 12–20, 2000 (<https://doi.org/10.1109/74.848944>).
- [34] M.A. Panduro, A.L. Mendez, R. Dominguez, and G. Romero, "Design of non-uniform circular antenna arrays for side lobe reduction using the method of genetic algorithms", *International Journal of Electronics and Communications*, vol. 60, no. 10, 713–717, 2006 (<https://doi.org/10.1016/j.aeeu.2006.03.006>).

Sanjay Kumar Dubey

Associate Professor

 <https://orcid.org/0000-0001-6066-6590>

K. K. University, Bihar, India

Debasis Mandal

Professor and Registrar

 <https://orcid.org/0000-0003-2170-3576>

E-mail: dr.debasis1984@gmail.com

Department of Electronics and Communication Engineering,
Ramchandra Chandravansi University,
Bishrampur, Palamu, Jharkhand, India

CG-SCMA Codebook Design Based on Maximized Euclidian Distance

Sura S. Mohamed and Hikmat N. Abdullah

Al-Nahrain University, College of Information Engineering, Baghdad, Iraq

<https://doi.org/10.26636/jtit.2023.168122>

Abstract — Sparse code multiple access (SCMA) is a multi-dimensional codebook based on a class of non-orthogonal multiple access (NOMA) technologies enabling the delivery of non-orthogonal resource elements to numerous users in 5G wireless communications without increasing complexity. This paper proposes a computer-generated sparse code multiple access (CG-SCMA) technique, where the minimum Euclidian distance (MED) of a star 16-point quadrature amplitude modulation is maximized by CG-SCMA, thus creating a complex SCMA codebook based on optimizing the difference between the first and other radiuses over rotated constellations. To specify the most suitable values for this constellation, it is divided into four sub-constellations using trellis coded modulation (TCM) in an effort to optimize MED. The new codebook has four sub-constellations with MED values of 3.85, 2.26, 2.26, and 3.85, respectively. Application of the message passing algorithm (MPA) ensures low complexity of the decoding process.

Keywords — codebook, Euclidian distance, MPA, multiple access, sparse code.

1. Introduction

The fifth generation (5G) of mobile communication systems as well as other important technologies, such as the Internet of Things, spread quickly throughout the world, combining various approaches and services to ensure high-throughput, mass-scale connectivity of contemporary networks [1], [2].

A novel code-domain non-orthogonal multiple access (NOMA) technique, referred to as sparse code multiple access (SCMA), has been created to allow for future massive machine-type communication networks. Multiple users in SCMA are distinguished by being assigned with unique sparse codebooks (CBs) serving as an upgraded version of code division multiple access (CDMA). The message passing algorithm (MPA), making use of the sparsity of CBs to produce error-free performance approaching the maximum potential of the receiver, is used at the receiver to perform efficient multiuser identification. To the best of our knowledge, despite multiple research initiatives undertaken in recent years, a thorough tutorial of SCMA covering the background, its fundamental concepts and new advancements, is still to be revealed [3].

The main contributions of this work can be summarized as follows:

- in contrast to the existing CG-SCMA codebook design method, this work uses a dedicated software to optimize

the difference values between the first and other radiuses over the rotated constellations,

- a modified CG-SCMA symmetric codebook architecture is suggested to increase the MinED of star 16-QAM by employing TCM to subdivide the constellation,
- the proposed SCMA codebook is designed to maximize the space between star 16-QAM constellations, so that BER is decreased along with the increase in SNR.

2. Related Work

Several papers may be found in the literature concerned with the process of designing an efficient SCMA scheme. Kosasih *et al.* [4] proposed a low-complexity detector technique for SCMA that is capable of addressing difficulties associated with MPA-SCMA and allows for the use of SCMA in very large MU-MIMO systems. Lou *et al.* [5] proposed a procedure for combining data bits in QAM symbol mapping and spreading, while providing direct mapping of bits to codewords. Liu *et al.* in [6] generated an optimized SCMA codebook by dividing round 16-QAM into several subsets to maximize the MinED. Tabra *et al.* [7] worked with star QAM constellations and proposed an improvement in MinED values using trellis coded modulation, while Chen and Chen [8] created a minimum complexity algorithm for nearly ideal CBs. In [9], Hussain *et al.* examined the effectiveness of various strategies of codebook generation and difficulties of SCMA optimization in real-world applications, across various applications. Zhang *et al.* [10] reviewed selected codebook designs and available SCMA detectors after presenting a comprehensive survey on the state-of-the-art in SCMA and of its fundamental principles. They also used the less complicated max-log-MPA approach instead of the MPA techniques. Waghmare *et al.* in [11] presented an effective method for designing and assigning the SCMA joint codebook based on the optimization of a 32-QAM constellation with large MinMED, high channel capacity and low detection complexity in the downlink network. In [12], Lin *et al.* developed a DL-SCMA deep neural network (DNN) to decode SCMA-modulated signals contaminated by additive white Gaussian noise (AWGN). In [13], Mohamed and Abdullah suggested a new SCMA codebook design method based on chaotic Arnold's cat map interleaving, characterized by a lower degree of complexity compared with the traditional method of interleaving.

3. SCMA Codebook Design

SCMA uses a codebook similar to the signature matrix relied upon in the low-density signatures (LDS) technique. It uses spreading to allow multiple users to share a single resource block. In NOMA, overloading is permitted when factor S/K exceeds 1. Consequently, K resource block (RB) accessed from S sources competes with the $K > S$ rule. The following equation may be used to determine the maximum number of sources for a given number of resources [6], [7], [14]:

$$S = C_k^{dv} = \frac{K.K-1 \dots K-dv+1}{dv.dv-1 \dots 1}, \quad (1)$$

where dv represents the number of resources that may be utilized by a single user.

For each user, $\log_2(L)$ bits are transmitted by a K -dimensional complex codebook with L codewords. The expression for the signature matrix for $S = 6$ users sharing $K = 4$ resource blocks is:

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \quad (2)$$

The amount of non-zero items in each column of the usual SCMA signature matrix (dv) refers to RB as a function node (FN). The number of non-zero elements in each row (df) refers to the number of users that each RB can serve, identifying each user node as a variable node (VN). Matrix (2) is used to map $\log_2(L)$ bits to a K -dimensional complex SCMA codebook with size L . To create the sparser codebook, the number of non-zero values in a codeword dv should be as low as possible, such that $dv < K$ [6], [7], [14]. For the purposes of codebook generation, constellation shaping gain and receiver complexity reduction, SCMA uses multi-dimensional constellations [8]. Even when two users share one resource block, the other resource block used is capable of distinguishing between them. The design of the codebook allows to reduce the number of constellation points used. This leads to lower complexity, hence log-MPA can be used for decoding data at the receiver side [7], [14]. Due to the sparse property of the SCMA codebook, the log-MPA technique can be applied at the receiver to reduce complexity while maintaining the desired performance.

3.1. The Proposed CG-SCMA Codebook Generating Method

To serve a lot of users while relying on a reasonable amount of resources, SCMA was developed – an approach that combines both modulation and multiple access and is capable of operating in a MIMO environment [6], [7], [14]. In a situation in which multiple users require data at the same time, but there are only a few resource blocks available, multiple access techniques allow to coordinate access to these resources by numerous users. Modulation and MA are combined in the

proposed SCMA codebook to provide users with access to a small number of resource blocks.

To allow this, an SCMA complex constellation codebook is created by a dedicated algorithm that maximizes the minimal Euclidian distance for better decoding results. Demodulation and demultiplexing are performed at the decoder using the log-MPA block and modified CG-SCMA. As a rule, the number of users in OMA ought to be lower or equal to the number of RBs, but overloading is permitted with NOMA, by a factor of $J/K > 1$. Therefore, there is competition among J mobile sources to provide K with access when $J > K$.

The proposed CG-SCMA scheme is characterized by the following parameters: $J, K, L, df, dv = 6, 4, 4, 3, 2$, respectively, and the proposed factor graph is shown in Fig. 1 [6]. Each source can access only dv RBs, with a non-zero dv value in the codeword. Two users are allowed to share two RBs. This is done based on the $\mathbf{F}$ matrix in Eq. (2) that uniquely maps each user to the RBs, where the set of the resources that may be accessed by each user is represented by “1” in $\mathbf{F}$. The star QAM constellation design is generated by extending the standard 16-QAM to four rings, as presented in Fig. 2 [6], [7].

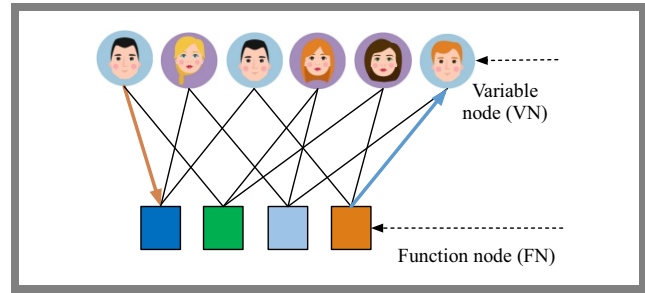


Fig. 1. Factor graph for $J = 6$ and $K = 4$.

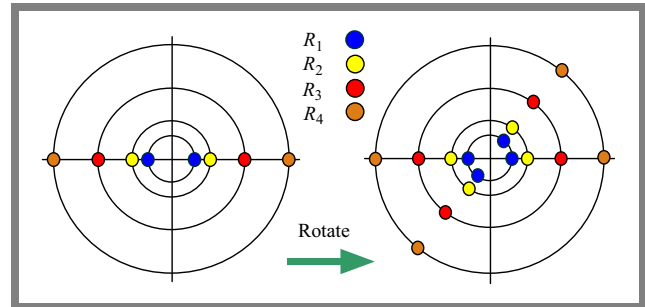


Fig. 2. 16-QAM modulation constellations [6], [7], [14].

The proposed CG-SCMA method follows the procedure presented below.

Let $R_1 = 1$, as proposed in [6], [7].

Let D_1 be the difference between R_2 and R_1 , and D_2 be the difference between R_3 and R_1 as:

$$D_1 = R_2 - R_1, \quad (3)$$

$$D_2 = R_3 - R_1. \quad (4)$$

Let us define the constellation's ratio between two successive points as $\frac{R_2}{R_1} = \frac{R_4}{R_3}$, which is called β [6], [7], and the ratio between non-successive points as $\frac{R_3}{R_1} = \frac{R_4}{R_2}$, which is

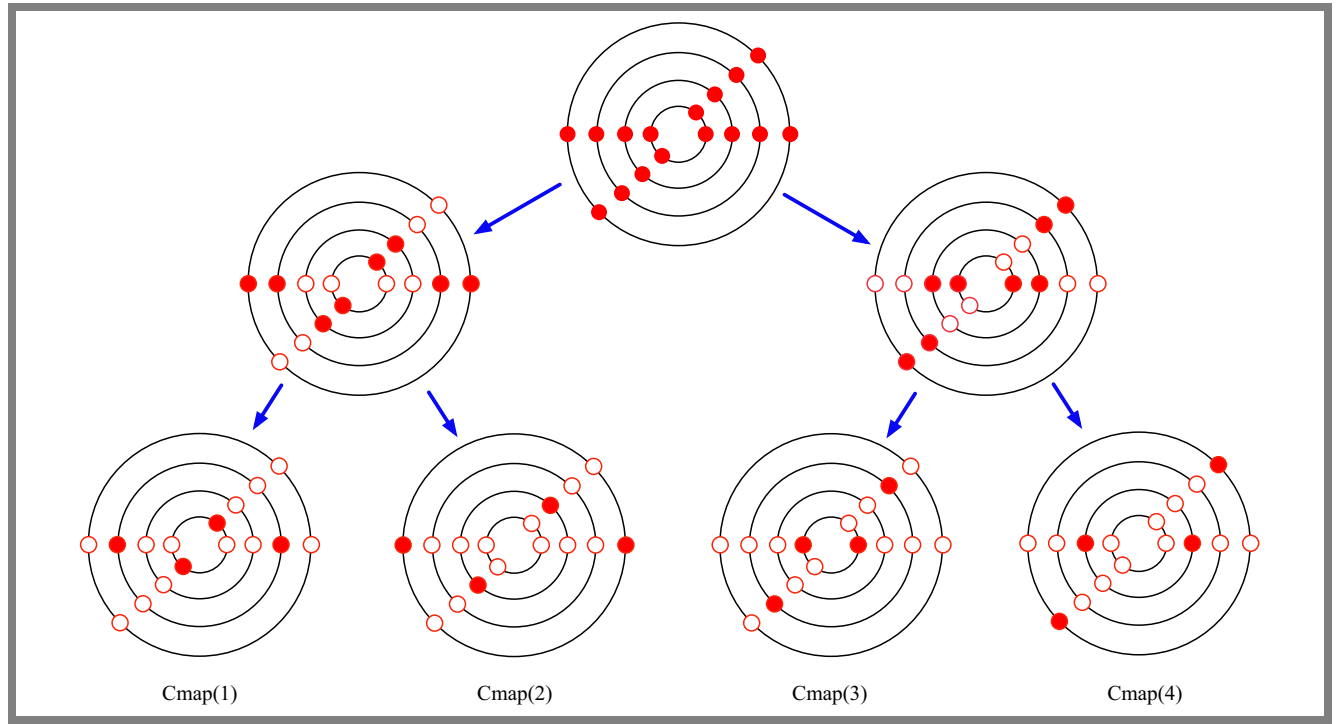


Fig. 3. The mother 16-QAM constellation is divided into four smaller constellations.

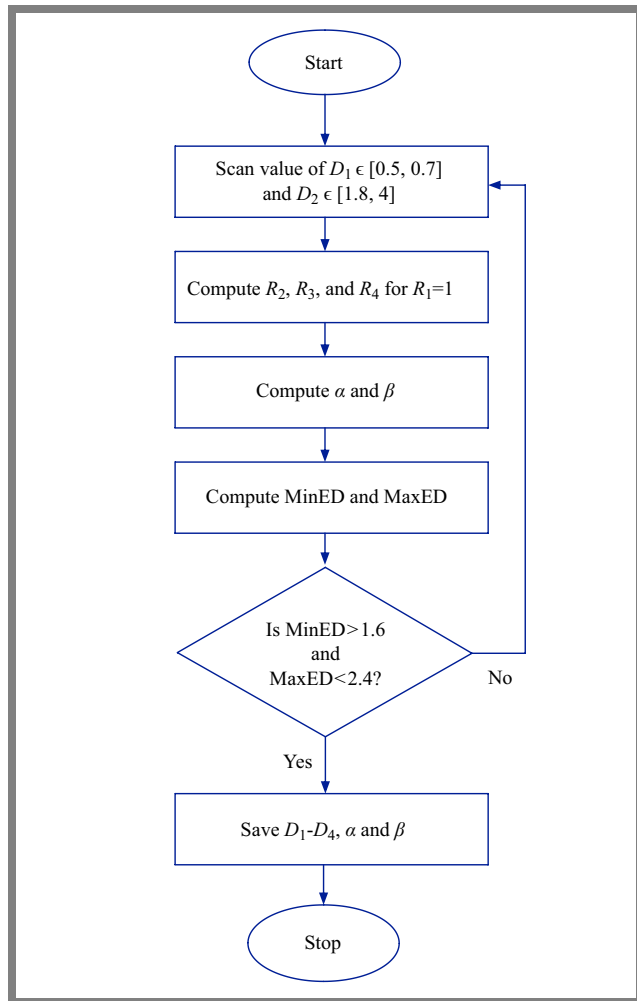


Fig. 4. Flowchart of the proposed CG-SCMA method.

called α . These ratios are then utilized to create the mother codebook [6], [7]. According to Eq. (4), $R_2 = D_1 + R_1$ and $R_3 = D_2 + R_1$, to keep the ratio of α and β constant $R_4 = \frac{R_3 R_2}{R_1}$. The next step is to search for $D_1 \in [0.5, 0.7]$ and $D_2 \in [1.8, 4]$ and then to determine α and β that achieved the minimum Euclidian distance (MinED) and the maximum Euclidian distance (MaxED) using the following formulas:

$$\begin{aligned} MinED = \frac{d_{\min}}{\sqrt{E}} &= \sqrt{\frac{(R_1 - R_1 \cos \emptyset)^2 + (R_1 \sin \emptyset)^2}{\frac{1}{4}(R_1^2 + R_2^2 + R_3^2 + R_4^2)}} \\ &= \sqrt{4 \frac{2 - 2 \cos \emptyset}{(1 + \alpha^2)(1 + \beta^2)}}, \end{aligned} \quad (5)$$

$$\begin{aligned} MaxED = \frac{d_{\max}}{\sqrt{E}} &= \sqrt{\frac{4R_4^2}{\frac{1}{4}(R_1^2 + R_2^2 + R_3^2 + R_4^2)}} \\ &= \sqrt{16 \frac{\beta^2}{1 + \beta^2} \cdot \frac{1}{1 + \frac{1}{\alpha^2}}}, \end{aligned} \quad (6)$$

where $\emptyset = \frac{\pi}{L \cdot dv}$ refers to the rotation angle. If MinED is greater than 1.58 and MaxED is lower than 2.4 [6], then these values are selected to generate the mother codebook, or are else used to select another value for D_1 and D_2 .

Using the obtained values, the following is a representation of the suggested mother codebook design relying on symmetric mapping [6], [7].

$$\text{Mother codebook} = \begin{bmatrix} a & b & -a & -b \\ c & d & -c & -d \\ e & f & -e & -f \\ g & h & -g & -h \end{bmatrix}. \quad (7)$$

where $a = R_3, b = R_1 \cos \emptyset, c = R_4, d = R_2 \cos \emptyset, e = R_1, f = R_3 \cos \emptyset, g = R_2$, and $h = R_4 \cos \emptyset$.

Partitioning the star 16-QAM constellation using trellis coded modulation subsets allows to obtain the users' codebooks, as shown in Fig. 3. User codebooks are generated by substituting Cmap(i) of the i -th subset of the mother codebook constellation for each 1 in the $\mathbf{F}$ matrix in Eq. (2), thus resulting in a new $\mathbf{F}'$ matrix, as presented in Eq. (8) [6], [7]:

$$\mathbf{F}' = \begin{bmatrix} \text{Cmap}(2) & \text{Cmap}(1) & \text{Cmap}(4) \\ \text{Cmap}(1) & 0 & 0 \\ 0 & \text{Cmap}(2) & 0 \\ 0 & 0 & \text{Cmap}(1) \\ 0 & 0 & 0 \\ \text{Cmap}(3) & \text{Cmap}(2) & 0 \\ \text{Cmap}(4) & 0 & \text{Cmap}(3) \\ 0 & \text{Cmap}(3) & \text{Cmap}(2) \end{bmatrix}. \quad (8)$$

Figure 4 shows the flow chart utilized to identify the ideal constellation points needed to create the CG-SCMA mother codebook [6], [7].

4. SCMA Encoder-Decoder Design

The method for encoding and multiplexing data for transmission and using a mother codebook to create user codebooks is illustrated in Fig. 5. A specific codebook is assigned to each user based on sparse matrix $\mathbf{V}$. When all six users provide data concurrently, information bits from user 1 through 6 are relayed in the order shown in Fig. 6 [11]. The user-specific codebook is added to the four time-frequency resource blocks currently in use. A single user owns each codebook, which is a 4×4 complex matrix. Each user's codebook contains four codes [11]. Users that need to send their encoded data should

construct codewords by multiplying their codebooks with encoded data. Finally, all users' codewords are multiplexed and sent to the destination [9], [14]. To guarantee good performance of the SCMA design, the potentially highest minimum Euclidian distance needs to be obtained. The general formula for estimating MED is:

$$MED = (\min \text{ or } \max) \|x - y\|, \quad \forall x, y \in \text{mother codebook and } x \neq y. \quad (9)$$

4.1. SCMA Decoder

The message passing algorithm (MPA) is an iterative parallel decoding method that works by moving extrinsic data from function nodes (FNs) to variables nodes (VNs) and vice versa [9], as shown in Fig. 1. Each FN of the factor graph computes its outgoing message to a particular VN in each iteration based on the incoming messages from the other VNs. Then, each VN responds with a different message in response to the FNs' reminder messages. The cycle is finally declared complete when one outgoing message has traveled over each edge, in both directions. The log likelihood rates (LLRs) of each coded bit are determined after many iterations to estimate the bits of each user. Algorithm 1 illustrates the three basic processes that make up the MPA [1] for an SCMA system of J users employing M constellation points through K orthogonal R resources.

Algorithm 1. Message passing algorithm

Input: $y, N_0, C_j, h_j, j = 1, \dots, J$

Result: Calculate each user's coded bit value for each user j , users are referred to as variable nodes (VNs), and resources or subcarriers are referred to as function nodes (FNs),

$U(k) = \{j, 1 \leq j \leq J \mid \text{VN}_j \text{ is connected to FN}_k\}$,

$R(j) = \{k, 1 \leq k \leq K \mid \text{FN}_k \text{ is connected to VN}_j\}$.

Step 1. Initialization.

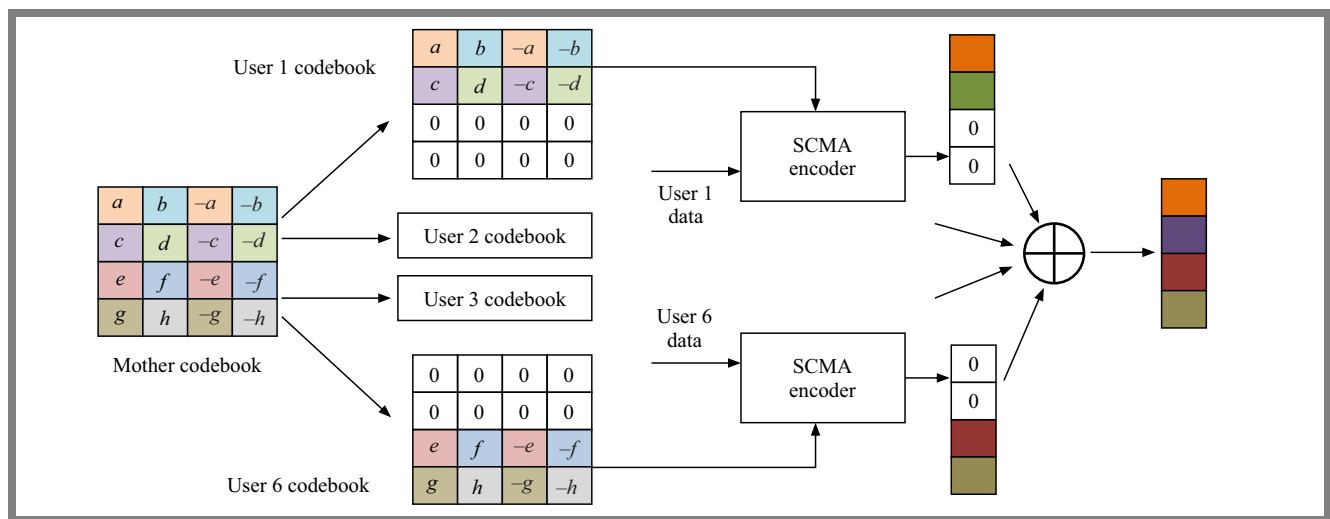


Fig. 5. SCMA encoder scheme.

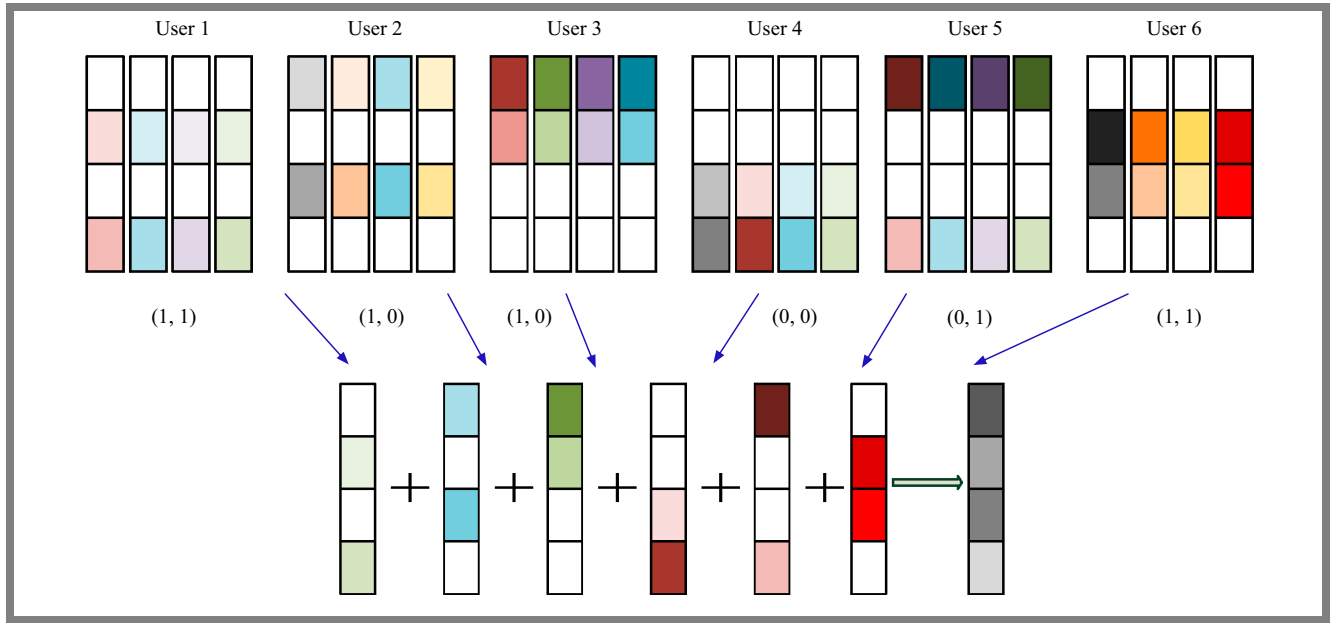


Fig. 6. SCMA encoding principle.

Each codeword's prior probability is given by:

$$V_{j \rightarrow k}^0(x_m) = P(x_m) = \frac{1}{M}, \quad j = 1, \dots, J, \quad k \in R(j).$$

Step 2. Iterative message passing along edges.

While N_{iter} **do**

FN update. The message from FN_k is to be passed to one of its neighbors VN_j , $k = 1, \dots, K$, $j \in U(k)$.

For a given codeword $x_m \in C_j$, $m = 1, \dots, MU_{j \rightarrow k}^t$ is:

$$U_{j \rightarrow k}^t = \left\{ \sum_{x_i | i \in U(k) \setminus j} e^{-\frac{1}{N_0}} \|y_k - \sum_j h_{j,k} x_{m,k}\|^2 \right\} \prod_{i \in U(k) \setminus j} U_{i \rightarrow k}^{t-1}(x_m).$$

VN update. The VN_j has a message that should be delivered to one of its neighbors FN_k , for a given codeword $m \in C_j$, $V_{j \rightarrow k}^t(x_m)$ is given by:

$$V_{j \rightarrow k}^t(x_m) = \frac{\prod_{i \in R(j) \setminus k} U_{i \rightarrow j}^{t-1}(x_m)}{\sum_{x_l \in C_j} \prod_{i \in R(j) \setminus k} U_{i \rightarrow j}^{t-1}(x_l)}.$$

To maintain the method numerically stable, normalization is required.

End while

Step 3. Decision on LLR at each VN.

For each layer j the *a posteriori* probability of codeword x_m is defined as:

$$P(x_m) = \prod_{k \in R(j)} U_{k \rightarrow j}^{N_{iter}}(x_m).$$

Each coded bit's log-likelihood rate bi , where $1 \leq i \leq \log_2(M)$ is represented by:

$$LLR(bi) = \log \frac{P(bi = 0)}{P(bi = 1)} = \log \frac{\sum_{\{x_m \in C_j | bi=0\}} P(x_m)}{\sum_{\{x_m \in C_j | bi=1\}} P(x_m)}.$$

The value of each bit's LLR is determined as follows:

$$bi = \begin{cases} 1 & \text{if } LLR(bi) \leq 0 \\ 0 & \text{otherwise} \end{cases}.$$

5. Simulation Results

The best values for the star 16-QAM constellation are determined by using dedicated software. The program outcome is that the best values of D_1 and D_2 produce the optimal ratio values $\alpha = 3.153$ and $\beta = 1.70$. These ratio values result in MinED = 1.60 and MaxED = 2.36 for the mother codebook. According to these ratio values, the four constellation point values are found to be $R_1 = 1.00$, $R_2 = 1.70$, $R_3 = 3.15$, and $R_4 = 5.36$, respectively. Based on these inputs, the mother codebook is:

$$\begin{bmatrix} 3.15 + i0.00 & 0.92 + i0.38 & -3.15 + i0.00 & -0.92 - i0.38 \\ 5.36 + i0.00 & 1.57 + i0.65 & -5.36 + i0.00 & -1.57 - i0.65 \\ 1.00 + i0.00 & 2.91 + i1.20 & -1.00 + i0.00 & -2.91 - i1.20 \\ 1.70 + i0.00 & 4.95 + i2.05 & -1.70 + i0.00 & -4.95 - i2.05 \end{bmatrix}$$

MinED is increased after the division of the main constellation into sub-constellations. After performing TCM, the four sub-constellations have achieved MinED = {3.85, 2.26, 2.26, 3.85} – see Fig. 7.

Two sub-constellations are assigned for each user to generate their codebooks. The improvement in MinED is higher than that of the sub-constellation proposed in [6]–[8]. Table 1 summarizes the obtained values of MinED, MaxED, D_1 , D_2 parameters for the proposed CG-SCMA and so-

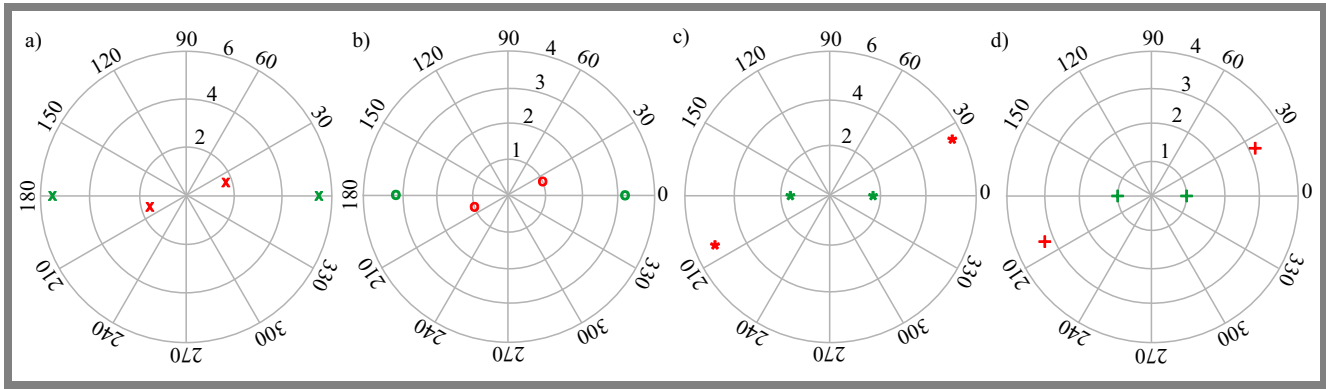


Fig. 7. Users' sub-constellations for: a) Cmap(1), b) Cmap(2), c) Cmap(3), and d) Cmap(4).

Tab. 1. Comparison of MinEd and MaxEd for the main constellation and sub-constellations.

	MinEd of MC	MaxEd of MC	D_1	D_2	MinEd for sub-constellation
Optimized SCMA [6]	1.58	2.40	0.58	2.00	3.35, 2.11, 3.35, 2.11
CG-SCMA [7]	1.58	2.33	0.60	2.05	3.46, 2.16, 2.16, 3.46
Proposed CG-SCMA	1.60	2.36	0.70	2.15	3.85, 2.26, 2.26, 3.85

lutions from [6], [7]. Furthermore, Fig. 8 presents the newly generated codebook, the CG-SCMA codebook and two other codebooks from [6], [7] in terms of BER. Figure 8 proves that the proposed method outperforms other known approaches.

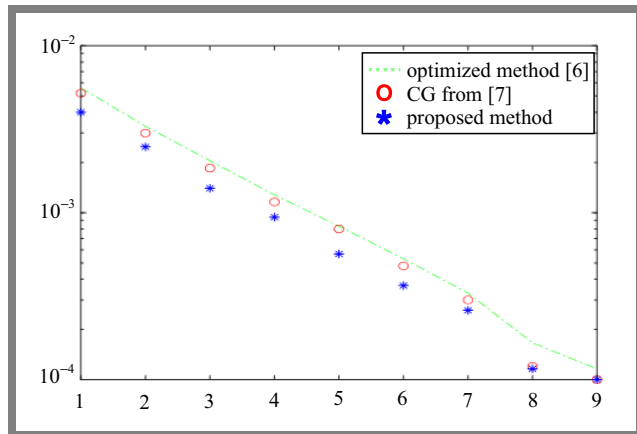


Fig. 8. BER comparison for classic and modified SCMA methods.

6. Conclusions

The paper presents an CG-SCMA codebook featuring maximization of MinEd and compares it with other existing codebooks described in the literature. Using the CG-SCMA codebook for encoding data from different users, the differences between the first and others radiuses are optimized over rotated constellations. To increase MinEd of a star 16-QAM, a novel CG-SCMA symmetric codebook architecture is proposed by employing TCM to subdivide the constellation. The SCMA codebook maximizes the space between star 16-QAM constellations, simultaneously decreasing BER and improving SNR.

References

- [1] M. Rebhi, K. Hassan, K. Raoof, and P. Charge, "Sparse code multiple access: Potentials and challenges", *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1205–1238, 2021 (<https://doi.org/10.1109/OJCOMS.2021.3081166>).
- [2] L. Zhang, "Sparse code multiple access for downlink multiple access of 5G wireless networks", *Computer Communications*, vol. 158, pp. 17–23, 2020 (<https://doi.org/10.1016/j.comcom.2020.04.022>).
- [3] S. Chaturvedi, Z. Liu, V.A. Bohara, A. Srivastava, and P. Xiao, "A tutorial on decoding techniques of sparse code multiple access", *IEEE Access*, vol. 10, pp. 58503–58524, 2022 (<https://doi.org/10.1109/ACCESS.2022.3178127>).
- [4] A. Kosasih, O. Setyawati, and R. Rahmadwati, "Low complexity multi-user MIMO detection for uplink SCMA system using expectation propagation algorithm", *TELKOMNIKA, Telecommunication Comput. Electron. Control*, vol. 16, no. 1, p. 182, 2018 (<https://doi.org/10.12928/telkomnika.v16.i1.7365>).
- [5] S. Lou, C. Gong, Q. Gao, and Z. Xu, "SCMA with low complexity symmetric codebook design for visible light communication", *2018 IEEE International Conference on Communications (ICC)*, USA, 2018, pp. 1–6 (<https://doi.org/10.1109/ICC.2018.8422652>).
- [6] S. Liu, J. Wang, J. Bao, and C. Liu, "Optimized SCMA codebook design by QAM constellation segmentation with maximized MED", *IEEE Access*, vol. 6, pp. 63232–63242, 2018 (<https://doi.org/10.1109/ACCESS.2018.2876030>).
- [7] Y.M. Tabra and B.M. Sabbar, "New computer generated-SCMA codebook with maximised Euclidian distance for 5G", *Iraqi Journal of Information and Communications Technology*, vol. 2, no. 2, pp. 9–24, 2019 (<https://doi.org/10.31987/ijict.2.2.64>).
- [8] Y.-M. Chen and J.-W. Chen, "On the design of near-optimal sparse code multiple access codebooks", *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2950–2962, 2020 (<https://doi.org/10.1109/TCOMM.2020.2974213>).
- [9] S.A. Hussain, N. Ahmad, I. Shaye, H.M. Kaidi, L.A. Latiff, N. Mohamed, and S.M. Sam, "A review of codebook design methods for sparse code multiple access", *Indonesian Journal of Electrical Engineering and Computer Science*, 2021 (<https://doi.org/22.927.10.11591/ijeecs.v22.i2.pp927-935>).
- [10] G. Zhang, Z. Gu, Q. Zhao, J. Ren, and W. Lu, "A threshold-based max-log-MPA low complexity multiuser detection algorithm", *Sensors*, vol. 20, no. 4, p. 1016, 2020 (<https://doi.org/10.3390/s20041016>).

- [11] P.P. Waghmare, "An optimized code book design and assignment based on 32-QAM constellation in downlink SCMA systems", *Digital Signal Processing*, vol. 109, p. 102919, 2021 (<https://doi.org/10.1016/j.dsp.2020.102919>).
 - [12] J. Lin, S. Feng, Y. Zhang, Z. Yang, and Y. Zhang, "A novel deep neural network based approach for sparse code multiple access", *Neurocomputing*, vol. 382, pp. 52–63, 2020 (<https://doi.org/10.48550/arXiv.1906.03169>).
 - [13] S.S. Mohamed and H.N. Abdullah, "Design low complexity SCMA codebook using Arnold's cat map", *Journal of Telecommunications and Information Technology*, no. 4, pp. 13–20, 2022 (<https://doi.org/10.26636/jtit.2022.164422>).
 - [14] Y.M. Tabra, "Beamforming with multiple access for 5G multimedia application", Ph.D. thesis, Al-Nahrain University, 2019 (<https://doi.org/10.13140/RG.2.2.13028.32647>).
-

Sura S. Mohamed, B.Sc.

E-mail: suhaibsur@gmail.com

Al-Nahrain University, College of Information Engineering,
Baghdad, Iraq

<https://coie-nahrain.edu.iq>

Hikmat N. Abdullah, Ph.D.

 <https://orcid.org/0000-0002-1133-2057>

E-mail: hikmat.abdullah@coie-nahrain.edu.iq

Al-Nahrain University, College of Information Engineering,
Baghdad, Iraq

<https://coie-nahrain.edu.iq>

A New Class of Fractional Cumulative Residual Entropy – Some Theoretical Results

Slimane Benmahmoud

Mohamed Boudiaf University at M'sila, Algeria

<https://doi.org/10.26636/jtit.2023.166622>

Abstract — In this paper, by differentiating the entropy's generating function (i.e., $h(t) = \int_{S_X} \bar{F}_X^t(x) dx$) using a Caputo fractional-order derivative, we derive a generalized non-logarithmic fractional cumulative residual entropy (FCRE). When the order of differentiation $\alpha \rightarrow 1$, the ordinary Rao CRE is recovered, which corresponds to the results from first-order ordinary differentiation. Some properties and examples of the proposed FCRE are also presented.

Keywords — cumulative residual entropy (CRE), entropy's generating function, fractional calculus, information measure, Riemann-Liouville/Caputo fractional integral/derivative, Tsallis/Rényi entropy.

1. Introduction

In keeping with the concept of entropy proposed by Clausius [1] in thermodynamics and Boltzmann [2] in classical statistical mechanics, Shannon proposed, in [3], a similar concept, in the context of communication theory, as a measure of surprise or uncertainty associated with the probability distributions of a random variable (RV). For a discrete RV X taking values in $\mathcal{X} = \{x_1, x_2, \dots, x_N\}$ and having a probability mass function $p_i = P(X = x_i)$ with $\sum_{i=1}^N p_i = 1$ and $p_i \geq 0$ for $i = 1, \dots, N$, it is given by:

$$S^{(S)}(X) = - \sum_{i=1}^N p_i \log p_i. \quad (1)$$

Due to its properties, which agree with the intuitive notions of randomness, such a measure of uncertainty suggested in Eq. (1) has proved to be useful in solving statistical problems related to the communication theory.

In the literature concerned with information theory and mathematical statistics, one can frequently encounter many contributions to generalize the concept of entropy in an attempt to explore and exploit the outcomes of its applications in different fields of engineering and physics. Perhaps the most popular among these contributions are the two developed by Tsallis [4] and Rényi [5], given by:

$$S_q^{(T)} = \frac{1}{q-1} \left(1 - \sum_i p_i^q \right), \quad q \in \mathbb{R}, \quad (2)$$

and

$$S_q^{(R)} = \frac{1}{1-q} \ln \left(\sum_i p_i^q \right), \quad q > 0, q \neq 1, \quad (3)$$

respectively.

Both entropies ($S_q^{(T)}$ and $S_q^{(R)}$) reduce to Shannon's entropy defined in Eq. (1), within the limit $q \rightarrow 1$. The Tsallis entropy $S_q^{(T)}$ has applications in Fokker-Plank systems [6] and diffusion equations [7]. Since the Rényi entropy $S_q^{(R)}$ has an inverse power law equilibrium distribution [8], it satisfies the zeroth law of thermodynamics [9]. Based on entropy $S_q^{(T)}$ defined in Eq. (2), Tsallis *et al.* have developed the theory of non-extensive statistical mechanics [4], where parameter q classifies a particular statistics, such that for $q \rightarrow 1$ it reproduces the standard Boltzmann-Gibbs statistics.

Abe in [10] has shown that while the Shannon entropy $S^{(S)}$ and the Tsallis entropy $S_q^{(T)}$ share the same generating function, $h(t) = \sum_i p_i^t$, the difference between their entropic forms is due to the differentiation operators considered. Namely, by applying the classical first-order differentiation operator, we can generate the following Shannon entropy $S^{(S)}$:

$$-\frac{d}{dt} h(t)|_{t=1} = - \sum_i p_i \log p_i \equiv S^{(S)}. \quad (4)$$

Otherwise, the Tsallis entropy $S_q^{(T)}$ can be generated by applying Jackson's q -differential operator given by:

$$({}^J D_q f)(x) = \frac{f(qx) - f(x)}{qx - x}, \quad (5)$$

instead of the first-order ordinary differentiation operator d/dt . This indicates that if one adopts other families of differentiation operators, new classes of entropies may be created.

Known as the differential entropy, the continuous analogue of the discrete entropy defined in Eq. (1), for a continuous RV, is given by [11]:

$$h(X) = - \int_{S_X} f_X(x) \log f_X(x), \quad (6)$$

where S_X and $f_X(x)$ are the support and probability density functions (PDF) of the RV X , respectively. However, this direct extension of the concept of entropy from discrete to

continuous schemes raises some concerns, including the following:

- it may be negative,
- it may become infinitely large,
- it does not necessarily remain invariant under the transformation of the coordinate systems,
- it is so difficult to be estimated through empirical samples.

For further discussions on these and more concerns, please refer to [12].

Among the proposed solutions intended to overcome these limitations, is the one named cumulative residual entropy (CRE) [13], [14]. It is given by:

$${}^{\text{CRE}}h(X) = - \int_0^{+\infty} \bar{F}_X(x) \log \bar{F}_X(x), \quad (7)$$

where $\bar{F}_X(x)$ is the complementary cumulative distribution function (CCDF) of the RV X . CRE is a non-negative quantity. It is well-defined for both discrete and continuous RVs and can be easily estimated using empirical entropy samples [13]. Some extensions and generalizations for the CRE have been proposed in [15]–[20]. Other CDF or CCDF-based information measures including the cumulative Shannon entropy [21]–[23], cumulative residual Tsallis/Rényi's entropy [24]–[26] and cumulative residual K-L divergence [27], [28], were proposed.

By benefiting from Abe's idea [10], Ubriaco in [29] seems to be the first who has proposed an FC-based entropy given by:

$$H_q^U(X) = \sum_{i=1} p_i (-\log p_i)^q, \quad 0 \leq q \leq 1. \quad (8)$$

In the limit $q \rightarrow 1$, we get Shannon's entropy $S^{(S)}(X)$ defined in Eq. (1).

Following Ubriaco's approach, Mitsuhiko in [30] proposed a new class of fractional entropy.

In this paper, inspired by Mitsuhiko's work in [30], we propose a new class of FCRE by relying on the CCDF-based entropy's generating function $h(t) := \int_{S_X} \bar{F}_X^t(x) dx$ and using the theory of fractal calculus (FC). The use of FC allows us to measure the information in a generalized metric space. The main area in which FCRE may be used directly is information theory. It can also be applied in computer vision solutions, where information-related measures, and in particular those related to CE, are of great importance. Among the recent works on FCRE are [31] and [32].

The remainder of this paper is organized as follows. In Section 2, from the theory of FC, we recall some definitions concerning the RL fractional integral and the Caputo fractional derivative. The generalized FCRE derived from a Caputo fractional-order differentiation of the aforementioned entropy's generating function with order α along with some of its properties and examples are then presented in Section 3. Finally, Section 4 concludes the paper.

2. Review of Fractional Integrals and Derivatives

Fractional calculus (FC) is a mathematical analysis branch which studies different possible approaches to define fractional-order integrals and derivatives. Based on FC, the theory of classical integer-order differential equations has been then generalized to the broader theory of fractional-order differential equations. FC can be traced back to a letter written to l'Hopital by Leibniz in 1695 [33]. In 1832, Liouville carried out a heavy-handed investigation on FC [34]. After that, the Riemann-Liouville (RL) fractional integrodifferential operator was introduced by Riemann in [35], along with a comprehensive theory of FC. FC has led to many breakthroughs in different fields of physics and engineering, where various processes can be modeled in a more accurate and authentic way [36].

The left-sided Riemann-Liouville (RL) fractional integral ${}^{\text{RL}}I_{a+}^\alpha f$ of order $\alpha \in \mathbb{R}(\alpha > 0)$ of an integrable function $f : [a, b] \rightarrow \mathbb{R}$, ($0 \leq a < b \leq \infty$) is defined as [37]:

$$\begin{aligned} ({}^{\text{RL}}I_{a+}^\alpha f)(t) &= {}^{\text{RL}}I_{a+}^\alpha [f(x)](t) \\ &= \frac{1}{\Gamma(\alpha)} \int_a^t (t-x)^{\alpha-1} f(x) dx, \end{aligned} \quad (9)$$

with $a \in \mathbb{R}$, $t > a$, $\alpha > 0$. $\Gamma(\cdot)$ is Euler's gamma function defined as $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$, $\alpha \in \mathbb{R}(\alpha > 0)$.

The left-sided Caputo fractional derivative ${}^{\text{C}}D_{a+}^\alpha f$ of order $\alpha \in \mathbb{R}(\alpha > 0)$ of an integrable and differentiable function $f : [a, b] \rightarrow \mathbb{R}$, ($0 \leq a < b \leq \infty$) is defined as [37]:

$$\begin{aligned} ({}^{\text{C}}D_{a+}^\alpha f)(t) &= {}^{\text{C}}D_{a+}^\alpha [f(x)](t) \\ &= \left({}^{\text{RL}}I_{a+}^{n-\alpha} \left(\frac{d}{dt} \right)^n f \right)(t), \end{aligned} \quad (10)$$

with $a \in \mathbb{R}$, $t > a$, $\alpha > 0$, $n = [\alpha] + 1$.

When $0 < \alpha < 1$, we get:

$$({}^{\text{C}}D_{a+}^\alpha f)(t) = \left({}^{\text{RL}}I_{a+}^{1-\alpha} \frac{d}{dt} f \right)(t), \quad (a \in \mathbb{R}, t > a). \quad (11)$$

3. Caputo Fractional Derivative-based FCRE

Definition 1 – CRE [13]. The FCRE of a non-negative continuous RV X with a PDF $f_X(x)$, a CCDF $\bar{F}_X(x)$ and support $S_X = [0, +\infty]$ is defined as:

$$\begin{aligned} {}^{\text{CRE}}h(X) &:= - \int_{S_X} \bar{F}_X(x) \log \bar{F}_X(x) dx \\ &= E_X \left[- \frac{\bar{F}_X(X) \log \bar{F}_X(X)}{f_X(X)} \right], \end{aligned} \quad (12)$$

when the integral exists.

$E_X[X]$ is the expected value of the RV X .

Remark. Throughout the entire paper, the base of the logarithm will be set to Euler's number $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Our basic idea consists of rewriting Eq. (12) as follows:

$${}^{CRE}h(X) := -\lim_{t \rightarrow 1} \frac{d}{dt} \int_{S_X} \bar{F}_X^t(x) dx. \quad (13)$$

Then, we deform the ordinary differential operator $\frac{d}{dt}$ in Eq. (13) to the Caputo fractional differential operator ${}^C D_{a+}^\alpha$ defined in Eq. (11) (which reduces to $\frac{d}{dt}$ in the limit $\alpha \rightarrow 1$).

Definition 2 – FCRE. The FCRE of a non-negative continuous RV X with a PDF $f_X(x)$, a CCDF $\bar{F}_X(x)$ and support $S_X = [0, +\infty]$ is defined as:

$${}^{FCRE}h^\alpha(X) := -\lim_{t \rightarrow 1} {}^C D_{a+}^\alpha \int_{S_X} \bar{F}_X^t(x) dx. \quad (14)$$

Based on these ideas, we derive in the following theorem a new class of FCRE.

3.1. Fractional CRE

Theorem 1 – FCRE. The finite-value FCRE of order α of a non-negative continuous RV X with a CCDF $\bar{F}_X(x)$ and support $S_X \subseteq [0, +\infty]$ is defined as:

$${}^{FCRE}h^\alpha(X) := \int_{S_X} \frac{(\alpha-1)}{\Gamma(2-\alpha)} {}_1F_1(1; 1-\alpha; \log \bar{F}_X(x)) dx, \quad 0 < \alpha \leq 1, \quad (15)$$

where ${}_1F_1(p; q; x)$ is the confluent hypergeometric function of the first kind – see Eq. (13.2.2) in [38].

Proof. Using the operator defined in Eq. (11), where the lower limit of the RL-integral is taken to zero, i.e. $a = 0$, without loss of generality, Eq. (13) can be re-written as follows:

$$- \lim_{t \rightarrow 1} \frac{d}{dt} \left({}^{RL} I_0^{1-\alpha} \left[\int_{S_X} e^{t \log \bar{F}_X(x)} dx \right] (t) \right), \quad (0 < \alpha \leq 1).$$

Therefore, we need to solve the following integral:

$${}^{FCRE}h^\alpha(X) := \frac{-1}{\Gamma(1-\alpha)} \lim_{t \rightarrow 1} \frac{d}{dt} \int_{S_X} \left(\int_0^t (t-y)^{-\alpha} e^{y \log \bar{F}_X(x)} dy \right) dx,$$

By letting $t-y = \frac{z}{\log \bar{F}_X(x)}$ and using the definition of the incomplete gamma function of the first kind, i.e. $\gamma(\mu, x) = \int_0^x e^{-t} t^{\mu-1} dt$, we get:

$${}^{FCRE}h^\alpha(X) := \frac{-1}{\Gamma(1-\alpha)} \times \lim_{t \rightarrow 1} \frac{d}{dt} \int_{S_X} \left[(\log \bar{F}_X(x))^{\alpha-1} e^{t \log \bar{F}_X(x)} \gamma(1-\alpha, t \log \bar{F}_X(x)) \right] dx.$$

Now, we use the following relation:

$$\gamma(\mu, x) = \mu^{-1} x^\mu e^{-x} {}_1F_1(1; 1+\mu; x).$$

Then, we refer to Eq. (13.3.18) in [38] to take the ordinary first derivative and we set $t = 1$ to get the results in Eq. (15). ■

The standard CRE defined in Eq. (12) can be recovered from the FCRE in Eq. (15), for $\alpha \rightarrow 1$, with the help of the limit:

$$\lim_{b \rightarrow 0} {}_1F_1(1; b; x) = x e^x.$$

The expression in Eq. (15) allows us to calculate the FCRE ($h^\alpha(X)$) for different values of α and compare them to each other and to the conventional CRE ($\alpha = 1$), simultaneously get further insights on its behavior.

3.2. Properties of FCRE

The FCRE defined in Eq. (15) is a generalized entropy. In the following section, we consider some of its mathematical properties.

3.2.1. Positivity of the FCRE

Let us consider the integrand of Eq. (15):

$$h^\alpha = \frac{(\alpha-1)}{\Gamma(2-\alpha)} {}_1F_1(1; 1-\alpha; \log \bar{F}_X(x)), \quad 0 < \alpha \leq 1. \quad (16)$$

With the help of the asymptotic expansion of ${}_1F_1(\cdot; \cdot; \cdot)$ for a large argument (see Sect. 13.7 in [38]) for the case where $\bar{F}_X(x) = 0$, we can easily find that $h^\alpha = 0$. For $\bar{F}_X(x)^\alpha = 1$, however, we get:

$$h^\alpha = \frac{(\alpha-1)}{\Gamma(2-\alpha)}.$$

In this case, h^α is negative for $\alpha \in (0, 1)$ and it is equal to zero for $\alpha \rightarrow 1$. These results imply that, as a unique property, the FCRE introduced in Eq. (15) is not positive unless $\alpha \rightarrow 1$.

3.2.2. Concavity of FCRE

The second ordinary derivative with respect to $\bar{F}_X(x)$ of the integrand given in Eq. (16) is:

$$\frac{\partial^2 h^\alpha}{\partial \bar{F}_X(x)^2} = \frac{-\alpha}{\Gamma(3-\alpha) \bar{F}_X(x)} {}_1F_1(1-\alpha; 3-\alpha; -\log \bar{F}_X(x)), \quad (17)$$

which is less than zero for $\alpha \in (0, 1)$, $0 < \bar{F}_X(x) \leq 1$. This means that the FCRE defined in Eq. (15) is concave.

3.2.3. Homogeneity and shift-independence of FCRE

Let X be an RV. If $Y = aX + b$, with $a > 0$ and $b \geq 0$, then:

$${}^{FCRE}h^\alpha(Y) = a {}^{FCRE}h^\alpha(X). \quad (18)$$

Proof. Recall that $\bar{F}_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right)$, $x \in [0, +\infty]$. From Eq. (15), the equality in Eq. (18) is obtained. ■

The result of Eq. (18) shows the effect of linear transformation of the RV on the FCRE. It demonstrates that the FCRE is homogeneous and shift-independent.

3.3. Examples on the FCRE

In the following section, we give a few examples of the FCRE for some common continuous probability distributions.

Example 1. Let X be a uniformly distributed RV in $[0, a]$, $a > 0$, i.e. $X \sim \text{Uniform}(0, a)$. Then, its FCRE is given by:

$${}^{FCRE}h^\alpha(X) = \frac{a\alpha(1-\alpha)}{2^{\alpha+1}\Gamma(2-\alpha)} B_{\frac{1}{2}}(\alpha; 1), \quad (19)$$

where $B(\cdot; \cdot)$ is the incomplete beta function (see Eq. (8.17.1) in [38]). To get further insights on the behavior of ${}^{FCRE}h^\alpha(X)$

as a function of α , one may refer to Fig. 1 where Eq. (19) is plotted.

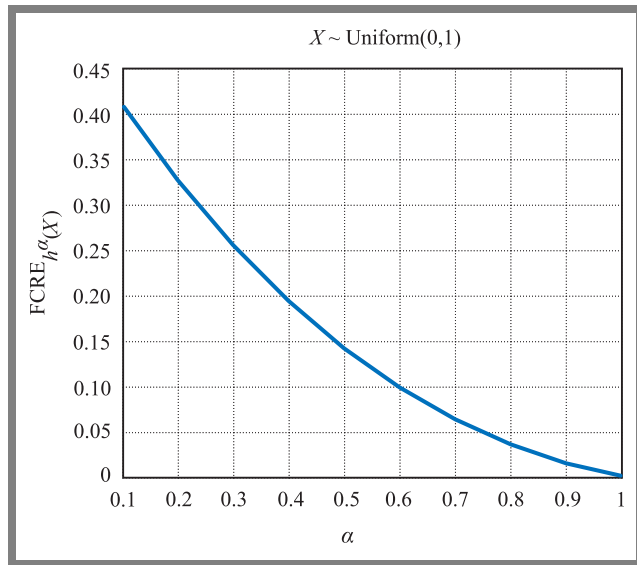


Fig. 1. The ${}^{\text{FCRE}}h^\alpha(X)$ as a function of α of a uniformly distributed RV in $[0, a]$, with $a = 1$.

Example 2. Let X be an exponentially distributed RV with a rate parameter λ , i.e. $X \sim e^\lambda$. Then, its FCRE is given by:

$${}^{\text{FCRE}}h^\alpha(X) = \frac{(\alpha - 1)}{\lambda\Gamma(2 - \alpha)} \int_0^{\infty} {}_1F_1(1; 1 - \alpha; -x) dx, \quad (20)$$

where $\Gamma(\cdot)$ is the Euler's gamma function (see Eq. (5.2.1) in [38]).

Example 3. Let X be a normally distributed RV with a mean μ and a standard deviation σ , i.e. $X \sim \mathcal{N}(\mu, \sigma)$. Then, its FCRE is given by:

$${}^{\text{FCRE}}h^\alpha(X) = \frac{(\alpha - 1)}{\Gamma(2 - \alpha)} \int_0^{\infty} {}_1F_1\left[1; 1 - \alpha; \log\left(\operatorname{erfc}\left(\frac{x - \mu}{\sigma}\right)\right)\right] dx, \quad (21)$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function (see Eq. (7.2.2) in [38]).

4. Conclusion

In this paper, we have introduced a new entropy functional ${}^{\text{FCRE}}h^\alpha(X)$ which generalizes the conventional CRE, ${}^{\text{CRE}}h^\alpha(X)$, by applying the Caputo fractional-order derivative. The FCRE has inherited the merits of the conventional CRE. We have verified that this entropy has useful features. Other properties, e.g. stability, are worth exploring separately in the future. The possibility of estimating it from empirical random samples merits attention. Testing its reliability through simulations on the logistic map and its applications in computer vision-related fields are also activities that are worthy of consideration.

References

- [1] R. Clausius, *The Mechanical Theory of Heat: With its Applications to the Steam-Engine and to the Physical Properties of Bodies*. London: J. Van Voorst, 1867 (ISBN: 9780343987374).
- [2] L. Boltzmann, *Vorlesungen Ueber Die Principe Der Mechanik*. Leipzig: Verlag von Johann Ambrosius Barth, vol. 1, 1897 (ISBN: 9781375769624).
- [3] C.E. Shannon, "A mathematical theory of communication", *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948 (<https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>).
- [4] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics", *Journal of Statistical Physics*, vol. 52, pp. 479–487, 1988 (<https://doi.org/10.1007/BF01016429>).
- [5] A. Rényi, "On measures of information and entropy", *Proceedings of the 4th Berkeley Symposium on Mathematics on Mathematical Statistics and Probability*, vol. 4.1, pp. 547–561, 1960.
- [6] T.D. Frank and A. Daffertshofer, "H-theorem for nonlinear Fokker-Planck equations related to generalized thermostatics", *Physica A: Statistical Mechanics and its Applications*, vol. 295, no. 3–4, pp. 455–474, 2001 ([https://doi.org/10.1016/S0378-4371\(01\)00146-7](https://doi.org/10.1016/S0378-4371(01)00146-7)).
- [7] A.R. Plastino, M. Casas, and A.R. Plastino, "A nonextensive maximum entropy approach to a family of nonlinear reaction-diffusion equations", *Physica A: Statistical Mechanics and its Applications*, vol. 280, no. 3–4, pp. 289–303, 2000 ([https://doi.org/10.1016/S0378-4371\(00\)00006-6](https://doi.org/10.1016/S0378-4371(00)00006-6)).
- [8] E.K. Lenzi, R.S. Mendes, and L.R. da Silva, "Statistical mechanics based on Rényi entropy", *Physica A: Statistical Mechanics and its Applications*, vol. 280, no. 3, pp. 337–345, 2000 ([https://doi.org/10.1016/S0378-4371\(00\)00007-8](https://doi.org/10.1016/S0378-4371(00)00007-8)).
- [9] A.S. Parvan and T.S. Biró, "Extensive Rényi statistics from non-extensive entropy", *Physics Letters A*, vol. 340, no. 5–6, pp. 375–387, 2005 (<https://doi.org/10.1016/j.physleta.2005.04.036>).
- [10] S. Abe, "A note on the q -deformation-theoretic aspect of the generalized entropies in nonextensive physics", *Physics Letters A*, vol. 224, no. 6, pp. 326–330, 1997 ([https://doi.org/10.1016/S0375-9601\(96\)00832-8](https://doi.org/10.1016/S0375-9601(96)00832-8)).
- [11] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley Series in Telecommunications and Signal Processing). Wiley, 1991 (ISBN: 9780471062592).
- [12] G. Jumarie, *Relative Information: Theories and Applications*. Berlin and Heidelberg: Springer, 2011 (ISBN: 9783642840197).
- [13] M. Rao, Y. Chen, B.C. Vemuri, and F. Wang, "Cumulative residual entropy: A new measure of information", *IEEE Transactions on Information Theory*, vol. 50, pp. 1220–1228, 2004 (<https://doi.org/10.1109/TIT.2004.828057>).
- [14] F. Wang, B.C. Vemuri, M. Rao, and Y. Chen, "Cumulative residual entropy, a new measure of information & its application to image alignment", *Proceedings 9th IEEE International Conference on Computer Vision*, vol. 1, pp. 548–553, 2003 (<https://doi.org/10.1109/ICCV.2003.1238395>).
- [15] S. Baratpour and Z. Bami, "On the discrete cumulative residual entropy", *Journal of the Iranian Statistical Society*, vol. 11, no. 2, pp. 203–215, 2012 (<http://jirss.irstat.ir/article-1-193-en.html>).
- [16] J. Navarro, Y. del Águila, and M. Asadi, "Some new results on the cumulative residual entropy", *Journal of Statistical Planning and Inference*, vol. 140, no. 1, pp. 310–322, 2010 (<https://doi.org/10.1016/j.jspi.2009.07.015>).
- [17] G. Psarrakos and J. Navarro, "Generalized cumulative residual entropy and record values", *Metrika*, vol. 76, no. 5, pp. 623–640, 2013 (<https://doi.org/10.1007/s00184-012-0408-6>).
- [18] G. Rajesh, E.I. Abdul-Sathar, K.R.M. Nair, and K.V. Reshmi, "Bivariate extension of dynamic cumulative residual entropy", *Statistical Methodology*, vol. 16, pp. 72–82, 2014 (<https://doi.org/10.1016/j.stamet.2013.07.006>).
- [19] M. Asadi and Y. Zohrevand, "On the dynamic cumulative residual entropy", *Journal of Statistical Planning and Inference*, vol. 137, no. 6, pp. 1931–1941, 2007 (<https://doi.org/10.1016/j.jspi.2006.06.035>).
- [20] S. Park and I. Kim, "On cumulative residual entropy of order statis-

- tics”, *Statistics & Probability Letters*, vol. 94, pp. 170–175, 2014 (<https://doi.org/10.1016/j.spl.2014.07.020>).
- [21] A. Di Crescenzo and M. Longobardi, “On cumulative entropies and lifetime estimations”, in *Methods and Models in Artificial and Natural Computation*, J. Mira, J.M. Ferrandez, J.R., Alvarez, F. de la Paz, and F.J. Toledo, Eds. A Homage to Professor Mira’s Scientific Legacy. IWINAC 2009. LNCS, vol 5601. Berlin, Heidelberg: Springer, 2009 (https://doi.org/10.1007/978-3-642-02264-7_15).
- [22] S. Kayal, “On generalized cumulative entropies”, *Probability in the Engineering and Informational Sciences*, vol. 30, pp. 640–662, 2016 (<https://doi.org/10.1017/S0269964816000218>).
- [23] A. Di Crescenzo and A. Toomaj, “Further results on the generalized cumulative entropy”, *Kybernetika*, vol. 53, no. 5, pp. 959–982, 2017 (<https://doi.org/10.14736/kyb-2017-5-0959>).
- [24] S. M. Sunoj and M. Linu, “Dynamic cumulative residual Rényi’s entropy”, *A Journal of Theoretical and Applied Statistics*, vol. 46, pp. 41–56, 2012 (<https://doi.org/10.1080/02331888.2010.494730>).
- [25] G. Rajesh and S. M. Sunoj, “Some properties of cumulative Tsallis entropy of order α ”, *Statistical Papers*, vol. 60, pp. 933–943, 2019 (<https://doi.org/10.1007/s00362-016-0855-7>).
- [26] M.M. Sati and N. Gupta, “Some characterization results on dynamic cumulative residual Tsallis entropy”, *Journal of Probability and Statistics*, vol. 2015, Article ID 694203, 2015 (<https://doi.org/10.1155/2015/694203>) 1–8.
- [27] S. Park, M. Rao, and D.W. Shin, “On cumulative residual Kullback-Leibler information”, *Statistics & Probability Letters*, vol. 82, no. 11, pp. 2025–2032, 2012 (<https://doi.org/10.1016/j.spl.2012.06.015>).
- [28] S. Park, H.A. Noughabi, and I. Kim, “General cumulative Kullback-Leibler information”, *Communications in Statistics – Theory and Methods*, vol. 47, no. 7, pp. 1551–1560, 2018 (<https://doi.org/10.1080/03610926.2017.1321767>).
- [29] M.R. Ubriaco, “Entropies based on fractional calculus”, *Physics Letters A*, vol. 373, no. 30, pp. 2516–2519, 2009 (<https://doi.org/10.1016/j.physleta.2009.05.026>).
- [30] M. Akimoto and A. Suzuki, “Proposition of a new class of entropy”, *Journal of the Korean Physical Society*, vol. 38, no. 5, pp. 460–463, 2001 (<https://www.jkps.or.kr/journal/download.php?spage=460&volume=38&number=5>).
- [31] H. Xiong, P. Shang, and Y. Zhang, “Fractional cumulative residual entropy”, *Communications in Nonlinear Science and Numerical Simulation*, vol. 78, 104879, 2019 (<https://doi.org/10.1016/j.cnsns.2019.104879>).
- [32] A. Di Crescenzo, S. Kayal, and A. Meoli, “Fractional generalized cumulative entropy and its dynamic version”, *Communications in Nonlinear Science and Numerical Simulation*, vol. 102, 105899, 2021 (<https://doi.org/10.1016/j.cnsns.2021.105899>).
- [33] G.W. Leibniz, “Letter from Hanover; Germany; to G.F.A. L’Hopital”, Sept. 30, 1695, in *Mathematische Schriften*, 1849; reprinted 1962, Olms Verlag, Hidesheim, Germany, vol. 2, pp. 301–302, 1965.
- [34] J. Liouville, *Mémoire sur le calcul des différentielles à indices quelconques*. 1832 (<https://books.google.dz/books?id=6jfbTwwAACAAJ>).
- [35] B. Riemann, R. Dedekind, and H.M. Weber, “Versuch einer allgemeinen Auffassung der Integration und Differentiation”, 1847, in *Bernard Riemann’s gesammelte mathematische Werke und wissenschaftlicher Nachlass*, R. Dedekind and H. Weber, Eds. (Cambridge Library Collection – Mathematics, pp. 331–344). Cambridge: Cambridge University Press, 2013 (<https://doi.org/10.1017/CB09781139568050.020>).
- [36] R. Hilfer, *Applications of Fractional Calculus in Physics*. World Scientific, Singapore, 2000 (<https://doi.org/10.1142/3779>).
- [37] A.A. Kilbas, H. M. Srivastava, and J.J. Trujillo, *Theory and applications of Fractional Differential Equations*. Elsevier Science, 2006 (ISBN: 9780444518323).
- [38] F.W.J. Olver, D.W. Lozier, R.F. Boisvert, and C.W. Clark, *NIST Handbook of Mathematical Functions*. New York: Cambridge University Press, 2010 (ISBN: 9780521192255).

Slimane Benmahmoud

A researcher with Signals and Systems Laboratory (LASS),
Department of Electronics Engineering
E-mail: slimane34a@gmail.com
Mohamed Boudiaf University at M’sila, Algeria
<https://www.univ-msila.dz/en/>

Concept of Joint Functioning of Access Control Systems

Mykhailo Strelbitskyi, Valentyn Mazur, Evgenii Prokopenko, Roman Rachok, and Dmytro Mul

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

<https://doi.org/10.26636/jtit.2023.164322>

Abstract — Modern integrated information and telecommunication systems are upgraded on a continuous basis. Such systems contain both new and old components. The approaches to developing individual components of access control systems are different in the majority of cases. As a rule, modernization of outdated but efficient systems that have been operating without any failures for long periods of time is economically unfeasible. Such an approach requires that different subsystems function based on shared data. This necessitates the coordination of various access control systems in order to ensure proper information security levels. This article examines how joint functioning of various versions of access control systems deployed in IT and telecommunication spheres may be achieved at the stage of their modernization. Potential ways in which information flows may bypass the security policies of one of the access control systems concerned are determined. The authors discuss traditional access control models. For role-based and thematic access control models, specific hypotheses are formulated to comply with security policies when different versions of access control systems work together. The structure of the model assuming that different versions of access control systems operate jointly has been developed. Based on the model, the necessary and sufficient conditions are determined under which unauthorized information flows are prevented. The security theorem for the joint functioning of different versions of access control systems is presented and proved. The results of the study showed that the methodological basis for coordinating access control models applicable to information and telecommunication systems undergoing modernization consists in observing, separately, the equality of information flows between shared objects in each of the versions of the access control systems. The approaches developed in this article can be extended to combined access control systems.

Keywords — access control model, concept, information and telecommunication system, modernization.

1. Introduction

Modern information and telecommunication systems (ITS) have a large number of subsystems (components) that are distributed across various geographical locations. A special feature of such systems is the requirement to operate in real time, and even short-lasting periods of unavailability or shut-downs may lead to serious damage on a nationwide scale. One of the ways in which reliability of such systems may be ensured is to guarantee the security of information flows. Information security can be described using such properties as: confidentiality, integrity, and availability [1], [2]. Compliance with these properties is ensured in ITS solutions,

in particular, by the access control system allowing users to access or prohibiting them from accessing a given organization's resources [3]. Many studies focused on analyzing existing access control systems. The traditional access control models are compared in Table 1 [4].

Tab. 1. Comparison of traditional access control models.

Criteria	DAC	MAC	RBAC	ABAC
Principle of least privilege	—	—	+	+
Dynamic behavior	—	—	—	+
Safety of models	—	+	+	+
Separation of duties	—	—	+	+
Capability delegation	+	—	—	—
Configuration flexibility	+	—	+	—
Auditing	+	+	+	+

where: DAC – discretionary access control, MAC – mandatory access control, RBAC – role-based access control, and ABAC – attribute-based access control.

Other authors considered theoretical aspects of the process of establishing information security models [4], [5] and examined the formation of combined models based on the traditional components [4]–[7]. However, all authors take into consideration one paradigm of the access control system modeling process, namely whether the system is of the traditional or combined variety (Fig. 1).

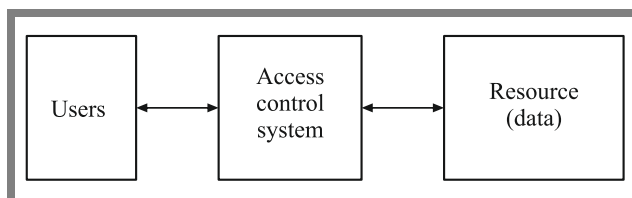


Fig. 1. Typical application of access control systems.

Some of the ITS that are in use currently were developed a long time ago. The integrated ITS of the State Border Guard Service of Ukraine may serve as a good example here. It consists of numerous subsystems [8] and was created by integrating individual ITS data sets into a higher-level data structure. The adoption of such an approach was conditioned

by completely different requirements that needed to be satisfied (Fig. 2).

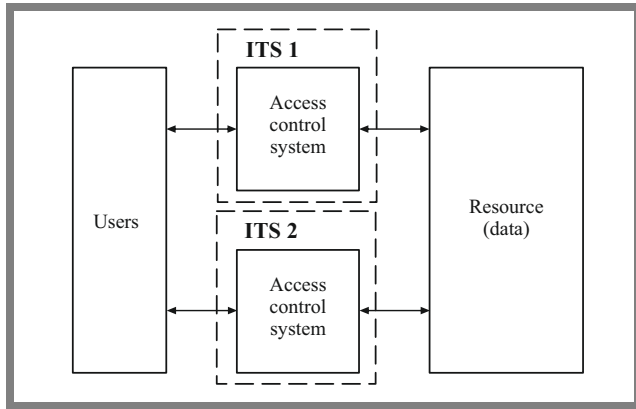


Fig. 2. Typical application of an access control system.

The development of access control systems (ACS) operating within various ITS that interact with one another as part of the same supersystem can be carried out both according to the same and different models of access differentiation. While upgrading the solution, the problem of coordinating various components needs to be taken into consideration, thus requiring that a concept ensuring their joint operation be designed. Unfortunately, information security models do not provide for joint functioning with other similar models in a common data field, and such a scenario may be encountered while performing modernization work. All models cover the functioning of a single system (security monitor, security core, etc.) and ensure compliance with its security policy. Recently, separate methods for coordinating access differentiation have been developed (as described in [9]–[11]), but a general approach has not been formed yet.

2. Aim of the Article

This paper analyzes the joint functioning of existing information security and focuses on the stage of modernization of specific information and telecommunication systems in order to:

- determine potential information flows bypassing one of the security models,
- determine hypothetical conditions under which their joint functioning is possible,
- define a concept for their safe joint functioning.

3. Joint Operation of Access Control Models

3.1. General Theoretical Provisions

It is not possible to define an ITS security policy which ensures the reliability of the information flow without relying on information security models serving as a formalized description of the basic principles of the solution to be adopted. Only with the help of formal models, the system's security

may be proven based on mathematical conditions. Security models define the basic principles of security policies, are used in their design and in the creation of technological solutions offering suitable levels of information security. In the context of ITS upgrades, if information security functionalities are implemented, in advance, in both the old and new versions of the software in order to eliminate conflicts caused by differences in models, it is necessary to perform their verification. So far, a number of strategies for modeling the behavior of security systems have been proposed in the literature [12]–[16]. Furthermore, each known model operates based on three elements: subject, object, and operations the former may perform on the latter (reading, writing, deletion, etc.). To represent an ACS, three sets of elements need to be defined:

- objects $\{O\}$ – terminals, network nodes, communication channels, external devices, applications, volumes, directories, files, records, recording fields,
- subjects $\{S\}$ – users or processes running on their behalf, i.e. administrators, applications, processes, terminals,
- operations $\{\Omega\}$ – sequences of the subject's actions concerning the object (information flow).

Based on such an assumption, an ACS is an object of set-theoretic calculus, i.e. a Cartesian product of the form: $S \times O \times \Omega$. This means that access control rules will be described by an information security model, the results of which will cause operation ω_i performed by outside s_i and related to o_i being permitted or prohibited.

If the old version of the ACS is represented as: $S_1 \times O_1 \times \Omega_1$, and the new version as: $S_2 \times O_2 \times \Omega_2$, then for the case where such an expression is true:

$$(S_1 \cap S_2 = \emptyset \wedge (O_1 \cap O_2 = \emptyset) \wedge (\Omega_1 \cap \Omega_2 = \emptyset) \vee (S_1 \setminus S_2 = \emptyset) \wedge (O_1 \setminus O_2 = \emptyset) \wedge (\Omega_1 \setminus \Omega_2 = \emptyset)). \quad (1)$$

Equation (1) assumes that there are no conflicts between the new and old versions of the ACS in ITS. Other cases are not included in the scope of the research.

The studied contradictions occur if the result of taking the Cartesian difference between $S_1 \times O_1 \times \Omega_1$ and $S_2 \times O_2 \times \Omega_2$ is different from $\emptyset$ subset $S_R \times O_R \times \Omega_R$ which, generally, will consist of components of both generations of the ACS of a heterogeneous system. The elements of this set form a common field through which the interaction of different components of the ITS is carried out. Let us analyze what properties the ACS structure should have and what requirements it should meet $S_R \times O_R \times \Omega_R$ to eliminate inconsistencies that lead to the possibility of a violation of the security policy or the occurrence of a hidden information flow.

By hidden information flow, we mean a mechanism by which information may be transferred between various entities within the ITS, bypassing the access control rules (policies).

In this article, the following models are selected in order to study the joint functioning of different computer system security models: discretionary, role-based, mandated, and thematic access control. Information flow security-based and

subject-oriented models used in isolated software environments are taken into consideration as well.

3.2. Discretionary Access Control Models

The known discretionary access control models, i.e. Harrison-Ruzzo-Ullman, typed access matrices, Take-Grant, and extended Take-Grant varieties, operate with an access matrix $M[s, o]$, where rows correspond to subjects, and columns relate to system objects. The cells define the rights that subject s has to access object o [5].

In a scenario in which different versions of these models work together, two situations may take place in which a security policy violation may occur. First, the values of access rights of different versions of ACS for familiar entities and objects are different:

$$\exists s_i, \exists o_j \text{ when } M_1[s_i, o_j] \neq M_2[s_i, o_j],$$

where $s_i \in S_1, s_i \in S_2, o_j \in O_1, o_j \in O_2$.

In this case, it is a clear violation of the security policy applying to the subject and concerning a given object. A given operation is allowed in one version of the ACS and is prohibited in the other. In the second case, the values of access rights of different ACS versions for familiar entities and objects are the same, i.e. $M_1[s_i, o_j] = M_2[s_i, o_j]$, for $\forall s_i \in S_1, \forall s_i \in S_2, \forall o_j \in O_1, \forall o_j \in O_2$. At a given moment, the presence of entities and objects not included another ACS version means that information flows may bypass the security policy of one of the versions (Fig. 3).

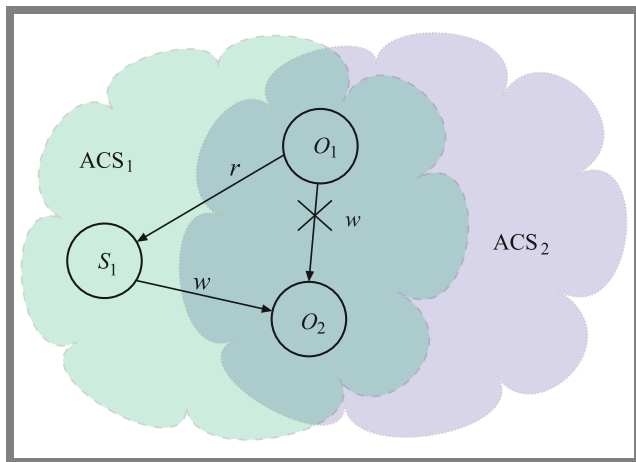


Fig. 3. A variant in which information flow bypasses the security policy of one of the versions of the access control system.

Figure 3 shows one element (subject S_1) present in the first ACS and absent another ACS₂, through which the information flow is carried out, bypassing the access matrix $M_2[s, o]$.

So, when determining the “bypass” data stream, it is necessary to consider all objects and subjects of the system.

3.3. Role-based Access Control Models

Role-based access control models are a further development of discretionary access control policies. The difference is that access rights assigned to specific system entities are

combined to form groups with the same rights. The said groups are determined taking into account the specifics of their functional tasks within the system [5]. Thus, when two ACS solutions relying on the principle of role-based access control work together, an information flow may bypass one of the ACS versions in the same way as it was the case with discretionary access control.

However, equality of rights for each role is not mandatory in the context of the joint functioning of different ACS versions. In this case, if there is an old role r_i^1 and a new ACS r_2^j under which the rights to object o_k are set (if those rights are different, then a clear violation of the security policy of one of the ACS versions is noticed. This is due to the fact that a user performs an operation with role r_i^1 concerning an object that one ACS version considers forbidden in role r_2^j of a different version or vice-versa.

Such a hypothesis has been advanced after researching information flows in scenarios involving joint functioning of different versions of role-based access control systems. To comply with the security policy and ensure coordinated operation of both versions of role-based access control systems, equal sets of access rights for roles of different versions of ACS for combined objects are necessary and sufficient.

3.4. Mandated Access Control Models

The classical Bell-LaPadula model, the secure transition model, the military message system model, the low-watermark policy, the Bib integrity model, and others allow to analyze the conditions under which information flows from objects with a higher level of privacy to the ones with a lower level of privacy are impossible [5]. The occurrence of an unauthorized information flow in one of the cooperating versions of ACS solutions is possible if:

- there are inconsistencies in the lattices of the privacy levels of ACS objects which may lead to an information flow prohibited in one of the versions (Fig. 4),

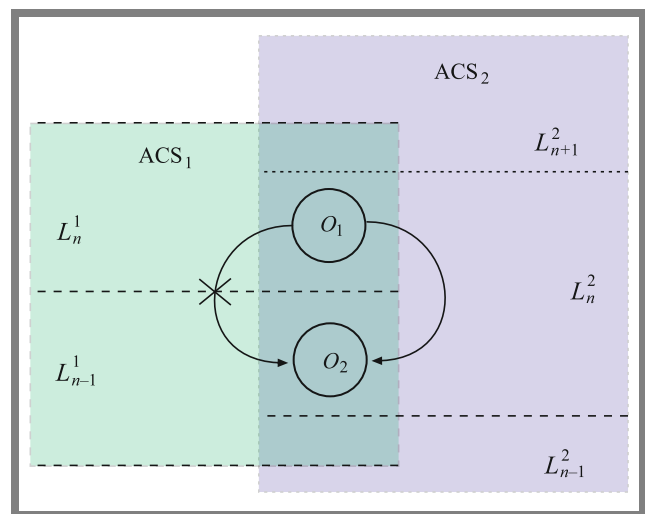


Fig. 4. Information flow bypassing the security policy of one of the ACS versions in a scenario in which the lattices do not match the privacy levels.

- there is a mismatch between access matrices $M[s, o]$.

It is worth noting that the existing regulatory framework clearly defines the grid of privacy levels of specific objects. Hence, the occurrence of an unauthorized information flow in the first method is rare. At the same time, the second method may be implemented in a manner similar to the discretionary access control model.

3.5. Thematic Access Control Models

Thematic access control models operate with a thematic hierarchical classifier (rubricator) which includes a finite set of thematic headings on which a partial order is established and determined by the root tree. Many entities in thematic access control models $X = S \cup O$ are classified based on their mapping to a set of multi-rubrics defined in the root tree of a hierarchical rubricator.

The occurrence of an unauthorized information flow in one of the versions of thematic access control systems during their joint operation is possible if the topics of the system entities do not match (Fig. 5).

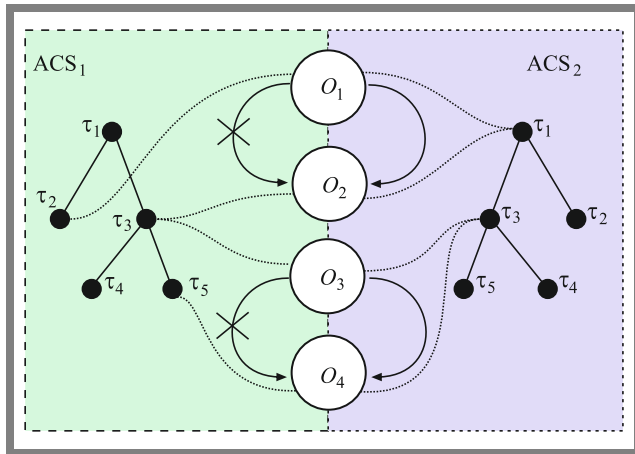


Fig. 5. Information flow in a scenario in which thematic hierarchical classifiers do not match in different versions of access control systems.

As an example, two cases of contradictions in information flows in different versions of thematic access control systems with four shared objects are:

- in the first ACS₁ objects o_1 and o_2 are located in entities that are not comparable in terms of their subject matter. Consequently, such flows are prohibited. At the same time, in ACS₂, these items are related to the same topic, so information flow between them is allowed;
- in the second ACS₁ object o_3 is located in an entity with a broader theme and o_4 with a narrower one, so this flow is prohibited. While in ACS₂ these items are on the same topic, the information flow between them is allowed.

To comply with the security policy when coordinating both thematic access control arrangements, the hierarchical equality of thematic classifiers for objects shared in different ACS versions is necessary.

3.6. Security-based Information Flow Models

Non-withdrawal and non-interference models are based on the distribution of all potential flows between objects within the system into two non-intersecting sets, i.e. authorized and unauthorized flows [5]. Thus, the main task of the security system relying on the security-based information flow model is to prevent the occurrence of unauthorized information flows. This information flow segregation model is mainly based on other security policies, such as discretionary or mandated access control. Thus, when different versions of access control systems relying on the security-based information flow model work together, a violation of security policies is possible. For example, such a breach may occur if the privacy levels do not match or if the access matrices in other types of security policies are not consistent.

3.7. Subject-oriented Model of an Isolated Software Environment

The subject-oriented model of an isolated software environment is similar to the information flow security model and operates with two non-overlapping sets: streams that represent unauthorized access and streams that represent legal (authorized) access. It is worth noting that this model does not specify any known security policies. The security policy describes only the criterion for dividing all information flows into the two subsets. Thus, the subject-oriented model of an isolated software environment is invariant to the security policy adopted in the system and has its disadvantages when different versions of the access control system work together.

4. Methodology for Coordinating Access Control Models

According to [17], threats related to unauthorized access to information constitute a privacy violation. That is why research focusing on access control systems aims to ensure the confidentiality of information. The “information threat” concept requires, in essence, that a subject-object model of a computer system be used, as the need exists for the subject to be an element of the system that becomes acquainted with the information and an object – a material data medium. In the absence of at least one of those two components, we can no longer talk about violating confidentiality. Moreover, it is necessary to introduce a third concept – the flow of information from the object to the subject, since the absence of such flow offers no threat to confidentiality. There are different approaches to defining subjects, objects, and information flows. This work is based the following definitions:

- A subject is a representation of a physical user of a computer system that is created during the user’s entry into the system and is fully characterized by its context (alias, identification code, permissions, etc.);
- An object is a computer system resource that is managed by a set of security tools and is characterized by specific attributes and behaviors;

- Information flow is a transfer of information from one object of a computer system to another.

In order for a confidentiality threat to be present, the concept of “familiarization” of the subject with the information contained in the object needs to be taken into consideration as well. Familiarization consists in the user becoming acquainted with the process and with the information contained in the object. It is worth noting that according to the definitions adopted, a confidentiality violation occurs when the user has read the information and the respective process has received it.

When developing or upgrading an ITS, the developer implements the ACS based on proven security policies and ensures information protection based on the criteria established by the customer. At the same time, the upgraded components are implemented in the system in stages and over a certain period of time during which the old version of the ISS functions simultaneously with the upgraded ACS. Such an approach leads to uncontrolled actions concerning individual objects being undertaken by one of the versions of the ACS. This means that information flows may emerge that comply with the security policy framework and ones that fail to do so.

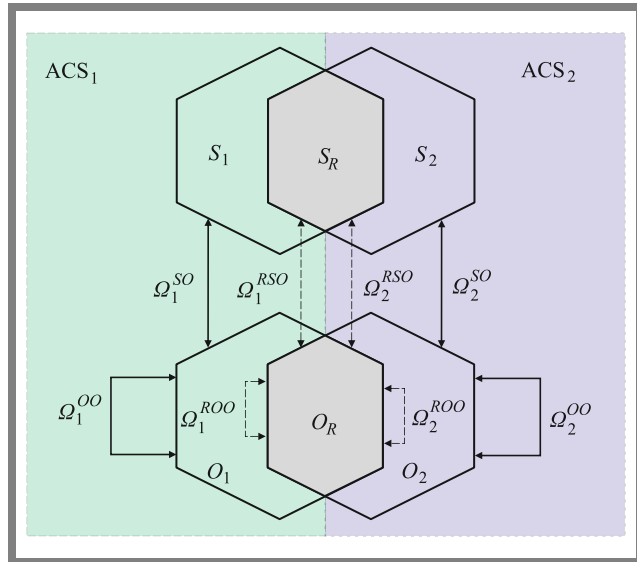


Fig. 6. Model of a scenario in which different versions of access control systems operate jointly.

In general, the modernization period is characterized by the occurrence of information flows that are not allowed in one of the versions, thus leading to a violation of the confidentiality of information. The confidentiality requirement will be met when equal information flows are ensured for common subjects and objects of the system in both versions. To prove the safety theorem for matching different versions of ACS, we describe a structural model of their joint functioning (Fig. 6):

$S_R = S_1 \cap S_2$ are common subjects for both versions,

$O_R = O_1 \cap O_2$ are shared objects for both versions,

$\Omega_1^{SO}, \Omega_2^{SO}$ are information flows between ACS subjects and objects,

$\Omega_1^{RSO}, \Omega_2^{RSO}$ are information flows between familiar entities and ACS objects,

$\Omega_1^{OO}, \Omega_2^{OO}$ are information flows between ACS objects,

$\Omega_1^{SOO}, \Omega_2^{SOO}$ are information flows between shared ACS objects.

In addition:

$$\Omega_1^{RSO} \subseteq \Omega_1^{SO} \subseteq \Omega_1, \Omega_2^{RSO} \subseteq \Omega_2^{SO} \subseteq \Omega_2,$$

$$\Omega_1^{ROO} \subseteq \Omega_1^{OO} \subseteq \Omega_1, \Omega_2^{ROO} \subseteq \Omega_2^{OO} \subseteq \Omega_2.$$

Note that a violation of confidentiality or the rules of the ACS security policy is only possible if there is an unauthorized information flow between the object containing the information and the subject. Therefore, information flows between subjects are not considered in this model. Obviously, if all information flows are equal, no violation of confidentiality takes place.

Let us define a necessary and sufficient condition under which an unauthorized information flow is impossible. To this end, we will consider four degenerate variants of joint functioning of both versions of the ACS, provided that each system fully ensures the confidentiality of information following the basic models of the ACS:

Option 1. In the case when $S_R = \emptyset, O_R = \emptyset$ the ACS involved do not have any common elements and function independently of each other. There are no unauthorized information flows.

Option 2. In the case when $S_R \neq \emptyset, O_R = \emptyset$ the ACS only have familiar users. The presence of an object containing information prohibited from being made available to the user is a necessary condition for violating the privacy requirement. Concurrently, the ACS security policy will not allow such information to be present in an object that belongs to the user (the has access to such an object). Thus, in this case, there is no violation of confidentiality.

Option 3. In the case when $S_R = \emptyset, O_R \neq \emptyset$ the ACS have only common objects. This option allows for an information flow between objects not provided for in one of the versions of the ACS, which in turn will lead to an unauthorized information flow to a particular subject. Thus, this option may violate confidentiality.

Option 4. In the case when $S_R \neq \emptyset, O_R \neq \emptyset$ the ACS have common subjects and objects. This option combines the second and third cases and allows for an information flow between objects not provided for in one of the versions of the ACS. Thus, this option may violate confidentiality.

The analysis of various options concerned with the sharing of ACS elements has shown that an unauthorized information flow can only occur if there are common objects in different versions of ACS. Thus, to prevent unauthorized information flows during the joint functioning of both versions of the ACS, equality of data flows between objects shared in each of the ACS is necessary. Such an approach serves as a formulation of the security theorem for the joint functioning of different versions of ACS. The proof of the theorem is clear from the above.

It is worth noting that the developed approach is invariant to the access differentiation models themselves, which makes it possible to coordinate not only systems of the same type, but also systems built based on different access control models.

5. Conclusions

The study was concerned with the joint operation of various versions of access control systems that are based on security models applicable to computer systems. This article has identified potential ways of ensuring that information flows bypass the security policies of one of the versions of the access control systems involved.

Discretionary access control models rely on an access matrix and use a privacy level grid. This means that some degree of coordination is required when different versions of ACS operate together. For role-based and thematic access control models, specific hypotheses are formulated allowing to achieve compliance with the security policies applicable to the joint operation of different versions of ACS.

The results of the study showed that the methodological basis for coordinating access control models applicable to information and telecommunication systems undergoing modernization consists in observing, separately, the equality of information flows between shared objects in each of the versions of the access control systems concerned. By relying on the methodological basis developed, we will be able to continue with the development of other methods for coordinating various access control models.

References

- [1] H. Huang, F. Shang, J. Liu, and H. Du, "Handling least privilege problem and role mining in RBAC", *Journal of Combinatorial Optimization*, vol. 30, no. 1, pp. 63–86, 2015 (<https://doi.org/10.1007/s10878-013-9633-9>).
- [2] J. Hassan *et al.*, "A lightweight proxy re-encryption approach with certificate-based and incremental cryptography for fog-enabled e-healthcare", *Security and Communication Networks*, vol. 2021, Article ID 936824, 2021 (<https://downloads.hindawi.com/journals/scn/2021/936824.pdf>).
- [3] H. Zhang, J. Wang, and J. Chang, "An access control model for multilevel security in multi-domain networking environments", *Proceedings of the 9th International Conf. on Modelling, Identification and Control (ICMIC)*, pp. 809–814, Kunming, China, 2017 (<https://doi.org/10.1109/ICMIC.2017.8321566>).
- [4] M.U. Aftab, A. Hamza, A. Oluwasanmi, X. Nie, M.S. Sarfraz, D. Shehzad, Z. Qin, and A. Rafiq, "Traditional and hybrid access control models: A detailed survey", *Security and Communication Networks*, vol. 2022, Article ID 1560885, 2022 (<https://doi.org/10.1155/2022/1560885>).
- [5] S. Pierangela and S. de Capitani di Vimercati, "Access control: Policies, models, and mechanisms", in *International School on Foundations of Security Analysis and Design*, R. Focardi and R. Gorrieri, Eds. LNCS, vol. 2171, pp. 137–196. Berlin, Heidelberg: Springer, 2000 (https://doi.org/10.1007/3-540-45608-2_3).
- [6] D.J. Bokefode, A.S. Ubale, S.S. Apte and G.D. Modani, "Analysis of DAC MAC RBAC access control based models for security", *International Journal of Computer Applications*, vol. 104, no. 5, pp. 6–13, 2014 (<https://doi.org/10.5120/18196-9115>).
- [7] A.K. Malik, N. Emmanuel, S. Zafar, H.A. Khattak, B. Raza, S. Khan, A.H. Al-Bayatti, M. O. Allassafi, A. S. Alfakeeh, and M. A. Alqarni,

"From conventional to state-of-the-art IoT access control models", *Electronics*, vol. 9, no. 10, 1693 (<https://doi.org/10.3390/electronics9101693>).

- [8] O.K. Yudin and M. A. Strelbitskiy, "Content and hierarchy of the register of information resources of the state border guard service of Ukraine", *Problems of Informatization and Management*, vol. 4, no. 56, pp. 85–91, 2016 (<https://doi.org/10.18372/2073-4751.4.13148>) [in Ukrainian].
- [9] M.A. Strelbitskiy, "Analysis of joint functioning of access differentiation models at the stage of modernization of information and telecommunication systems", *Collection of Scientific Works of Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine*, vol. 4, no. 70, pp. 276–287, 2016 [in Ukrainian].
- [10] V.V. Kuzavkov, M.A. Strelbitskiy, and V.O. Danko, "Method for harmonizing the privacy level grids of mandatory access control systems for information and telecommunication systems at the modernization stage", *Collection of Scientific Works of the Military Institute of Telecommunication and Informatization*, no. 1, pp. 56–60, 2017 (http://nbuv.gov.ua/j-pdf/Znpviti_2017_1_9.pdf) [in Ukrainian].
- [11] M.A. Strelbitskiy, "Method of coordination of access matrices of discretionary access control systems of information and telecommunication systems at the stage of modernization", *Modern Information Technologies in the Field of Security and Defense*, vol. 1, pp. 58–62, 2017 [in Ukrainian].
- [12] Y. Deng, J. Wang, J.J.P. Tsai, and K. Beznosov, "An approach for modeling and analysis of security system architectures", *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 5, pp. 1099–1119, 2003 (<https://doi.org/10.1109/TKDE.2003.1232267>).
- [13] J.E. Kobza and S.H. Jacobson, "Probability models for access security system architectures", *J. of the Operational Research Society*, vol. 48, no. 3, pp. 255–263, 1997 (<https://doi.org/10.2307/3010424>).
- [14] Ş. Bahtiyar and M.U. Çağlayan, "Extracting trust information from security system of a service", *Journal of Network and Computer Applications*, vol. 35, no. 1, pp. 480–490 (<https://doi.org/10.1016/j.jnca.2011.10.002>).
- [15] M. Fugini and G. Martella, "Acten: A conceptual model for security systems design", *Computers and Security*, vol. 3, no. 3, pp. 196–214, 1984 ([https://doi.org/10.1016/0167-4048\(84\)90041-5](https://doi.org/10.1016/0167-4048(84)90041-5)).
- [16] A. Boukerche and Y. Ren, "A trust-based security system for ubiquitous and pervasive computing environments", *Computer Communications*, vol. 31, no. 18, pp. 4343–4351, 2008 (<https://doi.org/10.1016/j.comcom.2008.05.007>).
- [17] T. Carlson, "Information Security Management: Understanding ISO 17799", International Network Services Inc. (INS), Whitepaper, 2001 (http://www.secureict.co.za/wp-content/uploads/2018/06/03_ins_info_security_iso_17799_1101-1.pdf).

Mykhailo Strelbitskiy, D.Sc.

Professor at the Department of Communications, Automation, and Cyber-security

 <https://orcid.org/0000-0001-8030-3228>

E-mail: m.strelb@ukr.net

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

Valentyn Mazur, D.Sc.

Professor at the Department of Border Security

 <https://orcid.org/0000-0002-3405-6200>

E-mail: vumazur154@gmail.com

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

Evgenii Prokopenko

Candidate of Technical Sciences, Associate Professor at the Department of Communications and Information Systems

 <https://orcid.org/0000-0002-6825-2357>

E-mail: mydocent@gmail.com

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

Roman Rachok, D.Sc.

Professor at the Department of Telecommunication and Information Systems

 <https://orcid.org/0000-0002-3283-9690>

E-mail: rrvnadpsu@i.ua

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

Dmytro Mul

Candidate of Technical Sciences, Associate Professor at the Department of Communications and Informatization

 <https://orcid.org/0000-0003-2662-4770>

E-mail: dmitry.mul@gmail.com

Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine, Khmelnytskyi, Ukraine

Adaptive Rider Feedback Artificial Tree Optimization-Based Deep Neuro-Fuzzy Network for Classification of Sentiment Grade

Sireesha Jasti and G.V.S. Raj Kumar

Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India

<https://doi.org/10.26636/jtit.2023.165222>

Abstract — Sentiment analysis is an efficient technique for expressing users' opinions (neutral, negative or positive) regarding specific services or products. One of the important benefits of analyzing sentiment is in appraising the comments that users provide or service providers or services. In this work, a solution known as adaptive rider feedback artificial tree optimization-based deep neuro-fuzzy network (RFATO-based DNFN) is implemented for efficient sentiment grade classification. Here, the input is pre-processed by employing the process of stemming and stop word removal. Then, important factors, e.g. SentiWordNet-based features, such as the mean value, variance, as well as kurtosis, spam word-based features, term frequency-inverse document frequency (TF-IDF) features and emoticon-based features, are extracted. In addition, angular similarity and the decision tree model are employed for grouping the reviewed data into specific sets. Next, the deep neuro-fuzzy network (DNFN) classifier is used to classify the sentiment grade. The proposed adaptive rider feedback artificial tree optimization (A-RFATO) approach is utilized for the training of DNFN. The A-RFATO technique is a combination of the feedback artificial tree (FAT) approach and the rider optimization algorithm (ROA) with an adaptive concept. The effectiveness of the proposed A-RFATO-based DNFN model is evaluated based on such metrics as sensitivity, accuracy, specificity, and precision. The sentiment grade classification method developed achieves better sensitivity, accuracy, specificity, and precision rates when compared with existing approaches based on Large Movie Review Dataset, Datafiniti Product Database, and Amazon reviews.

Keywords — *deep learning network, feedback artificial tree, natural language processing (NLP), rider optimization algorithm, sentiment grade classification.*

1. Introduction

Sentiment analysis [1], [2] is a growing field of research focusing on analyzing product and service reviews published online. Sentiment analysis is considered to be a “baggage issue”, which deals with polarity identification [2], [3] aspect extraction [4], [5], and natural language processing (NLP) tasks [2], [6]. Aspect-level sentiment classification is the basic approach in sentiment analysis, and it serves to deduce sentiment polarity. For instance, whether a given sentence is

classified as neutral, positive or negative in sentiment depends on finding a specific element therein [2]. Sentiment analysis is effective for evaluating users' opinions with regards to services or products described by text data [7], [8]. Therefore, its primary task is to identify and categorize the user's polarity based on content [9]–[11]. Sentiment analysis is also the main research area in NLP, relied upon to analyze the opinions and emotions of users from the texts they provide [12].

Sentiment analysis is usually harnessed by e-commerce websites and online social networks [13] for predicting their users' needs based on polarity or sentiment categories [4]. The analyses are performed with the intention to identify specific options and recognize their polarity. Sentiment is categorized into three sets (neutral, negative, and positive) with five different groups existing within each of those sets (strong opposition, opposition, support, strong support, and neutrality) [14].

Subjective information from the text concerned is recognized automatically using emotions, opinions, attitudes, etc. [15] and plays a vital part in NLP and data mining [16]. Three levels (target, sentence, and document) are taken into consideration when analyzing sentiment. The sentence and document levels are intended to categorize sentences and documents, and they consider the entire document or sentence as the basic information-carrying component. Polarity sentiment is also analyzed based on textual content retrieved from a sentence or document [17]. The features extracted from texts are subjected to classification in the corpus-based approach. In addition, the intensification and negation process is conducted by combining a lexicon-based technique for identifying scores from input text based on keywords and phrases dictionary [18]. Strength and polarity are specified, in sentiment dictionaries, using corpus words and phrases [12]. The entire process is relied upon for boosting sales and creating product rankings [19], [20].

Recently, neural networks have been used in several areas, such as discrete-time signals, computer vision, and NLP for sentiment classification [2], as a widely utilized approach for an attention model for aspect-level sentiment classifica-

tion to recognize the sentiment polarity of targets in context [21], [22]. The deep learning technique [23], [24] performs better than the machine learning approach when it comes to sentiment classification, due to the fact that GPU accelerators are capable of relying on large databases [8], [25]. Adaptive RNN is used in the target-based classification approach, where the sentiment of the wording is transferred to the target based on syntactic and context-related patterns [14], [26]. Tree structures of Long Short-Term Memory (LSTM) networks are also used in the sentiment classification process [16], [27], as are the Gated Recurrent Unit (GRU), LSTM, and gated RNN with LSTM [4].

The major goal of this work is to design and develop an A-RFATO-based DNFN for sentiment classification. The paper provides a meaningful contribution to two main topics:

- **Adaptive RFATO-based DNFN** for efficient sentiment grade classification with a DNFN trained by A-RFATO,
- **Adaptive RFATO** developed by integrating FAT and ROA approaches with the adaptive concept.

The rest of the paper is organized as follows. Section 2 reviews the literature and Section 3 elaborates on the proposed A-RFATO-based DNFN model. Section 4 presents the results and the discussion, while the conclusion is given in Section 5.

2. Literature Review

The CNN approach was developed by Kim and Jeong in [28] to classify sentiment. The input image was taken from a database and then its significant features were extracted for further processing. The flattening process was executed for converting a two-dimensional feature map to a one-dimensional structure. This approach easily confines higher level and local patterns for better precision. However, this method has not identified any improved patterns for sentiment classification. Fu *et al.* [29] introduced a semi-supervised aspect level sentiment classification approach using a variational auto encoder (AL-SSVAE) and employed in during the classification process. The topical word embedding (TWE) method was utilized for identifying aspect-specific word embedding, while joint sentiment topic (JST) was employed for obtaining sentiment from given words. The proposed method efficiently extracts sentiment features and provides global semantics for better performance. Nevertheless, this method did not include a particular neural network for identifying aspect-level sentiments of various texts.

A multi-layer attention-based CNN was modeled by Zhang *et al.* [14]. In this paper, every conventional layer output is considered to form a context representation for capturing additional context features. This method extracts high-level features, like fractional context, and low-level features, such as semantic relationships, phrases, and words, using these in the classification process. This technique utilizes sentiment-related vocabulary for the classification process, but offers poor results in complicated modeling.

Chen *et al.* [2] proposed a fusion model with multi-source data to classify sentiment with unified data structure. The model collects knowledge from different types of resources. After that, aspect-level sentiment is categorized using bidirectional encoder representations from transformers (BERT). The use of BERT means that the classification is performed in a precise manner, in the form of a good aspect-specific sentence representation. Unfortunately, this method fails to cover various resources, such as sentiment knowledge bases and document-level corpora in aspect-level classification.

A graph convolutional network (GCN) was introduced by Zhao *et al.* in [30]. Initially, aspect-level representations were captured by employing a bi-directional attention model using position encoding. Then, a multi-aspect sentiment classification structure was developed to efficiently confine sentiment dependencies among several features. This model effectively classifies the sentiment based on various aspects, but fails to overcome the problem of over-fitting. Salur and Aydin introduced, in [8], a hybrid deep learning technique to classify sentiment. At first, features were extracted using character-level embedding based on CNN and BiLSTM. Next, the extracted features were integrated and broadcast to the SoftMax layer for the classification process. The approach produced high level of classification accuracy. However, attention-based approaches were not deployed for increasing performance. Zhang *et al.* [31] introduced a bi-channel capsule network (SC-BiCapsNet) for sentiment classification. The Cot-Att scheme was used in text semantic representation to optimize text expression and coding patterns. Finally, a capsule network was developed for improving the weight value of significant features in the text classification phase. This method efficiently increases accuracy of the classification process, but failed to incorporate other auxiliary networks in order to decrease computational cost. A multi-task learning technique was devised by Jin *et al.* in [32], with multi-scale CNN and LSTM helping classify sentiment. This approach effectively enhances the quality of the encoder and emotion classification. Time complexity was reduced significantly, but it failed in enhancing sentiment classification outcomes for multitask learning.

3. Proposed A-RFATO-based DNFN

The proposed A-RFATO technique is a combination of the adaptive ROA concept and the FAT approach. The diagram of the classification model is presented in Fig. 1.

A dataset P containing reviews can be expressed by:

$$P = \{X_{a,b}\} \quad 1 \leq a \leq A, \quad 1 \leq b \leq B, \quad (1)$$

where $X_{a,b}$ represents data from the a -th review and with a b -th attribute, A and B indicate all data points and attributes.

The input review data is treated as input in the pre-processing stage and is subjected to stemming and stop word removal processes.

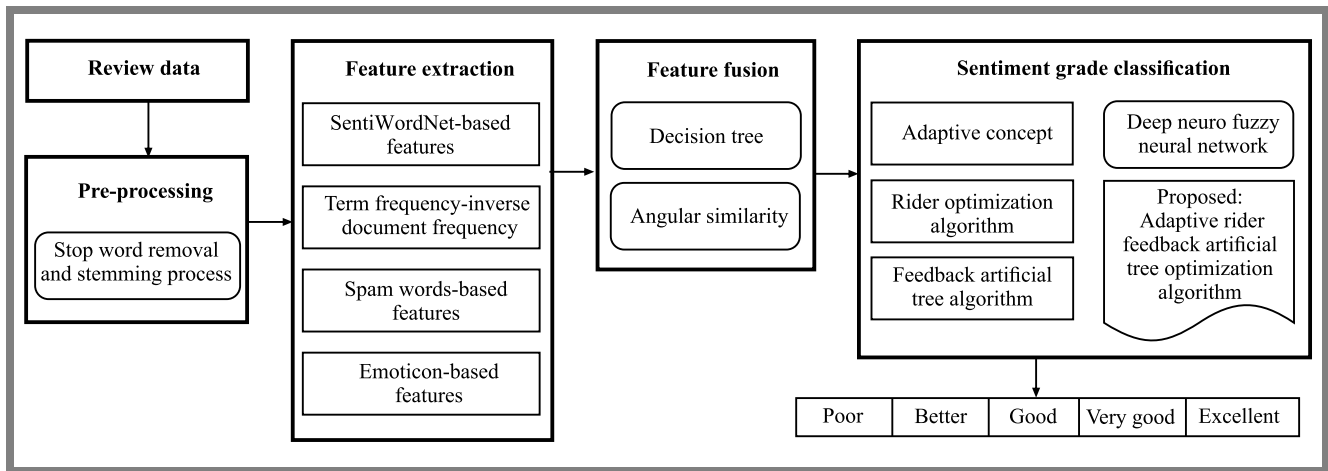


Fig. 1. Diagram of the A-RFATO-based deep neuro-fuzzy network for sentiment grade classification.

3.1. Stop Word Removal

It is an important process in the pre-processing stage, as it allows to avoid processing huge amounts of data in a speedy manner. The current goal is to perform the text examination phase while keeping the negation and intensity of the words in mind in order to effectively complete the sentiment analysis. The Python natural language toolkit (NLTK) package has been used to remove the stop words. Negation words are deleted from the list of stop words and are applied to the input review text. The stop word removal method decreased the amount of information noise by eliminating non-informational behavior from the sentences.

3.2. Stemming

The retrieved words are trimmed to their base form by applying the stemming process, as the processing of large data sets increases complexity of the sentiment classification process. The stemming process allows to avoid complications by eliminating unnecessary words from the input text file. The stemming process generates a compressed document by removing prefixes and suffixes. The output is formulated as:

$$X = \{C_d, 1 \leq d \leq n\}, \quad (2)$$

where C_d is the pre-processed data and n is the total number of reviews.

After data pre-processing, the extraction of features is performed. The features taken into consideration are explained below.

3.3. SentiWordNet Features

The establishment of a SentiWordNet structure is a two-step procedure comprising exploration of WordNet and determination of a subset of WordNet. Initially, WordNet term relationships, such as hyponyms, synonyms, and antonyms are explored for enlarging the seed word's core. Next, a subset of WordNet terms, i.e. negative and positive labels, is

identified [33]. Positive and negative scores of SentiWordNet features are given by:

$$\{M_x, N_x\} = SW(C_d), \quad (3)$$

where a positive score is given by M_x and a negative score is denoted by N_x , while $SW(C_d)$ indicates the SentiWordNet features with C_d data. The features identified from the negative and positive score are:

$$m_1^d = M_x, m_2^d = N_x, \quad (4)$$

where m_1^d and m_2^d are features acquired from positive and negative scores of SentiWordNet. These features are applied to train the classifier algorithm. After that, the mean, variance, and kurtosis of positive and negative scores of SentiWordNet are extracted.

The mean offers an estimation of the average amount of data in a document and is expressed by:

$$m_3^d = \frac{1}{e} \sum_{i=1}^e E_i, \quad (5)$$

where $i = 1, \dots, e$, $e = 1$ or 2 , which means positive and negative scores, respectively. The total amount of positive and negative data enclosed in a document is specified as E_i while m_3^d indicates the mean value.

Variance is based on the mean value that is specified as m_4^d , and is formulated as follows:

$$m_4^d = \frac{1}{e} \sum_{i=1}^e (E_i - m_3^d)^2. \quad (6)$$

Kurtosis is a measure of tailedness of a distribution, i.e. how much density there is in the tails, and is denoted by m_5^d .

3.4. TF-IDF Features

TF-IDF features are extracted using pre-processed data, based on the vector space model (VSM), while the weight of words

is estimated by IDF, whereas the occurrence of words is evaluated by DF. The TF feature is expressed by:

$$m_6^d = \frac{1}{e} \sum_{\substack{z=1 \\ y \in z}}^{|e|} TF_d^z, \quad (7)$$

where m_6 represents the TF features acquired from the review document. The word count and frequency of a z -th word in a d -th review is denoted as $|e|$ and TF_d^z . The IDF feature is calculated by:

$$m_7^d = \sum_{\substack{z=1 \\ y \in z}}^{|e|} IDF_d^z, \quad (8)$$

where m_7 is the IDF feature and it is extracted from the review document.

3.5. Feature Based on Spam Word

Identification of spam is an important step, as it identifies precise feedback using customers' product reviews. Spam reviews (not valid) also help estimate the best sentiment scores using various reviews. The spam-based feature is estimated by analyzing the content present in the reviews using such review properties as large word count, or detection of a higher number of pronouns. In contrast, reliable reviews involve more coordinating conjunctions, adjectives, nouns, determiners, and prepositions [34]. The features acquired from spam identification are represented by:

$$m_8^d = \frac{N_x}{|e|}. \quad (9)$$

The spam word and entire word counts are identified as N_x and $|e|$, respectively. For a d -th review:

$$m_9^d = \frac{e_{N_z}}{|e|}, \quad (10)$$

where e_{N_z} is the sum of the frequency of spam words.

3.6. Emoticon-based Feature

Emoticons are visual depictions of face expressions made up of such characters as letters, punctuation marks, and numbers. Each emoticon intensity is measured based on the text preceding that particular emoticon. The difference in the incidence of negative and positive emoticons is estimated as m_{10}^d .

3.7. Feature Fusion by Decision Tree and Angular Similarity

Once the feature extraction process is completed, emoticon-based features, SentiWordNet features, mean, variance, kurtosis, spam word-based features, and TD-IDF features are combined to form feature vectors in order to decrease complexity of sentiment estimation. Feature vectors are formulated as:

$$\mathbf{m} = \{m_1^d, m_2^d, m_3^d, m_4^d, m_5^d, m_6^d, m_7^d, m_8^d, m_9^d, m_{10}^d\}, \quad (11)$$

where $m_1^d, m_2^d, \dots$ indicate the feature set. The final fused feature vector is expressed as:

$$\mathbf{m}_i^{fused} = \sum_{y=1}^f \frac{\mu}{y} m_y, \quad (12)$$

$$y = 1 + \frac{N}{M}, \quad (13)$$

where $i = 1, \dots, z$, the selected features are indicated as z , N is total feature number and M is the number of features which are to be selected.

3.8. Decision Tree

This subsection describes the procedure of identifying the μ parameter using the decision tree. First, the review content is trained to group the data under division. The μ parameter is identified as the entropy function of the review data if it belongs to a specific group. The diagram illustrating the process of identifying μ via the decision tree and the ground truth is presented in Fig. 2 and is identified by:

$$\mu = AS(n_d, \beta_d), \quad (14)$$

where n_d is the review data and β_d is the average of the review data belonging to a specific class.

	m_1	m_2	$\dots$	m_y	Ground truth
n_1					μ_1
n_2					μ_2

Fig. 2. Determination of the μ parameter.

Then, angular similarity is computed based on the angular distance measure estimated by cosine similarity:

$$AD = \frac{\cos^{-1}(CS)}{\pi}, \quad (15)$$

$$AS = 1 - AD, \quad (16)$$

where AD is the angular distance, CS represents cosine similarity, and AS specifies angular similarity.

The decision tree is further executed for the purpose of the rule induction and classification process. The decision tree is basically a flowchart that has a tree-like shape, in which the test result is specified by an attribute test and a branch. The class label indicated by the leaf node and internal node is used to determine attribute test and branch. The source set is divided into subsets based on the attribute value test in the decision tree and every subset endures recursive way splitting. The recursive portioning ends when the subset node has a target with a similar variable. The decision is characterized by better accuracy and efficiently manages multi dimensional data. The target occurrence is categorized in the decision tree through arranging the tree from the root node followed by attribute testing and shift to the leaf node. This procedure is repeated for the sub trees of new nodes.

3.9. Classification of Sentiment Grade

The input for categorization of the sentiment grade using DNFN has the form of feature fused output. The DNFN is trained using the A-RFATO technique which incorporates ROA, FAT, and adaptive concepts to reduce computational complexity.

3.10. DNFN Structure

The DNFN [35] is a hybrid of fuzzy logic and a deep neural network. In this network, the first processing step is performed based on a deep neural network and then the second processing step is executed by means of fuzzy logic in order to estimate the objective of the system. This system involves three types of layers: input layer, a number of hidden layers for verification, and the output layer. Three rules are applied, including de-fuzzification layer, normalization layer, and rule layer. The output layer is the de-fuzzification layer. Consequents and premises constitute very important parameters of deep neural networks. Premises form the pedestal of the membership function is realized in the modified fuzzification in the input layer, also it deals with the incidence levels. Besides, the consequent parameter depends on the de-fuzzification procedure. A neuro-fuzzy block contains the fuzzy interface system (FIS) used for rule base evaluation. The DNFN structure used for the classification of sentiment is shown in Fig. 3.

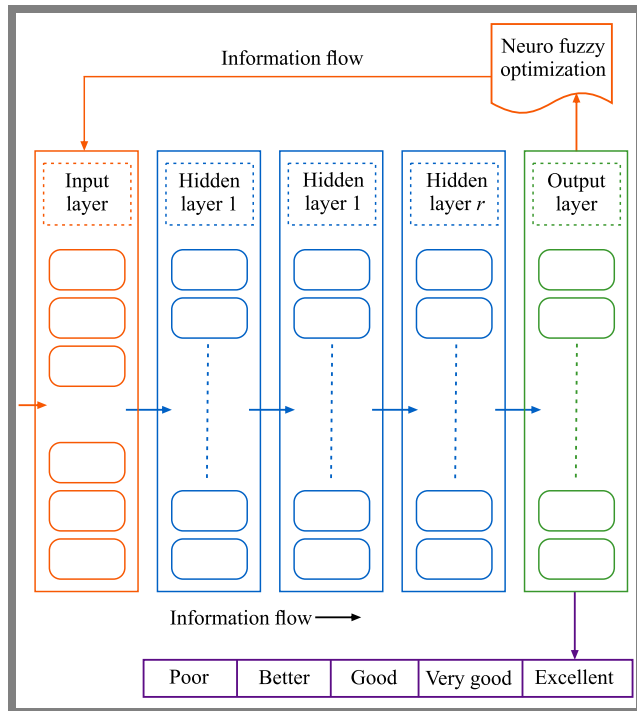


Fig. 3. DNFN structure diagram.

The degree of each input is distributed within the 0–1 range. Output follows every entity of the first layer. Let us consider a case in which the number of premises is two, with one consequent as:

$$J_{1,h} = kC_h(r) \text{ or } J_{1,h} = kC_{h-2}(t), \quad \forall h = 1, 2, 3, \quad (17)$$

where r represents input to each h -th entity, kC and kC_{h-2} specify antecedent membership function and degree of the membership function is indicated by $J_{1,h}$.

The bell-shaped function is used to model the membership function which is distributed with a minimum value of 0 and a maximum value of 1. It can be formulated by:

$$kC_h(r) = \frac{1}{1 + \left| \frac{h-A_h}{B_h} \right|^{2T_h}}, \quad (18)$$

where T_h , A_h , and B_h describe the membership function of the premise parameter, enhanced with the help of training.

The rule-based layer refers to layer 2, as it is used to express a group of rules. Each entity in the layer is multiplied by linguistic variables to satisfy the membership degree. The rules' firing strength is symbolized by the multiplication of variable membership values as:

$$J_{2,h} = \rho_h = kC_h(r)kC_{h-2}(t), \quad \forall h = 1, 2. \quad (19)$$

Layer 3 is the normalization level in which each entity calculates the ratio between the firing strength of the h -th rule and the total firing strength of all rules, and ρ_h refers to a generic network parameter. The result for each rule is normalized based on the firing strength of that rule and equals:

$$J_{3,h} = \bar{\rho}_h = \frac{\rho_h}{\rho_1 + \rho_2} \quad \forall h = 1, 2. \quad (20)$$

Layer 4 describes the de-fuzzification process, wherein the estimation of each rule consequent is performed for the outputs specified, and equals:

$$J_{4,h} = \bar{\rho}_h E_h = \bar{\rho}_h (H_h r + I_h r + K_h), \quad \forall h = 1, 2, \quad (21)$$

where H , I , and K indicate consequent parameter sets. The process of final outcome estimation is:

$$J_{5,h} = \sum_h \bar{\rho}_h E_h = \frac{\sum_h \rho_h E_h}{\sum_h \rho_h}. \quad (22)$$

The parameters are then distributed using random values and are tuned by means of training to obtain optimized results. The proposed A-RFATO is utilized in such training and the extent of hidden layers is used for efficient system training on a large amount of data.

3.11. DNFN Training

Below is presented an 8-step procedure used for training.

Step 1 – initialization of riders. The riders are initialized randomly by:

$$G_\sigma = \{G_\sigma(p, q)\}; \quad 1 \leq p \leq U; \quad 1 \leq q \leq V, \quad (23)$$

where σ is a moment in time, U is the total number of riders, p, q is the p -th rider location $G_\sigma(p, q)$, and V denotes the total coordinates of the dimension. After group initialization, rider parameters are initialized as well: gear I_σ , steering S_σ , brake R_σ , and accelerator A_σ .

Step 2 – computation of fitness. Estimation of the fitness function is performed by identifying the space between the

final destination and the location of the rider. The solution with a minimum error value is considered to be the optimal one. The following equation is used to compute the fitness value:

$$MSE = \frac{1}{v} \left[\sum_{j=1}^v Z_{target} - w^o \right], \quad (24)$$

where Z_{target} is the target output, v denotes the number of training samples, and w^o indicates the estimated output from DNFN.

Step 3 – leading rider determination. After computation of the fitness function, the leading rider is calculated. The rider with lowest fitness value is selected to be the leading rider and its location varies at each moment.

Step 4 – updating rider location. The location of the leading rider is estimated by updating the rider location at a given time and the location of the bypass rider in the group. Then, the follower is updated.

The bypass rider avoids the regular route, without bearing in mind the pathway of the leading rider. Location of the bypass rider is updated by:

$$G_{\sigma+1}(p, q) = \lambda \left[G_{\sigma}(\chi, q) \cdot \eta(q) + G_{\sigma}(\gamma, q) \cdot (1 - \eta(q)) \right], \quad (25)$$

where χ and γ are the random numbers from 1 to R . An adaptive concept is realized by using η and terms λ for better performance of sentiment classification, expressed as:

$$\eta = \frac{fitness_i}{\max(fitness)}, \quad (26)$$

$$\lambda = \frac{\tau}{T}, \quad (27)$$

The follower rapidly achieves its destination by updating the location of the leading rider. The area of the follower is modified using the values chosen and the location of the follower is influenced by the coordinate selector as:

$$G_{\sigma+1}(p, q) = G^K(K, q) + [\cos(KF_{p,q}^{\sigma}) \cdot G^K(K, q)] + g_p^{\sigma}, \quad (28)$$

where G^K and K indicate the leading rider's location and index, the distance of the p -th rider is denoted as g_p^{σ} , q is the coordinate selector and $KF_{p,q}^{\sigma}$ denotes the steering angle in the q -th coordinate of p -th rider.

The factors affecting the location update of the overtaker include the following: relative success rate, coordinate selector, and direction indicator. The overtaker provides faster convergence with a huge global neighborhood, thus the self-evaluation operator of FAT is used to update the value. The final equation of the overtaker updated with FAT is expressed as:

$$G_{pq}(\sigma + 1) = \frac{[1 - \text{rand}(0, 1)]}{\text{rand}(0, 1)} \left[\frac{G_{best,q} \text{rand}(0, 1)}{1 - \text{rand}(0, 1)} + D_p^Z(\sigma) G_{Kq}^K \right], \quad (29)$$

where $G_{pq}(\sigma)$ is the p -th rider location in the q -th coordinate and $D_p^Z(\sigma)$ denotes the direction indicator of the p -th rider at instant σ . The location of the best branch is indicated by $G_{best,q}$, $\text{rand}(0, 1)$ denotes a random number between 0 and 1.

Step 5 – updating attacker location. The attacker tries to reach the leader's location and uses the same update procedure as the follower. The attacker location is updated using the following formula:

$$G_{\sigma+1}(p, q) = G^K(K, q) + [\cos(KF_{p,q}^{\sigma}) \cdot G^K(K, q)] + g_p^{\sigma}, \quad (30)$$

where $G_{\sigma+1}(p, q)$ is the leading rider location, $KF_{p,q}^{\sigma}$ denotes the steering angle in the q -th coordinate of the p -th rider and the distance of the p -th rider is expressed as g_p^{σ} .

Step 6 – calculation of fitness function. Once the location of each rider has been estimated, the rider with the best fitness value is selected as the race champion.

Step 7 – updating rider parameters. The activity counter is computed to update the rider parameters in order to identify the optimal solution.

Step 8 – termination. The process described above is iterated until the time is reached.

The pseudo code of the adaptive RFATO technique is specified in Algorithm 1.

Algorithm 1. Pseudocode of the proposed A-RFATO method.

- 1: **Input:** rider's random position G_{σ} .
- 2: Initialize the population
- 3: Initialize rider parameter: gear I_{σ} , steering S_{σ} , brake R_{σ} , and accelerator A_{τ}
- 4: Fitness function computation
- 5: While $\sigma < \sigma_{off}$
- 6: For $p = 1$ to U
- 7: Position update of bypass rider using Eq. (25)
- 8: Include adaptive concept by $\eta = \frac{fitness_i}{\max(fitness)}$ and $\lambda = \frac{\tau}{T}$
- 9: Modify the area of follower based on Eq. (28)
- 10: Modify the location of overtaker based on Eq. (29)
- 11: Position update of attacker based on Eq. (30)
- 12: Ranking of riders based on fitness function
- 13: Rider with maximum fitness function is taken as leading rider
- 14: Modify I_{σ} , S_{σ} , R_{σ} , and A_{σ}
- 15: Return G^K
- 16: $\sigma = \sigma + 1$
- 17: End for
- 18: End while
- 19: **Output:** Leading rider G^K

4. Results and Discussion

To evaluate performance of proposed algorithm, the following metrics are used:

Accuracy is determined in order to describe the correctness of the sentiment classification process and is expressed as:

$$ACC = \frac{T_{positive} + T_{negative}}{T_{positive} + T_{negative} + F_{positive} + F_{negative}}. \quad (31)$$

Specificity is estimated to identify actual negative sentiments and is denoted in:

$$Spe = \frac{T_{negative}}{T_{negative} + F_{positive}}. \quad (32)$$

Sensitivity is the measure used for finding definite positives and is estimated by:

$$Sen = \frac{T_{positive}}{T_{positive} + F_{negative}}. \quad (33)$$

Precision is defined as the ratio between accurately classified outcomes and all classified positive values:

$$Pre = \frac{T_{positive}}{T_{positive} + F_{positive}}. \quad (34)$$

where true positives, false positives, true negatives and false negatives are specified as $T_{positive}$, $F_{positive}$, $T_{negative}$ and $F_{negative}$, respectively.

The Large Movie Review Dataset [36], the Datafiniti Product Database [37], and Amazon Reviews [38] are utilized for experimentation purposes and for testing the sentiment classification model developed.

The Large Movie Review Dataset includes 50,000 highly polarized movie reviews for used in testing and training processes. The number of positive and negative review files used in testing is 12,500, while 12,500 negative opinions and 12,500 positive review files are employed for the training process.

The Datafiniti Product Database contains 34,000 reviews of such Amazon products as Kindle, Fire TV stick and so on. It consists of the review text, product information and a rating assigned to each product.

Amazon Reviews contain reviews provided by several million Amazon customers. Star ratings, commonly used for sentiment analysis, as available as well. The database contains real business data on a reasonable scale.

The proposed A-RFATO-based DNFN is compared with other existing techniques, such as CNN [28], RNSA [11], AL-SSVAE [29], GRNN-SR [4], and RFATO-based deep RNN. In the comparison, different metrics are used, such as specificity, sensitivity, accuracy, and precision.

4.1. Analysis Based on the Large Movie Review Dataset

A comparative analysis of the A-RFATO-based DNFN, performed with the use of the Large Movie Review Dataset by changing training data percentages, is depicted in Fig. 4. Figure 4a shows the accuracy analysis performed by altering training data percentages. The accuracy score obtained by CNN is 0.647, by GRNN-SR is 0.777, by RNSA is 0.778, by AL-SSVAE is 0.778 and by RFATO-based deep RNN is 0.795, whereas the value achieved by the proposed A-RFATO-based DNFN is 0.813 with 70% of the training data considered. Additionally, improvement in the performance of Adaptive RFATO-based DNFN is 17.92% vs. CNN, 4.39% vs. GRNN-SR, 4.30% vs. RNSA, 4.26% vs. AL-SSVAE, and 2.23% vs. RFATO-based deep RNN techniques.

The sensitivity analysis performed by changing the training data percentage is illustrated in Fig. 4b. Sensitivity for CNN is 0.709, for GRNN-SR is 0.823, for RNSA is 0.824, for AL-SSVAE is 0.824, and for RFATO-based deep RNN is 0.845. The developed method achieves the score of 0.864. Improvement in the performance of the proposed technique versus existing methods, such as CNN, is 17.92%. It equals 4.75% vs. GRNN-SR, 4.68% vs. RNSA, 4.65% vs. AL-SSVAE and 2.19% vs. the RFATO-based deep RNN approach.

The specificity analysis is shown in Fig. 4c. In 70% of the training data, the proposed model achieves specificity of 0.746, while the score of CNN equals 0.55, GRNN-SR 0.7, RNSA 0.70, AL-SSVAE 0.701, and RFATO-based deep RNN 0.723. In addition, the improvement in performance of the proposed method is 26.35% compare with CNN, 6.27% with GRNN-SR, 6.14% with GRNN-SR, 6.10% AL-SSVAE, and 3.17% with RFATO-based deep RNN. Figure 4d presents an analysis of proposed method based on precision-related metrics, with the training data volumes altered. In 70% of training data, the A-RFATO-based DNFN achieves a precision score of 0.854, while the value of existing techniques equal 0.701 for CNN, 0.814 for GRNN-SR, 0.814 for RNSA, 0.815 for AL-SSVAE and 0.836 RFATO-based deep RNN.

A comparative analysis of the A-RFATO-based DNFN, using the Large Movie Review Dataset and with a varying feature size is shown in Fig. 5. The analysis is based on the comparison of accuracy with the changing feature size (Fig. 5a). The accuracy value achieved by CNN is 0.777, by GRNN-SR is 0.855, by RNSA is 0.856, by AL-SSVAE is 0.856, by RFATO-based deep RNN is 0.879, and the rate obtained by the proposed A-RFATO-based DNFN equaled 0.898, with the feature size 3. Furthermore, the developed sentiment classification model achieves a higher performance improvement rate of 13.41%, compared with the results of 4.74%, 4.61%, 4.59% and 2.12% attained by the existing sentiment classification approaches, respectively.

A comparative analysis of sensitivity with varying feature sizes is shown in Fig. 5b. The proposed method scores 0.935, while CNN achieves the score of 0.832, GRNN-SR of 0.888, RNSA of 0.888, AL-SSVAE of 0.889, and RFATO-based deep RNN methods of 0.910, for the feature size of approx. 3. The improvement in performance of the developed sentiment classification approach versus CNN is 11.99%, vs. GRNN-SR is 5.08%, vs. RNSA is 5.03%, vs. AL-SSVAE is 4.95%, and vs. RFATO-based deep RNN is 2.70%. Figure 5c presents the results of a specificity analysis based on a varying feature size. With a feature size of 3, the proposed method obtains the specificity rate of 0.845, thus outperforming other classification models: CNN achieves the score of 0.7, GRNN-SR of 0.798, RNSA of 0.799, AL-SSVAE of 0.800, and RFATO-based deep RNN of 0.822. The proposed technique achieved a higher performance improvement of 17.23% compared with CNN, 5.54% with GRNN-SR, 5.44% with RNSA, 5.30% with AL-SSVAE, and 2.74%, with RFATO-based deep RNN. Figure 5d illustrates the results of a precision analysis based on a varying feature size. With the feature size of 3, the pro-

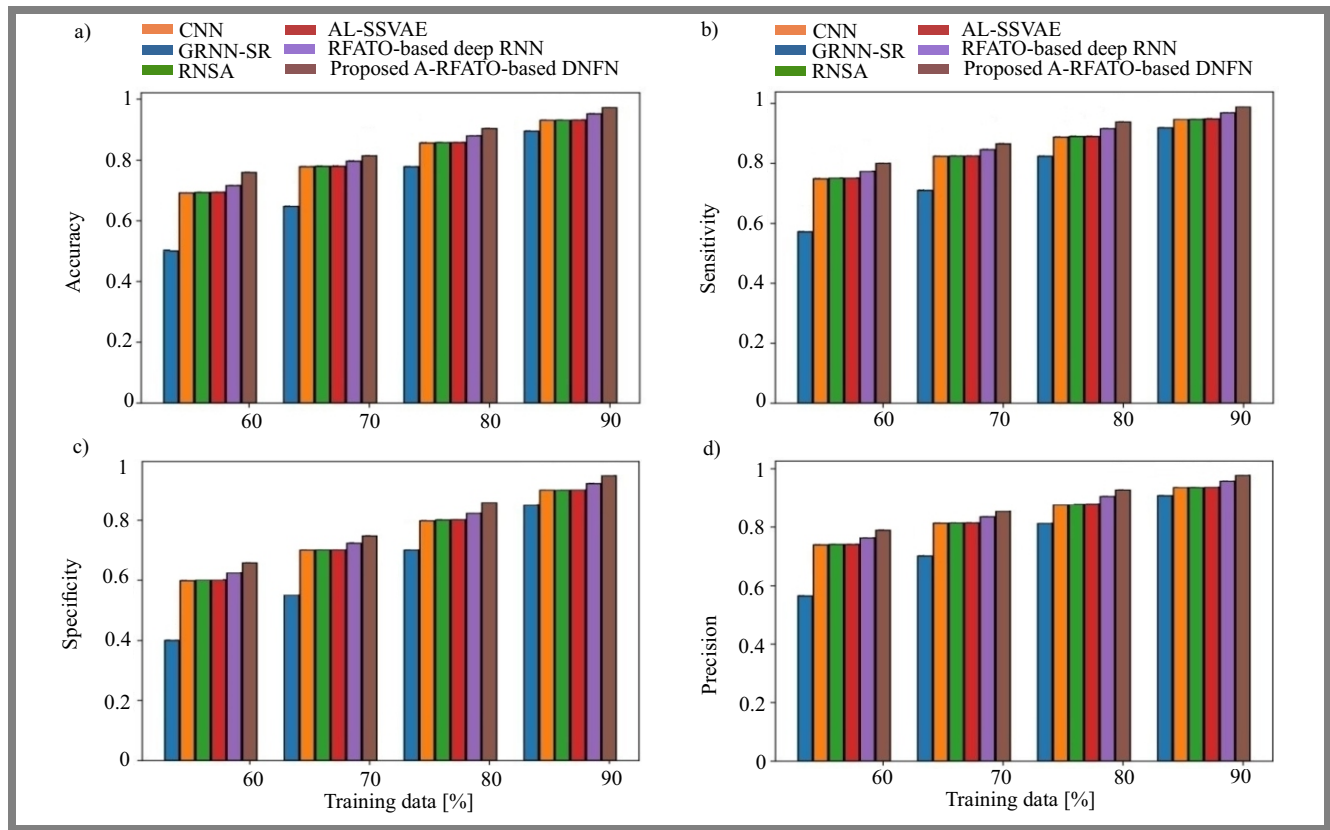


Fig. 4. Analysis of A-RFATO-based DNFN with the Large Movie Review Dataset depending on: a) accuracy, b) sensitivity, c) specificity, and d) precision with varying training data percentages.

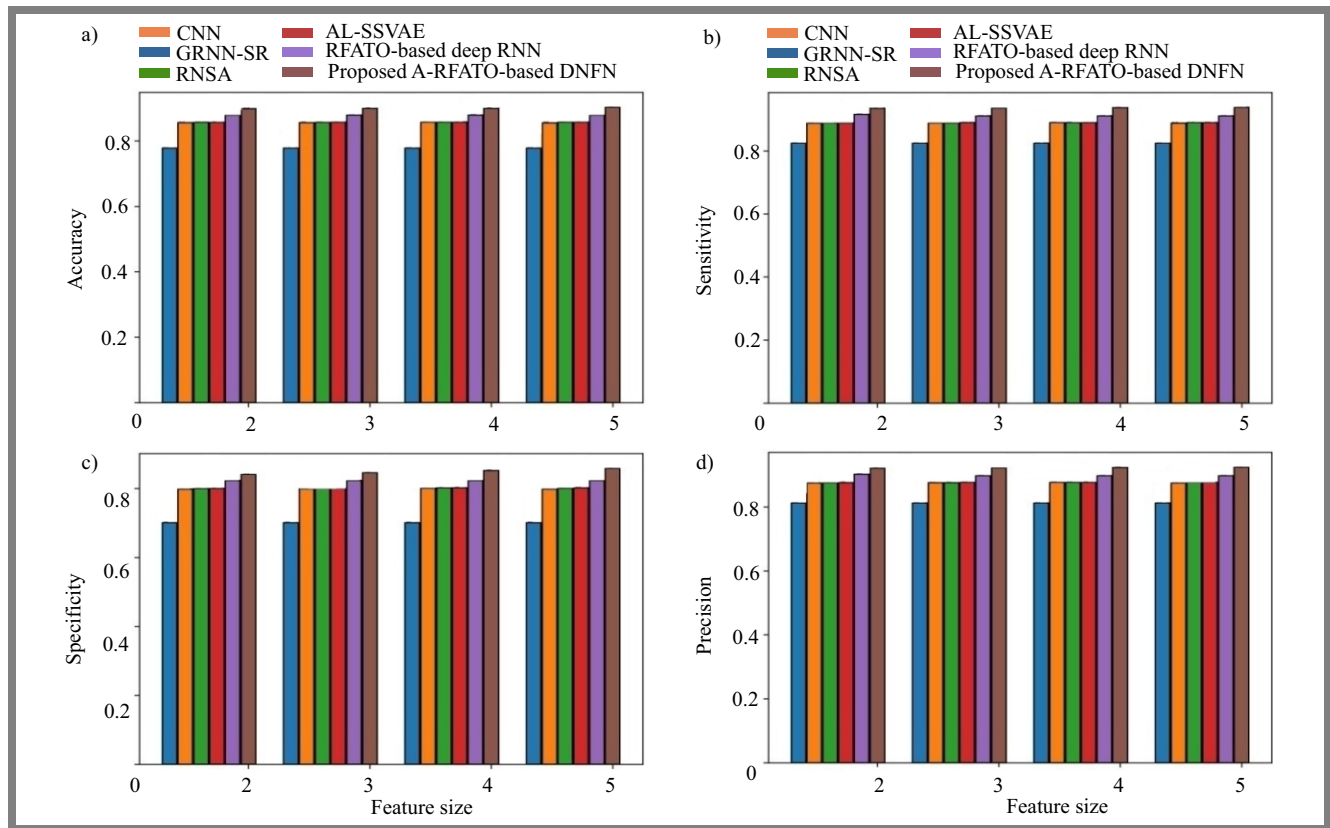


Fig. 5. Analysis of the proposed A-RFATO-based DNFN using the Large Movie Review Dataset with respect to: a) accuracy, b) sensitivity, c) specificity, and d) precision based on a varying feature size.

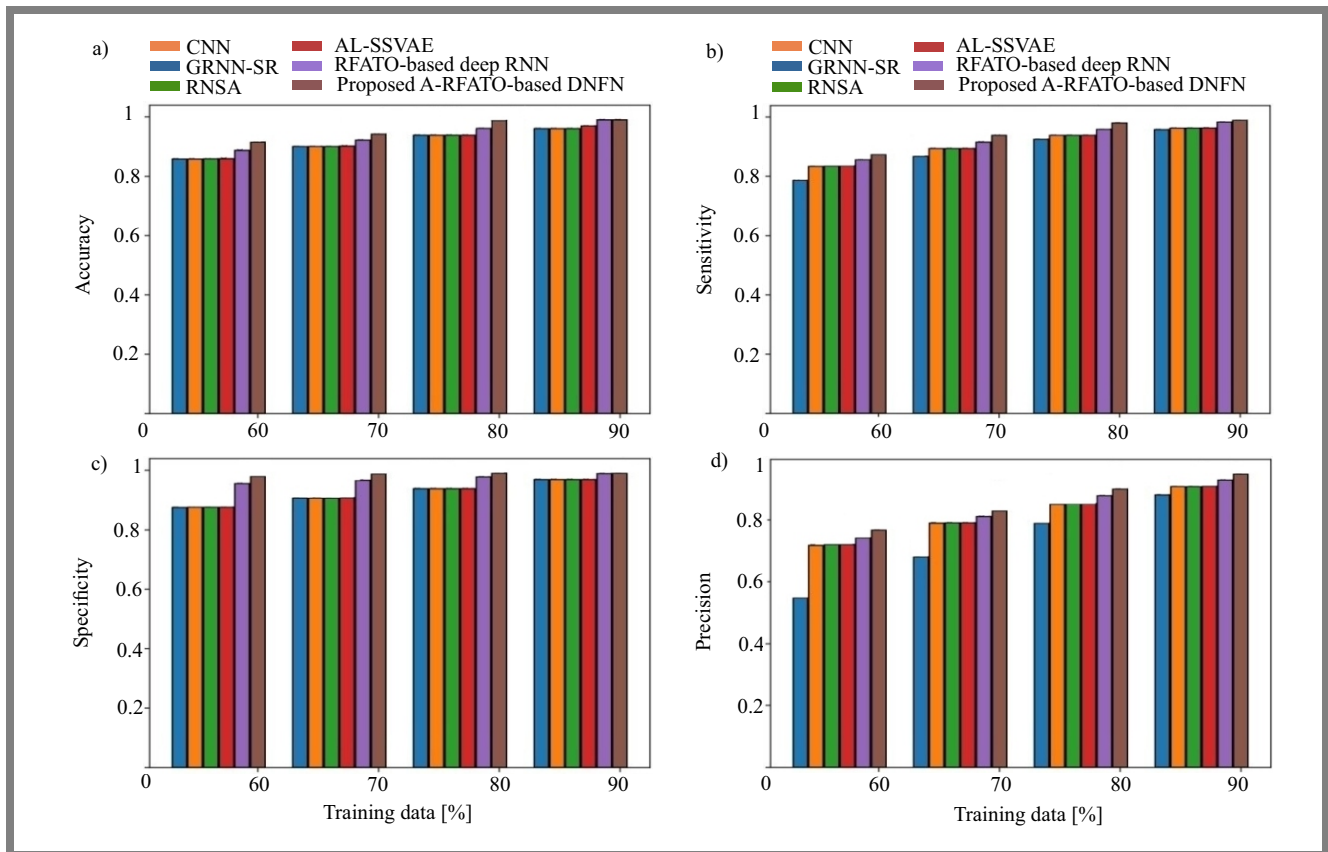


Fig. 6. Analysis of the proposed method using the Datafiniti Product Database, considering: a) accuracy, b) sensitivity, c) specificity, and d) precision with altering training data percentages.

posed method achieves the precision rating of 0.923 and also overperforms other sentiment classification models: CNN has the rating of 0.812, GRNN-SR of 0.877, RNSA of 0.877, AL-SSVAE of 0.877, and RFATO-based deep RNN of 0.898.

4.2. Analysis Based on the Datafiniti Product Database

A comparative analysis of the proposed A-RFATO-based DNFN using the Datafiniti Product Database with altering training data percentages is represented in Fig. 6. Figure 6a shows the accuracy analysis based on using different training data percentages. The existing techniques achieve the following accuracy scores: 0.937 for CNN, 0.937 for GRNN-SR, 0.938 for RNSA, 0.938 for AL-SSVAE, and 0.961 for RFATO-based deep RNN, whereas the proposed A-RFATO-based DNFN scores a ratio of 0.988 with 80% of training data. Moreover, A-RFATO-based DNFN shows an improvement in performance of 5.58% vs. CNN, 5.10% vs. GRNN-SR, 5.09% vs. RNSA, 5.06% vs. AL-SSVAE, and 2.68% vs. RFATO-based deep RNN.

Results of the sensitivity analysis with changing training data percentages are presented in Fig. 6b. The sensitivity score of CNN is 0.924, of GRNN-SR is 0.937, of RNSA is 0.937, of AL-SSVAE is 0.938, and of RFATO-based deep RNN is 0.958. The proposed method achieves the ratio of 0.979. A specificity analysis of the proposed method with altering training data percentages is presented in Fig. 6c. With 80%

of training data, the proposed model achieves the score of 0.99, where the existing techniques, such as CNN, RNSA and GRNN-SR, achieve the score of 0.937 each, AL-SSVAE reaches the value of 0.938, and RFATO-based deep RNN of 0.977. The improvement in performance of the proposed method versus CNN is 5.26%, vs. GRNN-SR is 5.26%, vs. RNSA is 5.25%, vs. AL-SSVAE is 5.23%, and vs. RFATO-based deep RNN is 1.23%. Figure 6d shows the precision measure for altering training data percentages. With 80% of training data, the developed model achieves the precision rating of 0.899, while CNN achieves the score of 0.790, RNSA of 0.851, GRNN-SR of 0.853, AL-SSVAE of 0.853, and RFATO-based deep RNN of 0.878.

Figure 7 shows a comparative analysis of the A-RFATO-based DNFN approach using the Datafiniti Product Database with a varying feature size. The comparative accuracy analysis with an altering feature size is shown in Fig. 7a. The values achieved by existing approaches are: CNN 0.937, GRNN-SR 0.937, RNSA 0.937, AL-SSVAE 0.938, RFATO-based deep RNN 0.961, with the A-RFATO-based deep neuro fuzzy network scoring 0.985 with the feature size of 4. The proposed sentiment classification model achieves a higher improvement in performance (4.89%, 4.88%, 4.87%, 4.82%, and 2.42%) compared with the present sentiment classification approaches, respectively.

Figure 7b shows a comparative sensitivity analysis based on a varying feature size. The proposed approach achieves

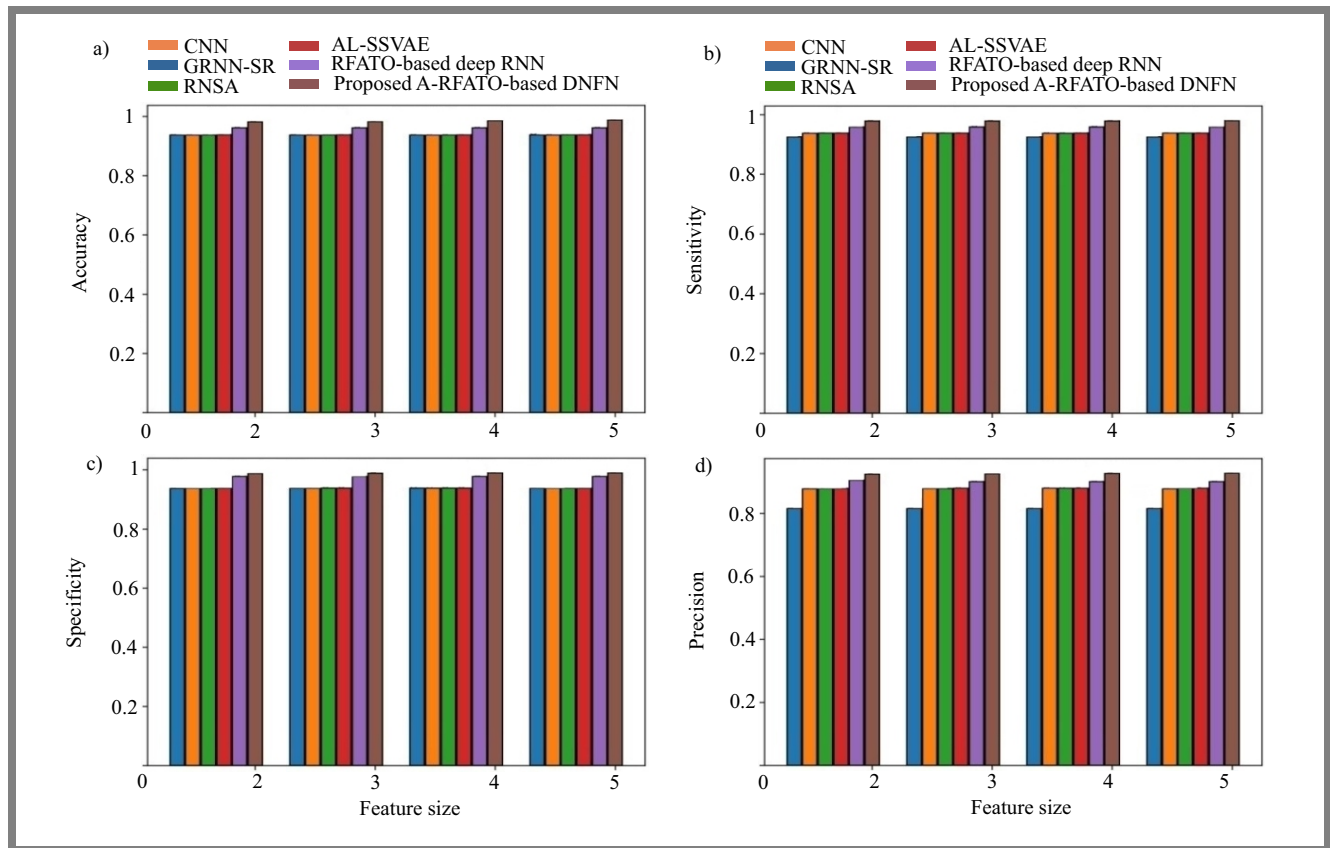


Fig. 7. Analysis of the proposed method using the Datafiniti Product Database, considering: a) accuracy, b) sensitivity, c) specificity, and d) precision with a varying feature size.

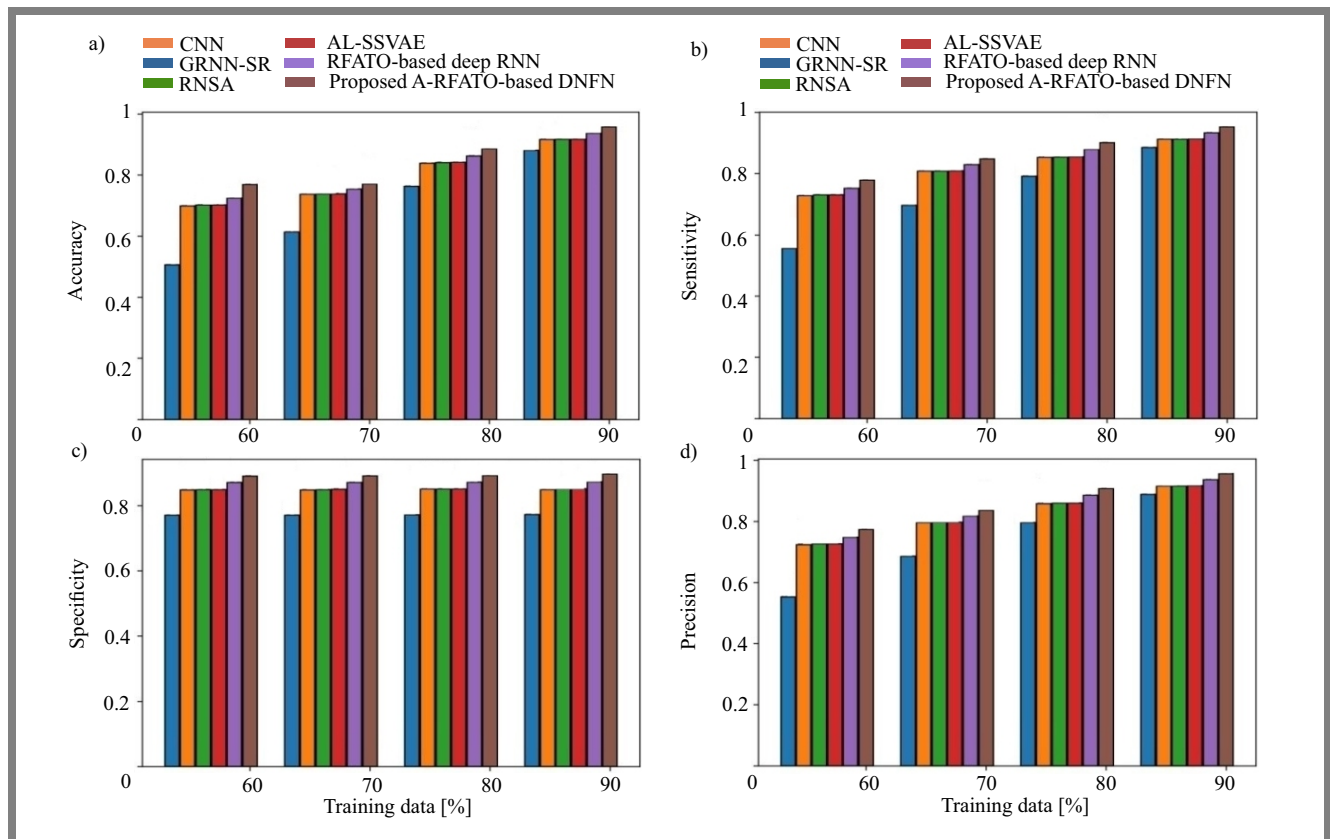


Fig. 8. Analysis of the proposed method based on Amazon Reviews, considering: a) accuracy, b) sensitivity, c) specificity, and d) precision, with varying training data percentages.

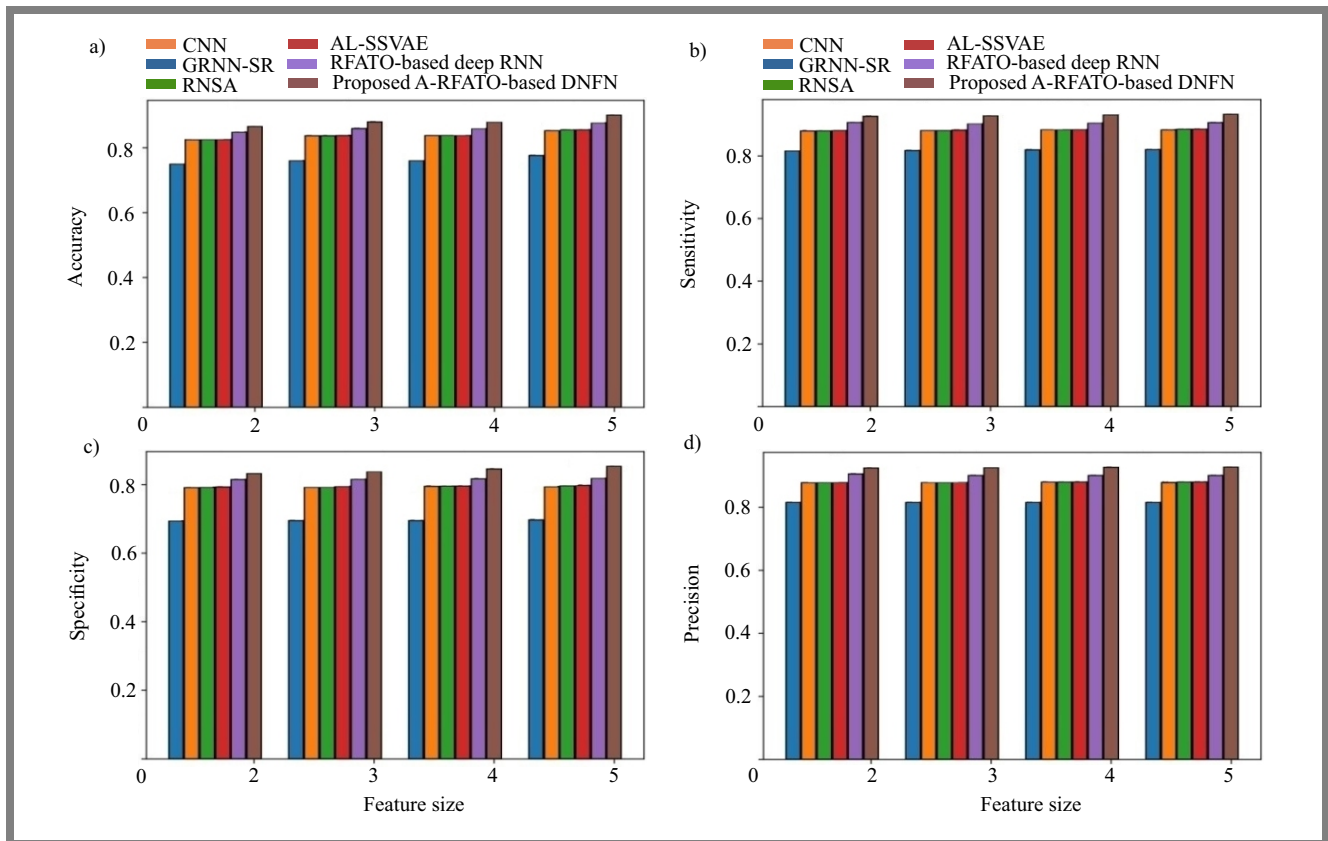


Fig. 9. Analysis of the proposed method using Amazon Reviews, considering: a) accuracy, b) sensitivity, c) specificity, and d) precision with varying feature sizes.

the value of 0.978, while CNN of 0.924, RNSA, GRNN-SR and AL-SSVAE of 0.937, and RFATO-based deep RNN of 0.958 (feature size 4). The improvement in performance of the proposed sentiment classification approach versus CNN is 5.51%, vs. GRNN-SR is 4.22%, vs. RNSA is 4.21%, vs. AL-SSVAE is 4.20%, and vs. RFATO-based deep RNN is 2.10%.

Figure 7c shows the results of a specificity analysis based on a varying feature size. For the feature size of 4, the proposed method reaches the specificity score of 0.99, CNN of 0.937, RNSA of 0.937, GRNN-SR of 0.937, AL-SSVAE of 0.938, and RFATO-based deep RNN of 0.977. The proposed technique achieves a performance improvement of 5.30% versus CNN, 5.29% vs. GRNN-SR, 5.28% vs. RNSA, 5.23% vs. AL-SSVAE, and 1.23% vs. RFATO-based deep RNN. Figure 7d shows the results of a precision analysis based on the feature size. For the feature size of 4, the proposed method reaches the precision score of 0.928, while CNN achieves the rating of 0.815, RNSA of 0.878, GRNN-SR of 0.880, AL-SSVAE of 0.881, and RFATO-based deep RNN of 0.901.

4.3. Analysis Based on the Amazon Reviews Dataset

Analysis of A-RFATO-based DNFN using the Amazon Reviews with an altering training data percentage is shown in Fig. 8. Figure 8a shows the results of the accuracy analysis with different training data percentages. The existing tech-

niques achieve the following accuracy ratings: 0.763 for CNN, 0.839 for GRNN-SR, 0.841 for RNSA, 0.842 for AL-SSVAE, and 0.863 for RFATO-based deep RNN. The proposed A-RFATO-based DNFN achieved a score of 0.886 with 80% of training data. Figure 8b shows the results of the sensitivity analysis with changing training data percentages. Sensitivity of the existing sentiment classification methods are as follows: CNN 0.696, GRNN-SR 0.808, RNSA 0.809, AL-SSVAE 0.809, and RFATO-based deep RNN 0.830. The developed method obtained the score of 0.848 for 70% of training data.

An analysis of specificity-related performance of the proposed method, with varying training data percentages, is presented in Fig. 8c. With 80% of training data used, the proposed model achieves specificity of 0.893, while the existing techniques achieved the following results: CNN 0.772, RNSA 0.851, GRNN-SR 0.851, AL-SSVAE 0.851, and RFATO-based deep RNN 0.872. Analysis of the precision of the proposed method is presented in Fig. 8d. Using 80% of training data, the proposed model scores 0.907, while CNN achieves the rating of 0.796, RNSA of 0.858, GRNN-SR of 0.860, AL-SSVAE of 0.860, and RFATO-based deep RNN of 0.886.

Figure 9 illustrates results of the analysis of A-RFATO-based DNFN using Amazon Reviews, with varying feature sizes. The level of accuracy vs. the altering feature size is presented in Fig. 9a. The accuracy scores achieved by exist-

Tab. 1. Comparative study for the Large Movie Review Dataset.

Analysis	Metrics	GRNN-SR	CNN	RNSA	RFATO-based deep RNN	AL-SSVAE	A-RFATO-based deep neuro fuzzy network
Based on training data	Accuracy	0.930	0.894	0.931	0.951	0.931	0.972
	Sensitivity	0.947	0.918	0.947	0.968	0.947	0.988
	Specificity	0.899	0.85	0.900	0.921	0.900	0.947
	Precision	0.936	0.908	0.936	0.956	0.936	0.977
Based on feature size	Accuracy	0.855	0.777	0.857	0.879	0.857	0.903
	Sensitivity	0.887	0.823	0.889	0.910	0.889	0.937
	Specificity	0.800	0.797	0.801	0.857	0.822	0.797
	Precision	0.875	0.812	0.877	0.898	0.877	0.925

Tab. 2. Comparative summary for the Datafiniti Product Database.

Analysis	Metrics	GRNN-SR	CNN	RNSA	RFATO-based deep RNN	AL-SSVAE	A-RFATO-based deep neuro fuzzy network
Based on training data	Accuracy	0.960	0.960	0.960	0.989	0.970	0.990
	Sensitivity	0.962	0.957	0.962	0.982	0.962	0.988
	Specificity	0.968	0.968	0.968	0.988	0.969	0.990
	Precision	0.908	0.881	0.909	0.928	0.909	0.948
Based on feature size	Accuracy	0.937	0.937	0.938	0.961	0.938	0.988
	Sensitivity	0.937	0.924	0.937	0.958	0.938	0.979
	Specificity	0.937	0.937	0.937	0.977	0.938	0.990
	Precision	0.878	0.815	0.880	0.901	0.881	0.928

ing approaches are: CNN 0.760, GRNN-SR 0.838, RNSA 0.838, AL-SSVAE 0.838, RFATO-based deep RNN 0.859, and A-RFATO-based deep neuro fuzzy network 0.879, with the feature size of 4. Figure 9b shows a similar analysis focusing on sensitivity. Sensitivity of the proposed approach is 0.931, while that of CNN equals 0.819, RNSA 0.884, GRNN-SR 0.884, AL-SSVAE 0.937, and RFATO-based deep RNN 0.905 (feature size 4).

Figure 9c shows the specificity analysis. The developed method obtained a specificity score of 0.845, while CNN of 0.695, RNSA of 0.795, GRNN-SR of 0.796, AL-SSVAE of 0.796, and RFATO-based deep RNN of 0.817. Part “d” of Fig. 9 shows a precision analysis. The developed method scored 0.926, CNN 0.814, RNSA 0.879, GRNN-SR 0.880, AL-SSVAE 0.880, and RFATO-based deep RNN 0.901.

4.4. Comparative Summary

A summary of the performance of different techniques relied upon in the classification of sentiment using the Large Movie Review Dataset is presented in Tab. 1. The proposed method obtains better performance with 90% of training data and the feature size of 5, and achieves the maximum values of sensitivity, accuracy and specificity.

Comparison of the performance of different techniques used for sentiment grade classification based on the Datafiniti Product Database is presented in Tab. 2. The accuracy score achieved by the developed sentiment classification approach

equals 0.99 (feature size 5), and its sensitivity ratio is 0.988. Likewise, the developed A-RFATO-based deep neuro fuzzy network obtains a specificity ratio of 0.99, which is better than that of its competitors. The proposed method obtains the best accuracy and sensitivity levels of 0.988 and 0.979, respectively, with the feature size of 5, when the percentage of training data considered is 90%. Hence, the table clearly indicates that the proposed A-RFATO-based DNFN approach achieves better accuracy, specificity, sensitivity, and precision scores than the existing methods.

The comparison of the techniques used in sentiment grade classification based on the Amazon Reviews dataset is shown in Tab. 3. The maximum accuracy, sensitivity, specificity, and precision scores attained by the A-RFATO-based deep neuro fuzzy network are 0.958, 0.953, 0.897, and 0.956, respectively.

5. Conclusion

This paper describes a technique for sentiment grade classification, known as A-RFATO-based DNFN. The classification process involves four stages: pre-processing stage, extraction of features, and feature fusion along with the classification of sentiment. The proposed A-RFATO scheme is designed by integrating the FAT technique and ROA with the adaptive concept. As a result, the developed sentiment grade classification technique allows to achieve the maximum values of

Tab. 3. Comparative data for the Amazon Reviews dataset.

Analysis	Metrics	GRNN-SR	CNN	RNSA	RFATO-based deep RNN	AL-SSVAE	A-RFATO-based deep neuro fuzzy network
Based on training data	Accuracy	0.917	0.881	0.917	0.938	0.917	0.958
	Sensitivity	0.913	0.886	0.913	0.933	0.914	0.953
	Specificity	0.849	0.772	0.852	0.873	0.852	0.897
	Precision	0.916	0.889	0.916	0.936	0.916	0.956
Based on feature size	Accuracy	0.853	0.775	0.855	0.877	0.855	0.901
	Sensitivity	0.883	0.819	0.885	0.906	0.885	0.933
	Specificity	0.793	0.697	0.797	0.819	0.797	0.854
	Precision	0.878	0.814	0.879	0.901	0.880	0.927

accuracy (0.99), sensitivity (0.988) and specificity (0.99). Future work related to this field of research will focus on optimizing the technique in order to enhance the classification process and tune it for data involving sarcasm.

References

- [1] B. Liu, *Sentiment Analysis and Opinion Mining*. Springer Cham, serie *Synthesis Lectures on Human Language Technologies*, 2012 (ISBN: 9783031010170, <https://doi.org/10.1007/978-3-031-02145-9>).
- [2] F. Chen, Z. Yuan, and Y. Huang, "Multi-source data fusion for aspect-level sentiment classification", *Knowledge-Based Systems*, vol. 187, pp. 104831, 2020 (<https://doi.org/10.1016/j.knosys.2019.07.002>).
- [3] Y. Wang, M. Huang, Z. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification", *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 606–615, 2016 (<https://doi.org/10.18653/v1/D16-1058>).
- [4] C. Chen, R. Zhuo, and J. Ren, "Gated recurrent neural network with sentimental relations for sentiment classification", *Information Sciences*, vol. 502, pp. 268–278, 2019 (<https://doi.org/10.1016/j.ins.2019.06.050>).
- [5] X. Li X and W. Lam, "Deep multi-task learning for aspect term extraction with memory interaction", *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2886–2892, 2017 (<https://doi.org/10.18653/v1/D17-1310>).
- [6] E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment analysis is a big suitcase", *IEEE Intelligent Systems*, vol. 2, no. 6, pp. 74–80, 2017 (<https://doi.org/10.1109/MIS.2017.4531228>).
- [7] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media", *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, 2019 (<https://doi.org/10.1007/s10115-018-1236-4>).
- [8] M.U. Salur and I. Aydin, "A novel hybrid deep learning model for sentiment classification", *IEEE Access*, vol. 8, pp. 58080–93, 2020 (<https://doi.org/10.1109/ACCESS.2020.2982538>).
- [9] Z. Wang and Z. Lin, "Optimal feature selection for learning-based algorithms for sentiment classification", *Cognitive Computation*, vol. 12, no. 1, pp. 238–248, 2020 (<https://doi.org/10.1007/s12559-019-09669-5>).
- [10] C. Zhang, Q. Li, and D. Song, "Aspect-based sentiment classification with aspect-specific graph convolutional networks", *Proceedings of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th International Joint Conf. on Natural Language Processing (EMNLP-IJCNLP)*, pp. 4568–4578, 2019 (<https://doi.org/10.18653/v1/D19-1464>).
- [11] A. Abdi, S.M. Shamsuddin, S. Hasan, and J. Piran, "Deep learning-based sentiment classification of evaluative text based on multi-feature fusion", *Information Processing & Management*, vol. 56, no. 4, pp. 1245–1259, 2019 (<https://doi.org/10.1016/j.ipm.2019.02.018>).
- [12] Y. Zhang, Z. Zhang, D. Miao, and J. Wang, "Three-way enhanced convolutional neural networks for sentence-level sentiment classification", *Information Sciences*, vol. 477, pp. 55–64, 2019 (<https://doi.org/10.1016/j.ins.2018.10.030>).
- [13] J. Serrano-Guerrero, J.A. Olivas, F.P. Romero, and E. Herrera-Viedma, "Sentiment analysis: A review and comparative analysis of web services", *Information Sciences*, vol. 311, pp. 18–38, 2015 (<https://doi.org/10.1016/j.ins.2015.03.040>).
- [14] S. Zhang, X. Xu, Y. Pang, and J. Han, "Multi-layer attention based CNN for target-dependent sentiment classification", *Neural Processing Letters*, vol. 51, no. 3, pp. 2089–2103, 2020 (<https://doi.org/10.1007/s11063-019-10017-9>).
- [15] C. Zhao, S. Wang, and D. Li, "Multi-source domain adaptation with joint learning for cross-domain sentiment classification", *Knowledge-Based Systems*, vol. 191, pp. 105254, 2020 (<https://doi.org/10.1016/j.knosys.2019.105254>).
- [16] H. Chen, M. Sun, C. Tu, Y. Lin, and Z. Liu, "Neural sentiment classification with user and product attention", *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 1650–1659, 2016 (<https://doi.org/10.18653/v1/D16-1171>).
- [17] M. Gangappa, K.C. Mai C, and P. Sammulal, "Enhanced crow search optimization algorithm and hybrid NN-CNN classifiers for classification of land cover images", *Multimedia Research*, vol. 2, no. 3, pp. 12–22, 2019 (<https://doi.org/10.46253/j.mr.v2i3.a2>).
- [18] S. Vasamsetti and C. Santhirani, "Hybrid particle swarm optimization-deep neural network model for speaker recognition", *Multimedia Research*, vol. 3, no. 1, pp. 1–10, 2020 (<https://doi.org/10.46253/j.mr.v3i1.a1>).
- [19] G. Parente, T. Gargano, G. Ruggeri, M. Maffi, S. D'Antonio, E. Sacchet, and M. Lima, "Anastomotic stricture definition after esophageal atresia repair: role of endoscopic stricture index", *Journal of Surgical Research*, vol. 257, pp. 572–578, 2021 (<https://doi.org/10.1016/j.jss.2020.08.035>).
- [20] G. Parente, T. Gargano, S. Pavia, C. Cordola, M. Vastano, Baccelli, G. Gallotta, L. Bruni, A. Corvaglia, and M. Lima, "Pyelonephritis in pediatric uropathic patients: Differences from community-acquired ones and therapeutic protocol considerations. A 10-year single-center retrospective study", *Children*, vol. 8, no. 6, 2021 (DOI: 10.3390/children8060436).
- [21] R.K. Bakshi, N. Kaur, R. Kaur, and G. Kaur, "Opinion mining and sentiment analysis", *Proceedings of 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 452–455, 2016.
- [22] D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification", *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17)*, pp. 4068–4074, 2017 (<https://www.ijcai.org/proceedings/2017/0568.pdf>).
- [23] R. John Martin, "EEG Feature Engineering Methods-A Comprehensive Review", *Multimedia Research*, vol. 5, no. 2, 2022 (URL: <https://publisher.resbee.org/mr/archive/v5i2/a5.html>).
- [24] S.A.J. Al Raisi, "A review on congestion management methodologies and its applications", *Journal of Computational Mechanics, Power System and Control*, vol. 3, no. 3, 2020 (<https://doi.org/10.46253/jcmps.v3i3.a3>).

- [25] M. Adil, R. Madani, S. Tavakkol, and A. Davoudi, "A First-Order Numerical Algorithm without Matrix Operations", arXiv preprint arXiv:2203.05027, 2022 (<https://arxiv.org/pdf/2203.05027>).
- [26] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent twitter sentiment classification", *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 49–54, 2014 (<https://doi.org/10.3115/v1/P14-2009>).
- [27] K.S. Tai, R. Socher, C.D. Manning, "Improved semantic representations from tree-structured long short-term memory networks", *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conf. on Natural Language Processing*, vol. 1, pp. 1556–1566, 2015 (<https://doi.org/10.3115/v1/P15-1150>).
- [28] H. Kim and Y.S. Jeong, "Sentiment classification using convolutional neural networks", *Applied Sciences*, vol. 9, no. 11, pp. 2347, 2019 (<https://doi.org/10.3390/app9112347>).
- [29] X. Fu, Y. Wei, F. Xu, T. Wang, Y. Lu, J. Li, and J.Z. Huang, "Semi-supervised aspect-level sentiment classification model based on variational autoencoder", *Knowledge-Based Systems*, vol. 171, pp. 81–92, 2019 (<https://doi.org/10.1016/j.knosys.2019.02.008>).
- [30] P. Zhao, L. Hou, and O. Wu, "Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification", *Knowledge-Based Systems*, vol. 193, pp. 105443, 2020 (<https://doi.org/10.1016/j.knosys.2019.105443>).
- [31] K. Zhang, M. Jiao, X. Chen, Z. Wang, B. Liu, and L. Liu, "SC-BiCapsNet: a sentiment classification model based on bi-channel capsule network", *IEEE Access*, vol. 7, pp. 171801–171713, 2019 (<https://doi.org/10.1109/ACCESS.2019.2953502>).
- [32] N. Jin, J. Wu, X. Ma, K. Yan, Y. Mo, "Multi-task learning model based on multi-scale CNN and LSTM for sentiment classification", *IEEE Access*, vol. 8, pp. 77060–77072, 2020 (<https://doi.org/10.1109/ACCESS.2020.2989428>).
- [33] B. Ohana and B. Tierney, "Sentiment classification of reviews using SentiWordNet", *9th IT&T Conference*, 2009 (<https://doi.org/10.21427/D77S56>).
- [34] N.M.K. Saeed, N.A. Helal, N.L. Badr, and T.F. Gharib, "The impact of spam reviews on feature-based sentiment analysis", *Proceedings of 13th International Conference on Computer Engineering and Systems (ICCES)*, 2018 (<https://doi.org/10.1109/ICCES.2018.8639343>).
- [35] S. Javaid, M. Abdullah, N. Javaid, T. Sultana, J. Ahmed, and N.A. Sattar, "Towards buildings energy management: using seasonal schedules under time of use pricing tariff via deep neuro-fuzzy optimizer", *Proceedings of 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 2019, pp. 1594–1599 (<https://doi.org/10.1109/IWCMC.2019.8766673>).
- [36] –, Large movie review dataset taken from (<http://ai.stanford.edu/~amaas/data/sentiment>), accessed on Dec. 2020.
- [37] –, Consumer Reviews of Amazon Products dataset taken from (<https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products>), accessed on Dec. 2020.
- [38] –, Amazon Reviews for Sentiment Analysis dataset taken from, 2022 (<https://www.kaggle.com/datasets/bittlingmayer/amazonreviews>), accessed on Nov. 2022

Sireesha Jasti

Research Scholar

 <https://orcid.org/0000-0002-4755-5033>

E-mail: sireeshajasti14@gmail.com

Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India

G.V.S. Raj Kumar, Prof.

E-mail: gganapav@gitam.edu

Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India

Improved Association Rule Mining-Based Data Sanitization for Privacy Preservation Model in Cloud

Rajkumar Patil and Gottumukkala HimaBindu

Department of Computer Science Engineering, GITAM (Deemed to be) University, Hyderabad, Telangana, India

<https://doi.org/10.26636/jtit.2023.166922>

Abstract — Data security in cloud services is achieved by imposing a broad range of privacy settings and restrictions. However, the different security techniques used fail to eliminate the hazard of serious data leakage, information loss and other vulnerabilities. Therefore, better security policy requirements are necessary to ensure acceptable data protection levels in the cloud. The two procedures presented in this paper are intended to build a new cloud data security method. Here, sensitive data stored in big datasets is protected from abuse via the data sanitization procedure relying on an improved apriori approach to clean the data. The main objective in this case is to generate a key using an optimization technique known as Corona-integrated Archimedes Optimization with Tent Map Estimation (CIAO-TME). Such a technique deals with both restoration and sanitization of data. The problem of optimizing the data preservation ratio (IPR), the hiding ratio (HR), and the degree of modification (DOM) is formulated and researched as well.

Keywords — data security, improved apriori, modified data restoration, sanitization.

1. Introduction

Cloud computing (CC) creates a massive virtualized data resource pool with various services connecting a huge number of resources [1]–[3] to achieve the required levels of portability and reliability. The three fundamental service delivery models include the following: infrastructure as a service, platform as a service, and software as a service, and the four development trends observed in the world of CC are public cloud, private cloud, cloud platform, public clouds, and virtual private cloud [4]–[6].

IaaS is a service-managed, computer hardware and fixed network provisioning concept with significant expansion capabilities. PaaS is a service paradigm that uses the middleware of the service model to offer integrated development environments, frameworks, applications, and development tools [7]–[9], while SaaS is a category of remote computing services. All three technologies are used in cloud computing. However, because users are not aware of the sources or owners of specific resources in such a structure, it is more challenging to protect such resources and services against attacks. The task is even more complicated, because many companies owning and running community groups, as well as outside stakeholders, operate in the capacity of cloud administrators. To benefit from the advantages of multiple cloud deploy-

Tab. 1. List of abbreviations used in this paper.

Abbreviation	Description
AOA	Archimedes optimization algorithm
ARM	Association rule mining
BMO	Blue monkey optimization
BES	Bald eagle search optimization
CHIO	Corona virus herd immunity algorithm
CIAO-TME	Corona integrated Archimedes optimization with tent map estimation
CNN	Convolutional neural network
CPA	Chosen plaintext attack
CP-ABE	Cipher text policy attribute-based encryption
CHIO	Corona virus herd immunity optimization
CC	Cloud computing
DOM	Degree of modification
EHR	Electronic health record
IDS	Intrusion detection system
IPR	Information preservation ratio
HLPN	High-level Petri nets
HBA	Honey badger algorithm
HPA	Honey pot algorithm
HR	Hiding ratio
IF	Impact factor
IM	Information masking
IPR	Information preservation ratio
GWT	Gabor wavelet transform
GMGW	Genetically modified glowworm swarm optimization
GLCM	Grey level co-occurrence matrix
JA	Jaya algorithm
J-SSO	Jaya-based shark smell optimization
KPA	Known plaintext attack
KP-ABE	Key policy attribute-based encryption
MIQP	Mixed-integer quadratic programming
PRE	Proxy re-encryption
PPX-AC	Privacy-preserving XACML-based access control model
PCA	Principal component analysis
SE	Searchable encryption
SMC	Secure multiparty computation
SSO	Shark smell optimization

ment methodologies, a hybrid cloud concept is introduced which integrates two or more clouds. Private virtual cloud is the term used to describe the common pool of resources in a cloud system [10]–[12].

Security systems, key distribution, encrypting, access, identity authentication, audit scheduling, as well as human and physical access control are some of the security issues affecting cloud data [13], [14], as defined by several privacy-protection strategies [15]–[17]. Improved data security levels have been achieved in the cloud through the development of a privacy-conscious access control system. Researchers have proposed a unique method for combining spectral band handprint images, depending upon the complex dual-tree transform as well as a feature extraction and minimization method based on the Gabor wavelet transform (GWT) and principal component analysis (PCA).

With such an environment taken into consideration, this article offers the following contributions:

- it employs a modified apriori approach to sanitize system information,
- it proposed the CIAO-TME method to generate the best key while adhering to the degree of modification (DOM), information preservation ratio (IPR), and hiding ratio (HR) requirements.

The rest of the paper is as organized follows. Section 2 evaluates the existing literature. The suggested model and its features are described in Section 3. Section 4 covers key generation processes and the data sanitization method. Results and discussion are given while Section 5 while conclusions are presented in Section 6.

2. Literature Review

Danish *et al.* [18] formulated their approach to privacy protection in the cloud environment by employing artificial intelligence (AI). The authors stated that AI abilities were helping companies achieve greater productivity in the corporate cloud environment. Data normalization and recovery were the two key steps of the recommended privacy-preserving system. The extraction of many co-functions, including such factors as DOM, HR, and IPR, leads to the generation of an optimal key. The study demonstrated the effectiveness of the suggested model in improving cloud security when compared to other methods.

Avijit and Radha in [19] employed the honeypot algorithm for data protection or IDS, which was a good strategy for predictions and privacy protection. The dataset was initially standardized using the normalization approach, a process which involves replacing missing values and removing unnecessary data. Following that, unique features were extracted and the best models were chosen using the GLCM algorithm. Predicting the target was done using a unique CNN classifier that offers high attack detection accuracy levels. The developed algorithm was used to protect information from infiltration and other assaults. In addition, a cryptographic mechanism was utilized to ensure the required secrecy protection, while encryption was performed using HPA. If the data

holder requests a specific file, the cloud server generates a key and verifies it by interacting with the user for authentication purposes. The performance of the solution was evaluated and was compared with that of other, existing methodologies to demonstrate efficacy of the proposed scheme.

Tehsin *et al.* [20] investigated the privacy-preserving authorization paradigm for the cloud and privacy-preserving strategies for cloud-based EHRs using defined taxonomy. Inner login control and outer security in outsourced system design for hybrid cloud have been formulated and then the PPX-AC algorithm was developed, combining fine login control with the multifunctional use of EHRs and a cutting-edge privacy mechanism. Using HLPN, the authors confirmed the efficiency of the proposed PPX-AC by invalidating known privacy threats. Furthermore, the described model demonstrates its efficacy and multifunctional application possibilities.

Tian *et al.* [21] proposed an IM-based methodology for using MIQP to solve the energy management problem. The viability and efficiency of using IM in MIQP were demonstrated by incorporating cloud-edge architecture. To better fit real applications, the general criteria of the MIQP IM were expanded in terms of security and implementation cost.

Luis *et al.* in [22] presented smart CAMPP data to achieve cloud authorization. Format-conserving encoding methods were used to outsource them discreetly. Furthermore, the observations demonstrated the applicability of the proposed technique, allowing to expect high accuracy levels. In contrast to a method that does not improve security, the authors' proposal has no substantial influence on encryption.

Pan in [23] presented various cloud privacy security vulnerabilities and then proposed a complete privacy security prevention architecture. The characteristics of several techniques were also compared, including network access tools, CP-ABE, KP-ABE, the fine-grain, polynomial number of authority, dismissal mechanism, the detect mechanism, PRE, various leveled encryption, and a mixture of other methods.

Chen *et al.* in [24] designed a lightweight encryption system that maintains an acceptable usefulness model while ensuring evidential privacy preservation. Using the specified prototype system, the recommended approach was deemed secure against a sincere but inquisitive host and a catastrophic collision. The effectiveness of the method was examined and compared to similar solutions using the MNIST and UCI human action recognition database. This strategy decreased the runtime by 20% and the communicated cipher text length by 85%, on average, while maintaining the accuracy of competitive SMC methods.

Yong *et al.* [14] presented a blockchain-based EHR sharing mechanism ensuring both secure and private features by employing encryption algorithms. Additionally, evidence of permission was intended to serve as the consensus protocol for consortium blockchains in order to guarantee the software's reliability. According to a study, the proposed protocol meets the security goals and has a good computational efficiency.

Table 2 summarizes the research on recent cloud data security methods.

Tab. 2. Summary on research in existing papers.

Paper	Proposed methods	Features	Drawbacks
Danish <i>et al.</i> [18]	J-SSO algorithm	It is an algorithm with fewer steps and strong global findings	When it comes to ensuring the security of every database, the strategies OI-CSA and BS-WOA strategies offer low convergence performance
Avijit and Radha [19]	Honeypot method	Due to efficient implementation of the system, the security rate of online services has been improved	It is only capable of detecting direct assaults
Tehsin <i>et al.</i> [20]	PPX-AC model	Enables fine-grained access control while maintaining privacy	When data is sent between parties, it is essential to make sure that it is secure, as the parties involved were unaware of the information shared between both original parties
Tian <i>et al.</i> [21]	MIQP algorithm	Is an optimum solution with good accuracy, while lowering the computational cost	Preserving privacy while decreasing computing costs to the highest possible extend is hard to achieve
Luis <i>et al.</i> [22]	Smart CAMPP algorithm	Builds a CA approach for smartphones	No positive feedback on motion sensors leads to challenging results
Sun [23]	CP-ABE, KP-ABE	CP-ABE allows data owners to exchange encryption with authenticated persons through cloud storage while keeping access control settings hidden. Senders can encrypt communications using KP-ABE with a characteristics set	The degree of privacy protection provided by sharing algorithms concerned with users' identities has to be improved
Chen <i>et al.</i> [24]	HFWP	It improves training efficiency and communication overhead while providing excellent privacy protection in a short span of time	It cannot gather raw data from consumers
Yong <i>et al.</i> [14]	EHR method	Data sharing method among medical institutions were identified	The developer fails to provide timely updates

Distributed locations with cloud processing infrastructures and mass data storage create privacy concerns. For instance, Google's cloud servers are spread out all over the world, including seven sites in the Americas, two in Asia, and three in Europe. Additionally, customers must be aware where the cloud hosting is located, because privacy defining laws vary in many countries. This shows the importance of cyber security aspects in cloud computing.

Many algorithms were introduced to solve the privacy preservation problem, but there is no optimal solution yet. For example the J-SSO algorithm [18] offers a poor convergence rate that causes failures in this sort of applications. Similarly, such algorithms as the honeypot method [19] are capable only of detecting direct assaults, while the smart CAMPP algorithm [23] often creates challenges, as no previous efforts has focused on motion sensors. Moreover, such encryption standards as KP-ABE and CP-ABE also need some improvement in the degree of privacy provided. Therefore, there is still a need of researching the advanced privacy preservation models.

3. Proposed Method Concerning Data Security in Cloud

This work proposes two procedures that attempt to develop a new cloud data security method: data sanitization and

restoration process. The data sanitization process sustains the security of sensitive data in the cloud by concealing it from unauthorized users and preventing it from being accessed. In this scenario, data is sanitized using an upgraded version of the a priori algorithm, then data restoration is incorporated for restoring or recovering the sanitized data. In both processes, key generation plays a very important role and should be performed optimally to guarantee data safety. Here, the recommended CIAO-TME technique is used to identify the best key for both data restoration and sanitization stages. Considering this to be an optimization problem, DOM, IPR, and HR-related objectives are used for obtaining the key.

Figure 1 shows the diagram of the concept proposed for ensuring preservation of data security in the cloud.

3.1. Data Sanitization

The association rule mining (ARM) approach is enhanced by using the apriori algorithm containing the following three steps:

- 1) Scan transaction database once and get the sampling method for each item. During the sampling procedure, the algorithm picks a random sample S from the database da and then searches for frequent item sets S . This can be reduced by lowering the so-called min-support. The

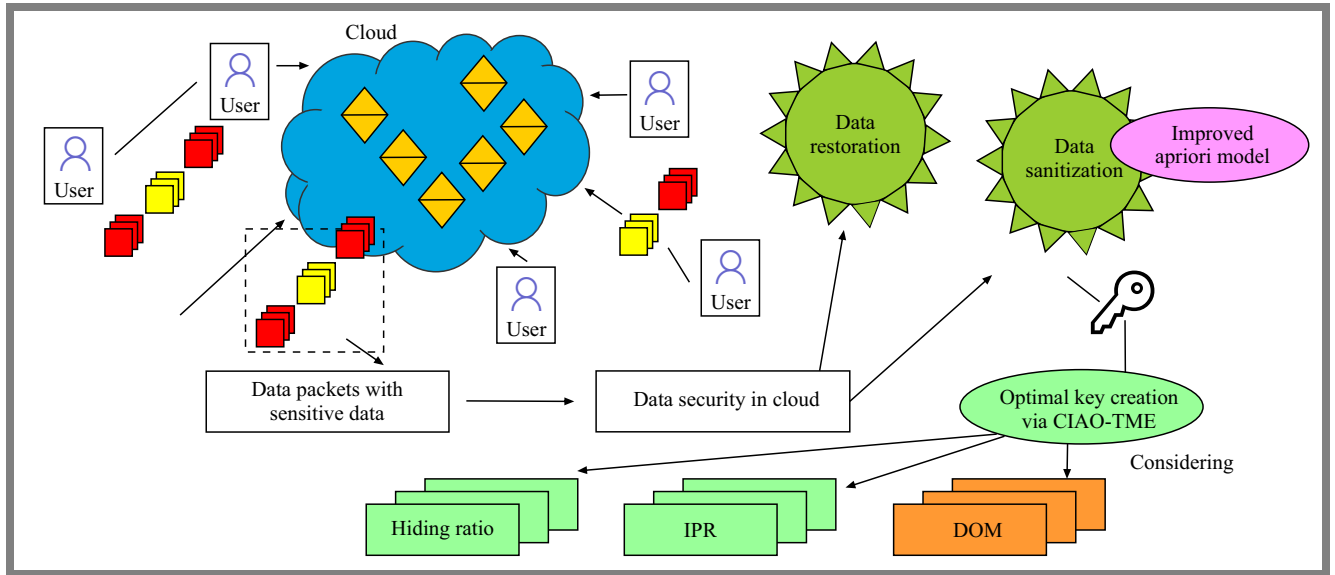


Fig. 1. The proposed concept ensuring data security in the cloud.

impact factor (IF) is found for removing the victim item set. It is fixed with a threshold value $\alpha_{\min}$ such as:

- if $\alpha_{\min}$ or IF is 1 or more, the item set should be continued (considered as sensitive data),
- if $\alpha_{\min}$ or IF equals 0 or less, the item set is considered as victim item set and it should be removed.

- 2) Using the overlap strategy for counting the support of candidate item set c_k , the sampling sets of L_{k-1} and of L_1 are created.
- 3) If the $|L_k| \leq k$, the algorithm is terminated.

By using the apriori model, the rules for sanitization are created and, similarly, the reverse rules are designed. Next, using the XOR function with the key values, the adopted data sanitization process is performed.

3.2. Key Generation

One of the important and computational time-consuming steps in the creation of a security mechanism is key generation. The creation of uncrackable and non-derivable secure keys is a complicated computational task. Creation of the key matrix and the initial cloud data are essential for preparing sanitized data, such as:

$$d'_S = d_S \oplus Ky_2, \quad (1)$$

where an optimally produced key is referred to as Ky_2 , d_S is the original data and d'_S is the sanitized data.

Generating the key is a part of the CIAO-TME model.

3.3. Restoration Procedure

Data sanitization is a technique used for concealing private or sensitive information in a cloud with the goal of preventing unwanted data leakages. Using the special key created during the data cleaning procedure, sensitive data is revealed during data restoration. The opposite is the data restoration process. The same key that the created the CIAO-TME model for the

purpose of generating sanitized data is used to recover the original information:

$$\hat{d}_S = d'_S \oplus Ky_2, \quad (2)$$

where $\hat{d}_S$ denotes the recovered data.

4. Novel CIAO-TME Optimal Key Selection

To identify the optimal key selection method, first the objective function is formulated as:

$$Obj = \text{Min}[\text{DOM} + (1 - \text{HR}) + (1 - \text{IPR})], \quad (3)$$

where HR, DOM, and IPR are the objective functions considered for data sanitization. Figure 2 presents the solution encoding scheme.

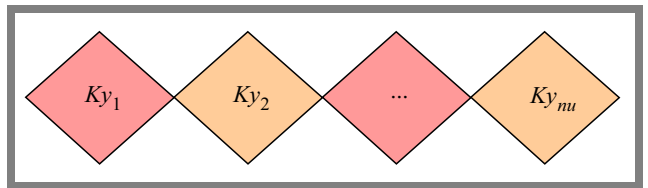


Fig. 2. The proposed data security in the cloud.

HR is the percentage of sensitive items that are properly concealed by d'_S :

$$\text{HR} = \frac{Nd}{tp}, \quad (4)$$

where tp is a total number of hidden data indexes, and Nd represents the size of non-zero indexes.

IPR is the inverse of data lost and the rate of non-sensitive rules not concealing in sanitized data set. It is formulated as:

$$\text{IPR} = \frac{1}{tp} N2, \quad (5)$$

where tp is the total number of saved data indexes and $N2$ equals the total amount of zero indexes.

The original dataset d_S and the sanitized dataset d'_S allow to determine the DOM function that is the Euclidean distance between d_S and d'_S :

$$\text{DOM} = -d'_S + d_S, \quad (6)$$

where d_S indicates the degree of modification taking place within the unique dataset and d'_S denotes the sanitized dataset d' .

4.1. Proposed CIAO-TME Model

The proposed CIAO-TME model is a hybrid optimization approach and is created by combining the traditional AOA [25] and CHIO techniques [26]. AOA is a high-performance optimization technique as far as convergence time and exploration-exploitation balance are concerned. The advantages of CHIO include high effectiveness when dealing with a large number of optimization issues across a wide range of optimization domains. To overcome the limitations of AOA and CHIO models, the hybrid optimization CIAO-TME model is proposed. It combines two common optimization models to speed up the convergence process offered by solutions [27]–[30]. The individual steps of the model are presented below.

In the first step, a population of N the search agents is initialized, and next the positions of the search agent and other algorithmic parameters are initialized:

$$S^I = LB^I + \text{Rand}(UB^I - LB^I), \quad (7)$$

where LB^I and UB^I denote the lower and upper bounds of I -th the search agent, respectively.

Then, density Den^I and volume Vol^I of the search agents are set randomly, and acceleration A^I of the I -th the search agent is assigned as:

$$A^I = \text{Rand}(UB^I - LB^I) + LB^I. \quad (8)$$

The initial population is then evaluated and the search agent characterized by the best fitness levels is selected as: a_{best} , Vol_{best} , and D_{best} . For the next position, density Den_{t+1}^I and volume Vol_{t+1}^I are determined by:

$$Den_{t+1}^I = \text{Rand}(-Den_t^I + Den_{best}) + Den_t^I, \quad (9)$$

$$Vol_{t+1}^I = Vol_t^I + \text{Rand}(Vol_{best} + Den_{best}) - Vol_t^I, \quad (10)$$

where Den_{best} and Vol_{best} points out to the best density and best volume, respectively. The Rand function points out to a random value with uniformly distribution. The transfer operator TF is modeled as:

$$TF = e^{\frac{itr - \max^{itr}}{\max^{itr}}}, \quad (11)$$

where itr and $\max^{itr}$ are the current and maximum iterations, respectively.

Step 2 is the exploitation phase. If $TF \leq 0.5$, then there occurs no collision between the searching agents. The acceleration factor is modeled as:

$$ACC_{t+1}^I = \frac{Den_{mr} + Vol_{mr} + A_{mr}}{Den_{t+1}^I Vol_{t+1}^I}, \quad (12)$$

where Den_{mr} , Vol_{mr} , and A_{mr} denote respectively density, volume, and acceleration of random search agents, respectively.

If $TF > 0.5$, the position of the search agent is updated as:

$$X_{t+1}^I = X_{t+1}^{best} + F \cdot C2 \cdot r \cdot A_{t+1}^{I-Norm} d(T \cdot X^{best} - X_t^{best}). \quad (13)$$

In CIAO-TME, the position is updated by combining the concepts of AOA and CHIO as:

$$X_{t+1}^I = \lambda(I) \frac{X_{t+1}^{best} + F \cdot C2 \cdot r \cdot A_{t+1}^{I-Norm}}{\tau \frac{1}{M}} d(T \cdot X^{best} - X_t^{best}), \quad (14)$$

where λ refers to the force of infection, τ is the transmission rate, and M refers to the birth rate percentage of individuals who are added to the entire population of M , A_{t+1}^{I-Norm} is the normalized acceleration, $C2 = 6$, $T = C3 \cdot TF$ and $C3 = 0.3$ are the constant values, while F is the flag that shows the direction. In addition, r points out the random value that is generated by means of the tent map to improve the convergence rate.

Step 3 defines the proposed exploration phase. When $TF > 0.5$, then there is no collision and the acceleration of the search agent $itr + 1$ is computed as:

$$ACC_{t+1}^I = \frac{Vol_{best} A_{best} + Den_{best}}{Vol_{t+1}^I Den_{t+1}^I}. \quad (15)$$

For CIAO-TME, the acceleration factor is computed using on sample variance S^2 and weight we^i as:

$$ACC_{t+1}^I = \frac{Den_{best} + Vol_{best} A_{best}}{Den_{t+1}^I \frac{Vol_{t+1}^I}{S^2}} we^i. \quad (16)$$

For $TF \leq 0.5$, the I -th search agent $itr + 1$ is computed as:

$$X_{t+1}^I C1 \cdot r \cdot d A_{t+1}^{I-Norm} (X^{rand} - X_t^I) + X_{t+1}^{best}. \quad (17)$$

As a result, the algorithm in step 4 returns the best solution found.

5. Results and Discussion

The proposed CIAO-TME data security model was implemented in Python using a sample dataset downloaded from [31]. The databases known as Cleveland, Hungary, Switzerland, and VA Long Beach were used. The accessible model was evaluated over AOA, CHIO, HBA, BES, and BMO for a variety of metrics, such as DOM, HR, IPR, and cost.

5.1. Convergence Analysis

Figure 3 shows the cost factors for three datasets (D1, D2, and D3), illustrating convergence of the implemented CIAO-TME and comparing it with the traditional schemes: AOA, CHIO, HBA, BES, and BMO. CHIO and AOA have revealed poor performance by acquiring more expensive assets in the first, second, and third scenarios. From the 12-th to the 50-th

iteration, CIAO-TME achieves a minimum convergence of 0.001 for all scenarios and slightly higher cost values before the 10-th iteration.

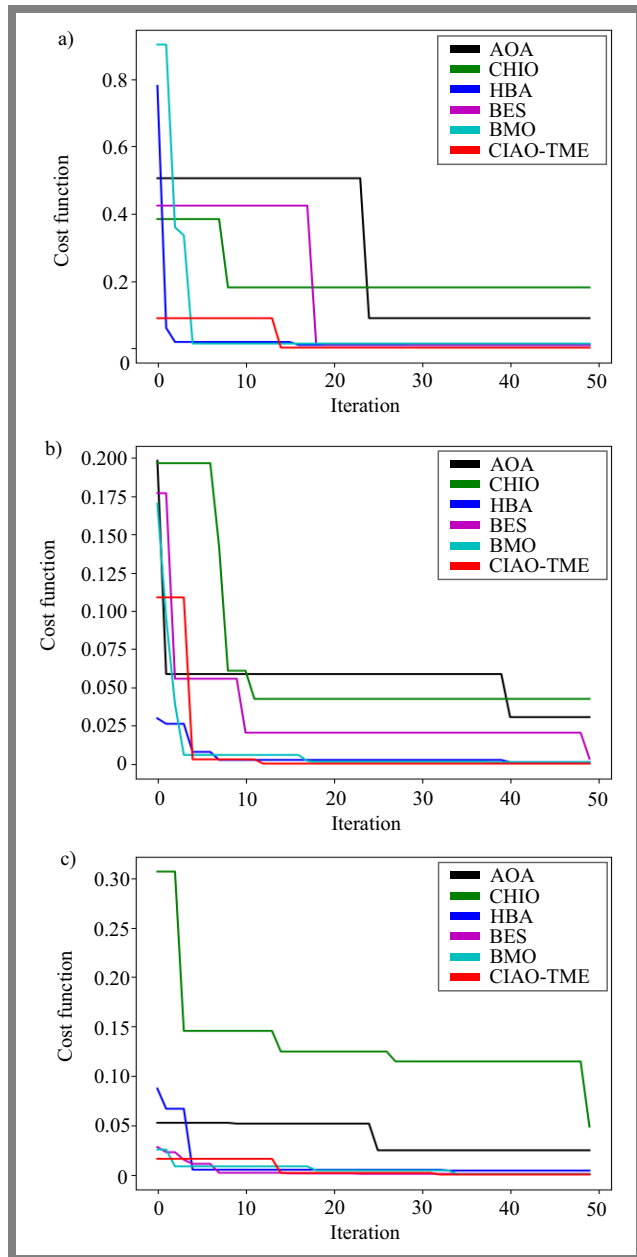


Fig. 3. Convergence of CIAO-TME for: a) D1, b) D2, and c) D3 datasets.

5.2. Analysis of DOM, HR, and IPR

Figure 4 illustrates the analysis of DOM, HR, and IPR parameters for the CIAO-TME method over known techniques, for D1, D2, and D3 datasets. The proposed CIAO-TME scheme achieved the lowest DOM (~ 0.45) – a result that better than the one characterizing other schemes, such as AOA, CHIO, HBA, BES, and BMO. Additionally, CIAO-TME achieved a higher HR (~ 1.0) for D2 and D3 and the lowest HR for D1, when compared to D2 and D3. Moreover, the proposed CIAO-TME scheme obtained a better IPR

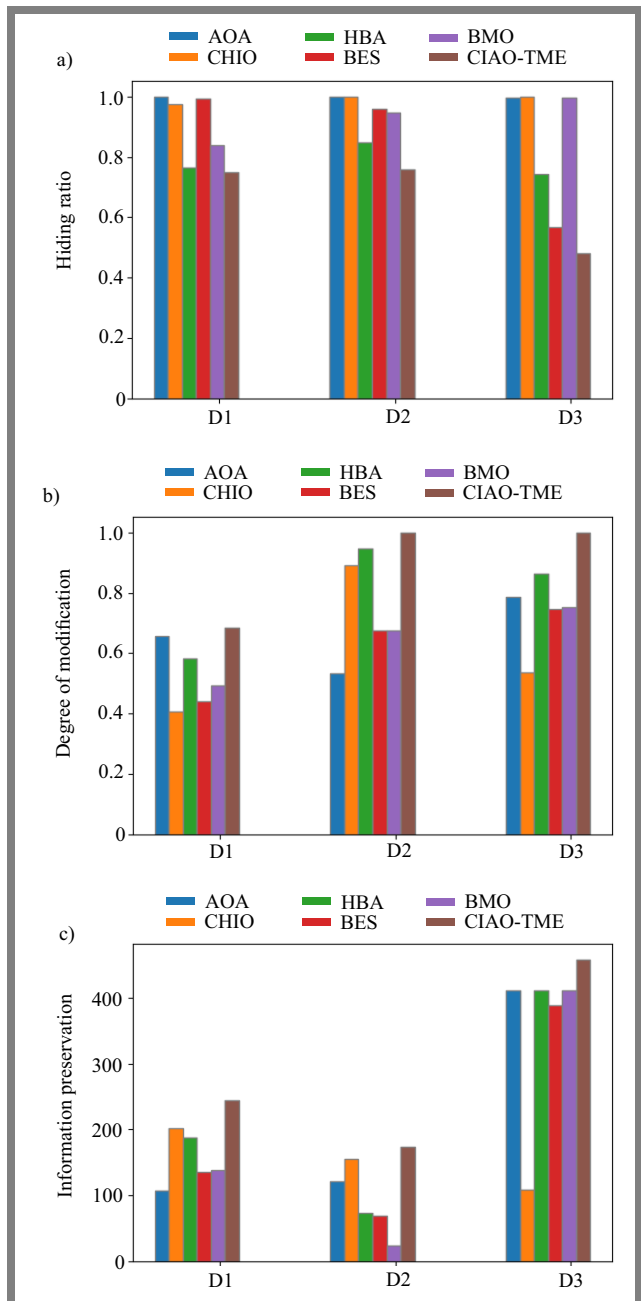


Fig. 4. Analysis of known schemes versus CIAO-TME for: a) DOM, b) HR, and c) IPR.

(450) for D3 and a relatively low IPR for D1 and D2, when compared to D3. The outcomes of using CIAO-TME are better when compared with AOA, CHIO, HBA, BES, and BMO models, especially with regard to the D3 dataset, due to the optimal generation of the key and an improved apriori-based ARM.

5.3. Fitness Analysis

Figure 5 shows the fitness examination defined by Eq. (3) for the CIAO-TME scheme over AOA, CHIO, HBA, BES, and BMO for D1, D2, and D3 datasets. The objective is to minimize the HR function while keeping IPR high and DOM at a relatively low levels. Here, the minimal outcomes

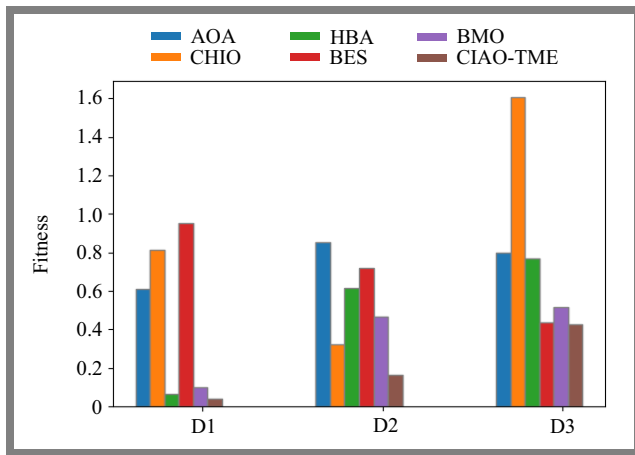


Fig. 5. Fitness analysis of CIAO-TME compared with others methods for D1, D2, and D3 datasets.

are achieved by CIAO-TME at D3. For D1 and D2, the CIAO-TME model has attained the best outcomes in terms of the objective. However, compared to D1, D2, D3-related outcomes were the best. This is the result of optimal generation of the key and an improved apriori-based ARM.

5.4. Comparison Summary

The proposed CIAO-TME cloud data security scheme is compared with existing ARM schemes for different metrics (Table 3). CIAO-TME-based security models are tested in terms of performance fitness at high HR, high IPR, and low DOM. An HR of 0.68 is achieved for CIAO-TME, i.e. a result that is better than the one obtained when using ARM or an approach with no CIAO-TME involved. The best results were achieved on D1 and D2 datasets. The merging of optimal key

Tab. 3. Comparison of the proposed scheme with other rule-mining methods and metrics.

D1 dataset			
Metrics	CIAO-TME	Proposed with existing ARM	Proposed with no CIAO-TME
Fitness	0.0429	0.8012	0.1987
HR	0.6859	0.5573	0.5573
IPR	244	108	108
DOM	0.7486	1	1
D2 dataset			
Metrics	CIAO-TME	Proposed with existing ARM	Proposed with no CIAO-TME
Fitness	0.1638	0.597	1.097
HR	1	0.9324	0.7891
IPR	174	5	54
DOM	0.7589	1	1
D3 dataset			
Metrics	CIAO-TME	Proposed with existing ARM	Proposed with no CIAO-TME
Fitness	0.4263	0.9637	0.5495
HR	1	0.8624	0
IPR	458	0	411
DOM	0.4823	1	0.9951

generation and the improved apriori-based ARM additionally enhanced the CIAO-TME scheme.

5.5. Analysis of Existing Works

For three test instances, Table 4 shows the comparison of the selected CIAO-TME with other existing schemes, such as GMGW [31] and J-SSO [18]. The adopted CIAO-TME-based security method is evaluated in terms of such metrics as fitness, HR, IPR, and DOM, with the lowest DOM (~ 0.48) achieved for D3. This result is superior to that of other currently used techniques. The D2 and D3 datasets allowed to obtain an improved HR (~ 1.0) by using CIAO-TME, and for the D1 dataset, the proposed solution achieved the lowest HR (~ 0.68). Additionally, the IPR on D3 was higher (~ 458), while for D1 and D2 IPR it was relatively low. Overall, the CIAO-TME produces superior results than AOA, CHIO, HBA, BES, and BMO models.

Tab. 4. Comparison of CIAO-TME with other methods.

D1 dataset			
Metrics	CIAO-TME	GMGW [31]	J-SSO [18]
Fitness	0.0429	0.7554	0.1023
HR	0.6859	0.3286	0.513
IPR	244	202	205
DOM	0.7486	0.975	0.9639
D2 dataset			
Fitness	0.1638	1.663	0.3699
HR	1	0.8531	0.647852
IPR	174	148	74
DOM	0.7589	1	0.8486
D3 dataset			
Fitness	0.4263	1.024	0.7383
HR	1	0.8122	0.8864
IPR	458	388	411
DOM	0.4823	1	0.7435

5.6. Attack Analysis

Tables 5–6 present the examination results for three datasets under simulated CPA and KPA attacks. A CPA is a cryptanalysis attack paradigm that assumes the attacker has access to the cypher texts for any plain. The KPA is a cryptanalysis attack type in which the attacker has access to both plain-text (also known as a crib) and encrypted version of the data (cipher text). These can be used to divulge additional classified information, including security codes and private keys. CIAO-TME attained minimal KPA attack and CPA attack vulnerability values versus AOA, CHIO, HBA, BES, and BMO for all datasets. Specifically, the third datasets revealed lower attack parameter values for the CPA, while for the KPA, dataset D1 is characterized by lower attack values than datasets D2 and D3. This significant improvement stems from due the enhanced apriori-based ARM and optimized key generation.

Tab. 5. CPA analysis results.

Methods	AOA	CHIO	HBA	BES	BMO	CIAO-TME
Dataset D1	0.1975	0.2834	0.2346	0.2701	0.2212	0.1686
Dataset D2	0.2125	0.2434	0.2653	0.2275	0.4494	0.1769
Dataset D3	0.2173	0.3794	0.3917	0.2928	0.2715	0.1629

Tab. 6. KPA analysis of the proposed and other known schemes.

Methods	AOA	CHIO	HBA	BES	BMO	CIAO-TME
Dataset D1	0.1862	0.3429	0.3003	0.3275	0.2793	0.1197
Dataset D2	0.1658	0.2658	0.2527	0.2049	0.2274	0.1285
Dataset D3	0.1835	0.2572	0.2713	0.386	0.2237	0.1612

6. Conclusion

This work proposes two procedures to protect sensitive data from unauthorized access and users by relying on the data sanitization process. The modified apriori approach is used to clean the data and the major goal is to generate the best key – a task accomplished by employing an optimization method. A reversible method, known as data restoration, allows to retrieve or obtaining cleaned content. IPR, HR, and DOM were the objectives set to tackle the optimization problem. In the simulations, the CIAO-TME scheme achieved the lowest DOM for the D3 dataset and overperformed other techniques that are currently in use, such as J-SSO and GMGW. The D2 and D3 datasets improved HR to approx. 1.0 with the CIAO-TME scheme. The CIAO-TME also achieved the lowest HR (0.68) on the D1 dataset.

References

- [1] Y. Wang, A. Zhang, P. Zhang, and H. Wang, “Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain”, *IEEE Access*, vol. 7, pp. 136704–136719, 2019 (<https://doi.org/10.1109/ACCESS.2019.2943153>).
- [2] X. Yang, M. Wang, X. Wang, G. Chen, and C. Wang, “Stateless cloud auditing scheme for non-manager dynamic group data with privacy preservation”, *IEEE Access*, vol. 8, pp. 212888–212903, 2020 (<https://doi.org/10.1109/ACCESS.2020.3039981>).
- [3] H. Liu, X. Yao, T. Yang, and H. Ning, “Cooperative privacy preservation for wearable devices in hybrid computing-based smart health”, *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1352–1362, 2019 (<https://doi.org/10.1109/JIOT.2018.2843561>).
- [4] H. Yan and W. Gui, “Efficient identity-based public integrity auditing of shared data in cloud storage with user privacy preserving”, *IEEE Access*, vol. 9, pp. 45822–45831, 2021 (<https://doi.org/10.1109/ACCESS.2021.3066497>).
- [5] C. Xu, N. Wang, L. Zhu, K. Sharif, and C. Zhang, “Achieving searchable and privacy-preserving data sharing for cloud-assisted e-healthcare system”, *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8345–8356, 2019 (<https://doi.org/10.1109/JIOT.2019.2917186>).
- [6] B.A. Jalil, T.M. Hasan, G.S. Mahmood, and H.N. Abed, “A secure and efficient public auditing system of cloud storage based on BLS signature and automatic blocker protocol”, *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 7, pp. 4008–4021, 2022 (<https://doi.org/10.1016/j.jksuci.2021.04.001>).
- [7] K. Wang, C.-M. Chen, Z. Tie, M. Shojafar, S. Kumar, and S. Kumari, “Forward privacy preservation in IoT-enabled healthcare systems”, *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1991–1999, 2022 (<https://doi.org/10.1109/TII.2021.3064691>).
- [8] Q. Kong, R. Lu, F. Yin, and S. Cui, “Privacy-preserving continuous data collection for predictive maintenance in vehicular fog-cloud”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5060–5070, 2021 (<https://doi.org/10.1109/TITS.2020.3011931>).
- [9] J. Wang, D. Shi, J. Chen, and C.-C. Liu, “Privacy-preserving hierarchical state estimation in untrustworthy cloud environments”, *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1541–1551, 2021 (<https://doi.org/10.1109/TSG.2020.3023891>).
- [10] M. Fernandes, J. Decouchant, M. Völz, F.M. Couto, and P. Esteves-Verissimo, “DNA-SeAI: Sensitivity levels to optimize the performance of privacy-preserving DNA alignment”, *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 3, pp. 907–915, 2020 (<https://doi.org/10.1109/JBHI.2019.2914952>).
- [11] R. Li, T. Song, B. Mei, C. Hu, W. Li, M. Larson, X. Cheng, and R. Bie, “A cloud-based framework for verifiable privacy-preserving spectrum auction”, *High-Confidence Computing*, vol. 2, 2022 (<https://doi.org/10.1016/j.hcc.2021.100037>).
- [12] S. Mewada, “Data mining-based privacy preservation technique for medical dataset over horizontal partitioned”, *International Journal of E-Health and Medical Communications (IJEHMC)*, vol. 12, no. 5, pp. 50–66, 2021 (<https://doi.org/10.4018/IJEHMC.202109.01.0a4>).
- [13] X. Xu, S. Fu, L. Qi, X. Zhang, Q. Liu, Q. He, and S. Li, “An IoT-oriented data placement method with privacy preservation in cloud environment”, *Journal of Network and Computer Applications*, vol. 124, pp. 148–157, 2018 (<https://doi.org/10.1016/j.jnca.2018.09.006>).
- [14] C. Fang, Y. Guo, N. Wang, and A. Ju, “Highly efficient federated learning with strong privacy preservation in cloud computing”, *Computers & Security*, vol. 96, Article 101889, 2020 (<https://doi.org/10.1016/j.cose.2020.101889>).
- [15] K.M. Prabha and P.V. Saraswathi, “Suppressed K-anonymity multi-factor authentication based Schmidt-Samoa cryptography for privacy preserved data access in cloud computing”, *Computer Communications*, vol. 158, pp. 85–94, 2020 (<https://doi.org/10.1016/j.comcom.2020.04.057>).
- [16] C.-T. Li, D.-H. Shih, and C.-C. Wang, “Cloud-assisted mutual authentication and privacy preservation protocol for telecare medical information systems”, *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 191–203, 2018 (<https://doi.org/10.1016/j.cmpb.2018.02.002>).
- [17] J. Mandala and M.V.P. Chandra Sekhara Rao, “Privacy preservation of data using crow search with adaptive awareness probability”, *Journal of Information Security and Applications*, vol. 44, pp. 157–169, 2019 (<https://doi.org/10.1016/j.jisa.2018.12.005>).
- [18] D. Ahamad, S.A. Hameed, and M. Akhtar, “A multi-objective privacy preservation model for cloud security using hybrid Jaya-based shark smell optimization”, *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 6, part A, pp. 2343–2358, 2020 (<https://doi.org/10.1016/j.jksuci.2020.10.015>).

- [19] A. Mondal and R.T. Goswami, "Enhanced HoneyPot cryptographic scheme and privacy preservation for an effective prediction in cloud security", *Microprocessors and Microsystems*, vol. 81, 2021 (<https://doi.org/10.1016/j.micpro.2020.103719>).
- [20] X. Yao, F. Farha, R. Li, I. Psychoula, L. Chen, and H. Ning, "Security and privacy issues of physical objects in the IoT: Challenges and opportunities", *Digital Communications and Networks*, vol. 7, no. 3, pp. 373–384, 2021 (<https://doi.org/10.1016/j.dcan.2020.09.001>).
- [21] T. Kanwal, A. Anjum, S.U.R. Malik, A. Khan, and M.A. Khan, "Privacy preservation of electronic health records with adversarial attacks identification in hybrid cloud", *Computer Standards & Interfaces*, vol. 78, 2021 (<https://doi.org/10.1016/j.csi.2021.103522>).
- [22] N. Tian, Q. Guo, and H. Sun, "Privacy preservation method for MIQP-based energy management problem: A cloud-edge framework", *Electric Power Systems Research*, vol. 190, 2021 (<https://doi.org/10.1016/j.epsr.2020.106850>).
- [23] L. Hernández-Álvarez, J. María de Fuentes, L. González-Manzano, and L.H. Encinas, "SmartCAMPP – Smartphone-based continuous authentication leveraging motion sensors with privacy preservation", *Pattern Recognition Letters*, vol. 147, pp. 189–196, 2021 (<https://doi.org/10.1016/j.patrec.2021.04.013>).
- [24] P.J. Sun, "Security and privacy protection in cloud computing: Discussions and challenges", *Journal of Network and Computer Applications*, vol. 160, 2020 (<https://doi.org/10.1016/j.jnca.2020.102642>).
- [25] F. A. Hashim, K. Hussain, E.H. Houssein, M.S. Mabrouk, and W. Al-Atabany, "Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems", *Applied Intelligence*, vol. 51, pp. 1531–1551, 2021 (<https://doi.org/10.1007/s10489-020-01893-z>).
- [26] F. Martínez-Álvarez, G. Asencio-Cortés, J.F. Torres, D. Gutiérrez-Avilés, L. Melgar-García, R. Pérez-Chacón, C. Rubio-Escudero, J.C. Riquelme, and A. Troncoso, "Coronavirus optimization algorithm: A bioinspired metaheuristic based on the COVID-19 propagation model", *Big Data*, vol. 8, no. 4, 2020 (<https://doi.org/10.1089/big.2020.0051>).
- [27] M.M. Beno, I.R. Valarmathi, S.M. Swamy, and B.R. Rajakumar, "Threshold prediction for segmenting tumour from brain MRI scans", *International Journal of Imaging Systems and Technology*, vol. 24, no. 2, pp. 129–137, 2014 (<https://doi.org/10.1002/ima.22087>).
- [28] R. Thomas and M.J.S. Rangachar, "Hybrid optimization based DBN for face recognition using low-resolution images", *Multimedia Research*, vol. 1, no. 1, pp. 33–43, 2018 (DOI: 10.46253/j.mr.v1i1.a5).
- [29] J. Devagnanam and N.M. Elango, "Optimal resource allocation of cluster using hybrid grey wolf and cuckoo search algorithm in cloud computing", *Journal of Networking and Communication Systems*, vol. 3, no. 1, pp. 31–40, 2020 (<https://doi.org/10.46253/jnacs.v3i1.a4>).
- [30] R.A. Mustafa, H.S. Chyad, and J.R. Mutar, "Enhancement in privacy preservation in cloud computing using apriori algorithm", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 3, pp. 1747–1757, 2022 (<https://doi.org/10.11591/ijeecs.v26.i3.pp1747-1757>).
- [31] M.M. Annie Alphonsa and P. Amudhavalli, "Genetically modified glowworm swarm optimization based privacy preservation in cloud computing for healthcare sector", *Evolutionary Intelligence*, vol. 11, pp. 101–116, 2018 (<https://doi.org/10.1007/s12065-018-0162-4>).

Rajkumar Patil

Assistant Professor, Department of Computer Science Engineering
E-mail: rajkumarpatil8827@gmail.com
Department of Computer Science Engineering, GITAM (Deemed to be) University, Hyderabad, Telangana, India

Gottumukkala HimaBindu, PhD

Assistant Professor, Department of Computer Science Engineering
E-mail: hgottumu@gitam.edu
Department of Computer Science Engineering, GITAM (Deemed to be) University, Hyderabad, Telangana, India

Single Linkage Weighted Steepest Gradient Adaboost Cluster-Based D2D in 5G Networks

Sridhar Varadala and S. Emalda Roslin

Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India

<https://doi.org/10.26636/jtit.2023.167222>

Abstract — Efficiency of data transmissions with minimum latency levels and better resource utilization is a challenging issue in 5G device-to-device (D2D) environments. A novel technique referred to as single linkage steepest gradient gentle AdaBoost cluster-based device (SLSGAC) is introduced to improve device-to-device communications with minimum latency. The proposed technique uses the ensemble clustering approach to group mobile devices by constructing a set of weak clusters, based on the Minkowski single linkage clustering technique. In the weak clustering process, residual energy, bandwidth and SINR are estimated, and mobile devices are grouped based on the Minkowski distance measure. Results of the weak clustering process are combined to provide the final ensemble's clustering output by applying the steepest gradient function to minimize the error rate. For each cluster, a head is selected from among the group members to improve the data transmission rate and minimize latency. Simulations are conducted comparing the proposed technique with the existing methods based on such metrics as energy efficiency, data delivery ratio, packet loss rate, throughput and latency.

Keywords — 5G networks, D2D communication, Minkowski single linkage clustering.

1. Introduction

Among the many modern solutions, device-to-device (D2D) communication is an upcoming technology for enhancing the performance of 5G cellular communication. D2D is currently the subject of many scientific works aimed at expanding the application range, improving parameters and performance. Below is a short review of the important works in this area.

The social-aware D2D video delivery model based on rapid sample-efficient measurement (DMSEM) was introduced in [1]. It relies on the fuzzy c-means clustering algorithm. However, the designed algorithm failed to consider the process of accounting for mobility to improve D2D data transmissions. A dynamic mode selection algorithm for improving the communication based on the fuzzy clustering process was designed in [2]. Unfortunately, good performance of efficient latency-aware data communication was not achieved. A novel cluster-based architecture was introduced in [3] for D2D mmWave communication to enhance the performance of 5G networks. The intended architecture succeeds in increasing throughput but not in increasing data delivery. A fuzzy

clustering-based approach was introduced in [4] for a 5G ultra-dense network in order to improve the data transmission rate. But accurate clustering of mobile devices with multiple characteristics was not considered. A novel weighted clustering algorithm for the distributed architecture was proposed in [5] to improve throughput and minimize energy consumption. But the designed method failed to apply machine learning techniques to further improve data delivery. In [6], random-based clustering (RBC) and channel-gain-based clustering (CG-BC) approaches were introduced to realize D2D clustering. The clustering algorithm decreases the delay but no higher delivery rates and lower packet loss levels were achieved. In [7], a random linear network coding-based cooperative relaying (RNCC) system was developed for network coding supported by D2D transmissions based on resource distribution. The drawback of this method is that the multiple-hop model-based D2D transmission for network coding-assisted D2D communications was not provided.

A centralized non-uniform clustering routing protocol was introduced in [8], along with the power estimation as well as communication cost. Unfortunately, the performance of latency-aware data transmission was not verified. Various issues, solutions and challenges were reviewed in [9] to efficiently utilize the available resources and thus minimize latency and enhance data transmission rates.

A VNF resource allocation system was introduced in [10] to obtain the best number of clusters for reducing delay in data transmission. An artificial neural network-based channel model framework was introduced in [11] for 5G wireless communication. But cluster-based data delivery was not performed to minimize latency. A data-driven multi-objective optimization approach was introduced in [12] for the purpose of developing hyperdense 5G networks. Unfortunately, no analyses concerned with a wider span of network communications were performed in order to minimize latency and improve reliability. A MIMO-NOMA cellular system approach for D2D communications was introduced in [13]. But the efficiency of the algorithms was not improved.

In [14], an intensive benchmarking of the incorporation of D2D communication was discussed to improve energy efficiency, throughput and latency. The reinforcement learning-based latency-controlled D2D connectivity (RL-LCDC) method was introduced in [15]. Though the designed

method minimizes the delay and energy consumption, the packet loss rate was not reduced. A distributed intelligent method was developed in [16] to enhance the data rate and reduce power consumption in D2D communications. But no latency-aware D2D communication was performed. Several machine learning techniques were introduced in [17] for 5G cellular networks to improve the data transmission rate. The sparse code multiple access (SCMA) method was introduced in [18] for enhancing the overall network performance of D2D-enabled cellular networks. The dynamic resource block (RB) sharing method was introduced in [19] to enhance the network throughput of the overall system. However, the designed method failed to minimize the latency of data communication.

To solve such issues, a technique called SLSGAC is introduced, offering the following contributions:

- it is based on the ensemble clustering approach to boost D2D communication,
- it uses weak clusters as well as the Minkowski single linkage clustering technique to divide the cellular network into different groups based on residual energy, bandwidth, and SINR. It uses the steepest gradient function into the ensemble clustering technique to minimize the error rate and obtain accurate clustering results,
- to improve data delivery and minimize latency, a cluster head is selected for each cluster,
- finally, a simulation is performed using several performance measures versus the traditional approaches.

The paper is organized into different sections. Section 2 describes the proposed SLSGAC technique. Section 3 presents the simulation of the proposed technique and conventional methods. Finally, Section 4 contains the summary of the paper.

2. Proposed Method

Figure 1 illustrates the data exchange between mobile devices and the base station. Two mobile devices communicate with each other in a direct manner and via the links provided from the base station. In such a way, communication is entirely handled by the base station.

In the course of direct communication between the devices, minimum utilization of resources is essential for increasing

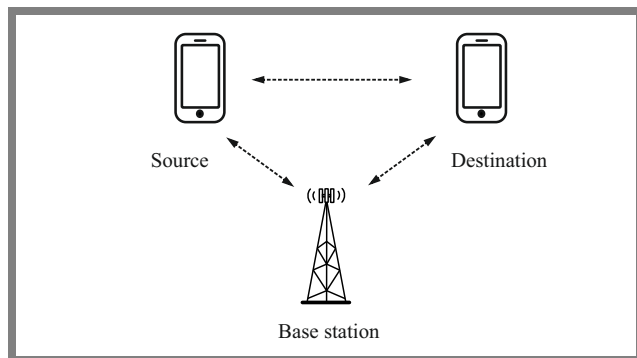


Fig. 1. Data flow in a 5G network.

the network's lifespan. In order to achieve resource-efficient data transmission rates, a novel cluster-based technique is introduced.

Let us consider a scenario in which 5G network comprises a number of mobile devices $d_1, d_2, d_3, \dots, d_n$ distributed randomly within the network. Each mobile device has different computational resources available, such as energy E , bandwidth B , and signal to noise-plus-interference ratio (SINR).

Figure 2 illustrates the concept behind the SLSGAC technique to enhance data communication between mobile devices within a 5G network. The number of mobile devices is considered as input. Then, the ensemble clustering approach is applied to partition the entire hole network into different clusters depending on such resources as energy, bandwidth, and SINR. For each cluster, a head is selected among the individual members to additionally boost data transmission.

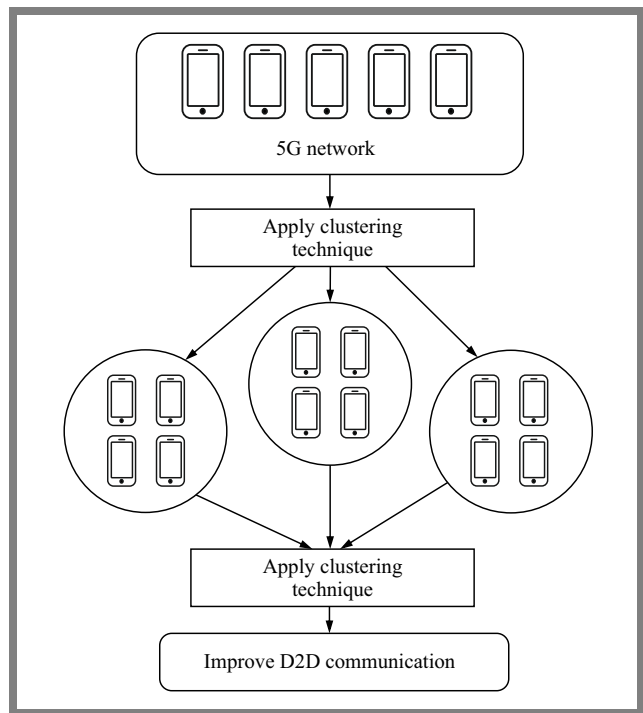


Fig. 2. Diagram of architecture of SLSGAC technique.

The single linkage AdaBoost clustering technique is an ensemble meta-algorithm that primarily converts weak learners' output into that of strong learners. A weak learner $z_1, z_2, z_3, \dots, z_k$ is defined as a base clustering technique that is only slightly correlated with the true clustering results. In contrast, a strong learner is a clustering technique that offers true clustering outcomes. The weak clustering results are merged and the final, improved clustering results are obtained.

Figure 3 illustrates the basic process of ensembling clustering results. The mobile devices are presented, as input, to the clustering technique. Then k weak learners are constructed using the Minkowski single linkage clustering technique to partition the entire network into different groups depending on energy, bandwidth, signal to interference, and noise ratio.

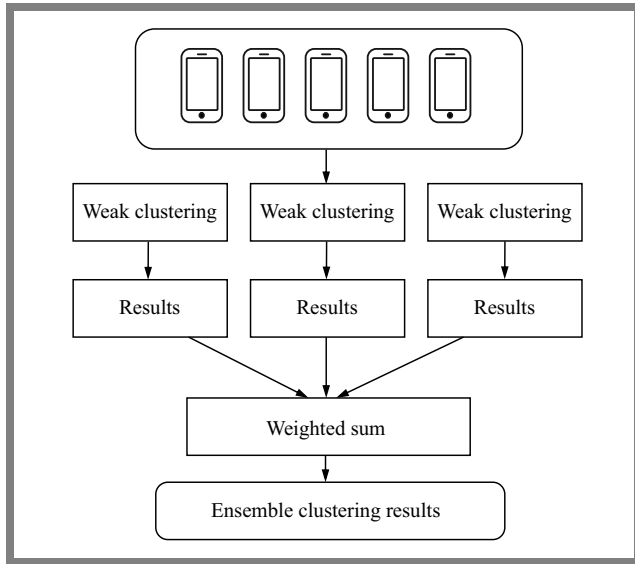


Fig. 3. Basic process of ensembling clustering results.

In the first step, energy is measured in terms of power and time as:

$$E_{device} = p_t t, \quad (1)$$

where E_{device} denotes the energy of the mobile device, p_t indicates the transmitting power of the mobile device, t represents time in seconds. The remaining (i.e. residual) energy of each device is measured as:

$$RE_{device} = E_t - E_C, \quad (2)$$

where RE_{device} indicates the remaining energy of the device in Joules, E_t indicates the total energy, and E_C denotes the consumed energy.

Bandwidth is estimated as the maximum amount of data transmitted in a given time slot and it expressed as:

$$BW = \frac{D_p}{t}, \quad (3)$$

where BW denotes the bandwidth, t is time in seconds, and D_p denotes a data packet.

The signal to interference and noise ratio R_{sin} is estimated as:

$$R_{sin} = \frac{s_p}{n_p + i_p}, \quad (4)$$

where s_p denotes signal power, n_p indicates noise power, and i_p denotes interference power.

Based on the above parameters, the single linkage clustering technique is applied to group the sensor nodes based on the Minkowski distance. The method identifies the minimum distance between mobile devices based on the energy, bandwidth, and SINR. The distance is formulated as:

$$D_{ij} = (|d_i - d_j|^r)^{\frac{1}{r}}, \quad (5)$$

where D_{ij} denotes the distance between two mobile devices d_i and d_j , r denotes the Minkowski distance of order (≥ 1). The minimum distance between two mobile devices is grouped and clusters are formed:

$$Q = \arg \min(D_{ij}), \quad (6)$$

where Q denotes a base clustering result, $\arg \min$ denotes an argument of the minimum function. Consequently, weak learner results are obtained. In order to attain the ensemble clustering results, the weak learners are summarized:

$$Z = \sum_{i=1}^k Q_i, \quad (7)$$

where Z denotes ensemble clustering results, Q_i denotes weak classification results. The weighted sum of the weak learner results is obtained as:

$$Z = \sum_{i=1}^k Q_i \beta_i, \quad (8)$$

where β_i indicates the weight of clustering results and the weight is an integer. For each weak clustering result obtained, the training error is estimated as:

$$E(f) = (R_A - Q_i)^2, \quad (9)$$

where $E(f)$ denotes an error function, R_A denotes the actual clustering result, Q_i denotes the clustering results observed.

Next, the proposed technique updates the initial weight based on the error value. Finally, the steepest gradient descent is applied to find the weak clustering results with a minimum error among the multiple weak clustering results:

$$Y = \arg \min E(f), \quad (10)$$

where Y indicates the strong clustering results, $\min E(f)$ denotes the minimum argument of the error function of the clustering results. In this way, all sensor nodes are grouped into a particular cluster. For each cluster, the mobile device with the highest level of energy, bandwidth, and SINR is selected as a cluster head to ensure efficient data transmission from the source to the destination.

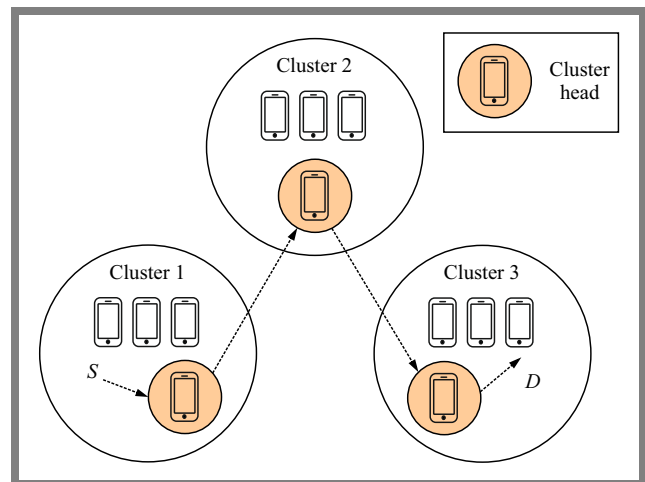


Fig. 4. Single linkage weighted steepest gradient AdaBoost cluster-based D2D communication.

Figure 4 illustrates the process of cluster-based data transmission from source mobile device S to destination D via a cluster head. As shown in Fig. 4, the source device in cluster 1 sends the data to the destination in cluster 3. This process is

carried out via the cluster head, resulting in increases in the data delivery and in minimized latency. Algorithm 1 describes the technique's algorithmic phase.

Algorithm 1. Description of the proposed technique.

Input: number of mobile devices: $d_1, d_2, d_3, \dots, d_n$

Initialization

- 1: **For** each mobile device d_i
- 2: Construct k weak learners $z_1, z_2, z_3, \dots, z_k$
- 3: **Measure** $RE_{device}, BW, SINR$
- 4: **Estimate** Minkowski distance D_{ij}
- 5: Find minimum distance $\arg \min[D_{ij}]$
- 6: Group sensor nodes into clusters
- 7: Obtain clustering results
- 8: **End for**
- 9: Sum all weak clustering results $\sum_{i=1}^k Q_i \beta_t$
- 10: Initialize weight $\sum_{i=1}^k Q_i \beta_t$
- 11: Compute error function $E(f)$
- 12: Update weight
- 13: Find the weak learner with minimum error $\arg \min E(f)$
- 14: Return (strong clustering results)

The ensemble boost clustering technique computes the number of weak learners. Next, residual energy, bandwidth, and signal to noise ratio of each device are estimated and the mobile devices are grouped into a particular cluster. The weak learner's results are summarized to create a strong one in order to achieve accurate clustering results. After combining the weak learners' results, the weight is initialized and measures the training error function. Then, the initial weight gets updated based on the error value. Finally, the steepest gradient function is applied to find the weak learner with the minimum error value. Once the devices are grouped, a cluster head is selected for efficient communication. Then, data transmission is performed via the cluster head to achieve a higher delivery ratio and minimum latency.

3. Simulation and Analysis

In this section, a Matlab simulation of the proposed SLSGAC technique is performed and the results are compared with existing DMSEM [1] and the dynamic mode selection method based on a fuzzy clustering algorithm [2], with performance measured in terms of energy efficiency, data delivery ratio, packet loss rate, throughput, and latency.

Energy efficiency is measured as the fraction of output energy required to perform communication:

$$EE = \frac{E_o}{E_I} \cdot 100 [\%], \quad (11)$$

where EE denotes energy efficiency, E_o indicates output energy, and E_I denotes input energy.

The simulation of energy efficiency of various methods based on 100, 200, 300... 1000 mobile devices is presented in Fig. 5. The obtained values indicate that the energy efficiency of all three methods increases with the growing number of mobile devices. For comparison, the proposed SLSGAC technique provides better performance than the conventional methods.

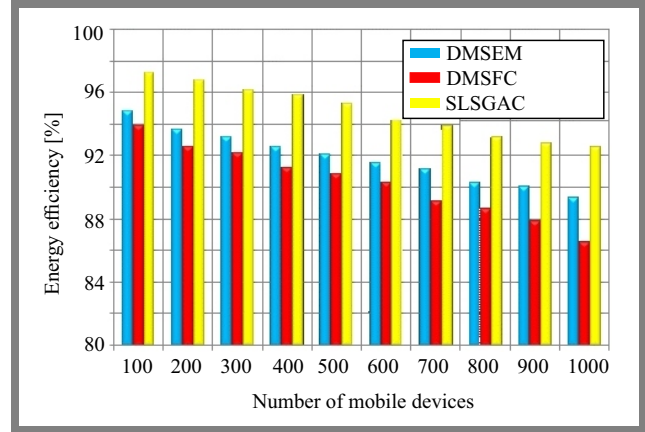


Fig. 5. Performance results – energy efficiency.

Data delivery ratio is a performance metric based on data packets successfully transmitted from the base station to a mobile device. This parameter is also used to estimate the confidential data transmission between devices:

$$DR = \frac{\text{Data received}}{\text{Data sent}} \cdot 100 [\%]. \quad (12)$$

Figure 6 presents experimental data delivery outcomes for varying numbers of mobile devices. For each method, ten iterations are performed relating to different counts of mobile devices. While increasing the input count, the data delivery ratio for all the methods becomes lower. a comparison of all methods shows that the SLSGAC technique outperforms other approach in achieving a higher delivery ratio.

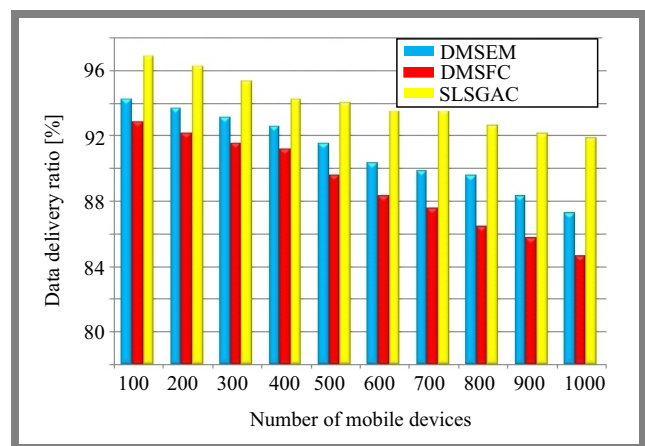


Fig. 6. Performance results – data delivery ratio.

Figure 6 illustrates the performance of the data delivery ratio versus the number of mobile devices in the simulation scenario. The SLSGAC technique achieves a delivery ratio of 96.8% in the first iteration, with 100 devices present.

Similarly, DMSEM [1] and dynamic mode selection based on a fuzzy clustering algorithm [2] generate delivery ratios of 94.2% and 92.8%, respectively. The results indicate that the SLSGAC technique provides better performance due to using the ensemble technique which groups the sensor nodes based on residual energy, bandwidth, and signal to interference ratio. For each cluster, a head is chosen for delivering the data packets from the source to the destination.

Packet loss rate is estimated based on the ratio between data packets lost and the number of data packets transmitted:

$$PLR = \frac{\text{Data lost}}{\text{Data sent}} \cdot 100 [\%]. \quad (13)$$

Figure 7 shows a comparison of the packet loss rates. 100 mobile devices were used for conducting the simulation in the first iteration. For SLSGAC, the packet loss rate is 3.2%, whereas for DMSEM and dynamic mode selection based on fuzzy clustering algorithm, the results equal 5.8% and 7.2%, respectively. The SLSGAC technique minimizes the packet loss rate more effectively than the existing methods, due to the selection of cluster heads among the group members. The source mobile device sends the data to its destination via a neighboring cluster head. This process helps minimize data transmission losses.

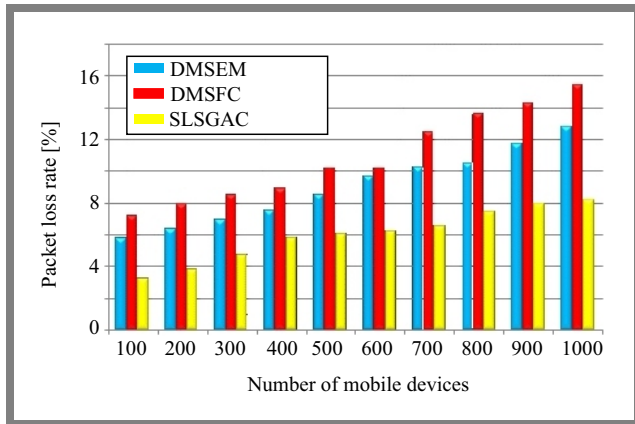


Fig. 7. Performance results – packet loss rate.

Throughput is estimated as the amount of data sent from one base station to mobile devices in a specified time frame. Throughput is computed in bits per second, using the following formula:

$$\text{Throughput} = \frac{\text{Size of data received}}{\text{Time}} [\text{bps}]. \quad (14)$$

Figure 8 shows that the proposed SLSGAC technique provides a higher network throughput due to the fact that higher residual energy requirements are applied when more bandwidth-efficient devices are selected as cluster heads, thus resulting in an improvement in data transmission speed.

Figure 9 illustrates the simulation results concerning latency measured for the number of mobile nodes varying from 100 to 1000. Latency of the SLSGAC technique is reduced when compared to other existing methods in 5G D2D communication. For 100 mobile devices, latency of the SLSGAC technique is 11.5 ms in the first iteration. Similarly,

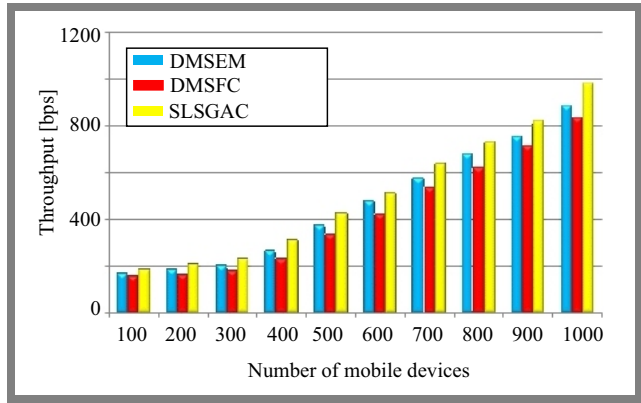


Fig. 8. Performance results – throughput.

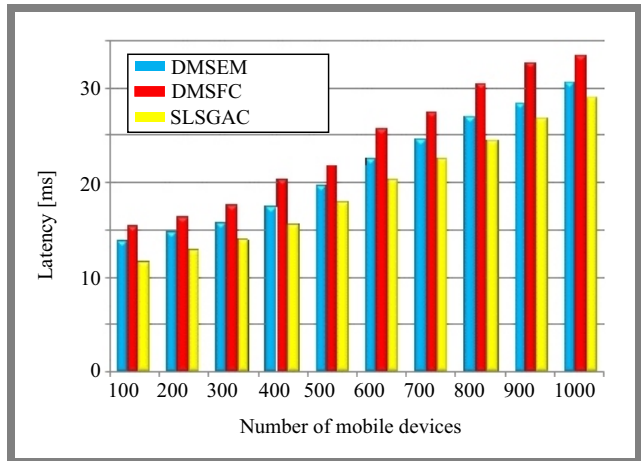


Fig. 9. Performance results – latency.

values of 13.7 ms and 15.3 ms were observed in the case of DMSEM and dynamic mode selection based on a fuzzy clustering algorithm, respectively. This is due to the application of the ensemble clustering technique, where residual energy, bandwidth, and signal to interference ratio are estimated and the clustering the mobile devices is based on the Minkowski distance measure. The weak clustering results are combined to provide the final ensemble clustering output by applying the steepest gradient function in order to find efficient mobile devices and achieve more efficient data communication between devices while maintaining minimum latency levels.

4. Conclusion

The simulation results prove that the proposed SLSGAC technique provides better performance than the traditional approaches in terms of achieving higher throughput, data delivery, and minimum latency. When compared to conventional cellular networks, D2D communication is anticipated to offer a number of advantages. Future networks are most likely to benefit from and be most successful with D2D technology. In the study, we focused on a thorough analysis of the D2D technologies now in use, along with their features, including device detection, interference management, security, mode selection, and power control. We examine a number of

ideas that have been proposed in an effort to ensure secure device-to-device (D2D) connectivity in 5G. By highlighting the drawbacks or faults, as well as the suggested fixes, we emphasize the current solutions.

References

- [1] R. Zhang, S. Jia, Y. Ma, and Ch. Xu, "Social-aware D2D video delivery method based on mobility similarity measurement in 5G ultra-dense network", *IEEE Access*, vol. 8, pp. 52413–52427, 2020 (<https://doi.org/10.1109/ACCESS.2020.2980865>).
- [2] A.A. Algedir and H.H. Refai, "Energy efficiency optimization and dynamic mode selection algorithms for D2D communication under HetNet in downlink reuse", *IEEE Access*, vol. 8, pp. 95251–95265, 2020 (<https://doi.org/10.1109/ACCESS.2020.2995833>).
- [3] G. Mesbahi and A.G. Rahbar, "Cluster-based architecture capable for device-to-device millimeter-wave communications in 5G cellular networks", *Arabian Journal for Science and Engineering*, vol. 44, pp. 9719–9733, 2019 (<https://doi.org/10.1007/s13369-019-03830-w>).
- [4] Q. Liu, Ch.F. Kwong, S. Zhang, L. Li, and J. Wang, "A fuzzy-clustering based approach for MADM handover in 5G ultra-dense networks", *Wireless Network*, vol. 28, pp. 965–978, 2022 (<https://doi.org/10.1007/s11276-019-02130-3>).
- [5] S. Aslam, F. Alam, S.F. Hasan, and M. Rashid, "A novel weighted clustering algorithm supported through a distributed architecture for D2D enabled content-centric networks", *Sensors*, vol. 20, no. 19, Article number 5509, 2020 (<https://doi.org/10.3390/s20195509>).
- [6] M. Gharbieh, A. Bader, H. ElSawy, H.-Ch. Yang, M.-S. Alouini, and A. Adinoyi, "Self-organized scheduling request for uplink 5G networks: A D2D clustering approach", *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1197–1209, 2019 (<https://doi.org/10.1109/TCOMM.2018.2876008>).
- [7] X. Xiao, M. Ahmed, X. Chen, Y. Zhao, Y. Li, and Z. Han, "Accelerating content delivery via efficient resource allocation for network coding aided D2D communications", *IEEE Access*, vol. 7, pp. 115783–115796, 2019 (<https://doi.org/10.1109/ACCESS.2019.2930728>).
- [8] W. Chen, B. Zhang, X. Yang, W. Fang, W. Zhang, and X. Jiang, "C-EEUC: A cluster routing protocol for coal mine wireless sensor network based on fog computing and 5G", *Mobile Networks and Applications*, vol. 27, no. 5, pp. 1853–1866, 2022 (<https://doi.org/10.1007/s11036-019-01401-9>).
- [9] M.H. Adnan and Z.A. Zuckarnain, "Device-to-device communication in 5G environment: Issues, solutions, and challenges", *Symmetry*, vol. 12, no. 11, 1762, 2020 (<https://doi.org/10.3390/sym1211762>).
- [10] S. Song, Ch. Lee, H. Cho, G. Lim, and J.-M. Chung, "Clustered virtualized network functions resource allocation based on context-aware grouping in 5G edge networks", *IEEE Transactions on Mobile Computing*, vol. 19, no. 5, pp. 1072–1083, 2020 (<https://doi.org/10.1109/TMC.2019.2907593>).
- [11] J. Huang, Ch.-X. Wang, L. Bai, J. Sun, Y. Yang, J. Li, O. Tirkkonen, and M.-T. Zhou, "A big data enabled channel model for 5G wireless communication systems", *IEEE Transactions on Big Data*, vol. 6, no. 2, pp. 211–222, 2020 (<https://doi.org/10.1109/TBDATA.2018.2884489>).
- [12] B.B. Haile, E. Mutafulungwa, and J. Hämäläinen, "A data-driven multi-objective optimization framework for hyperdense 5G network planning", *IEEE Access*, vol. 8, pp. 169423–169443, 2020 (<https://doi.org/10.1109/ACCESS.2020.3023452>).
- [13] S. Solaiman, A. Nassef, and E. Fadel, "User clustering and optimized power allocation for D2D communications at mm wave underlaying MIMO-NOMA cellular networks", *IEEE Access*, vol. 9, pp. 57726–57742, 2021 (<https://doi.org/10.1109/ACCESS.2021.3071992>).
- [14] H.H. Hussein, H.A. Elsayed, and S.M. Abd El-kade, "Intensive benchmarking of D2D communication over 5G cellular networks: Prototype, integrated features, challenges, and main applications", *Wireless Networks*, vol. 26, pp. 3183–3202, 2020 (<https://doi.org/10.1007/s11276-019-02131-2>).
- [15] A.G. Sreedevi and T.R. Rao, "Reinforcement learning algorithm for 5G indoor device-to-device communications", *Transaction on Emerging Telecommunication Technologies*, vol. 30, no. 9, pp. 1–10, 2019 (<https://doi.org/10.1002/ett.3670>).
- [16] I. Ioannou, V. Vassiliou, Ch. Christophorou, and A. Pitsillides, "Distributed artificial intelligence solution for D2D communication in 5G networks", *IEEE Systems Journal*, vol. 14, no. 3, pp. 4232–4241, 2020 (<https://doi.org/10.1109/JSYST.2020.2979044>).
- [17] S. Sevgican, M. Turan, K. Gökarslan, H.B. Yilmaz, and T. Tugcu, "Intelligent network data analytics function in 5G cellular networks using machine learning", *Journal of Communications and Networks*, vol. 22, no. 3, pp. 269–280, 2020 (<https://doi.org/10.1109/JCN.2020.000019>).
- [18] A. Sultana, I. Woungang, A. Anpalagan, L. Zhao, and L. Ferdouse, "Efficient resource allocation in SCMA-enabled device-to-device communication for 5G networks", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5343–5354, 2020 (<https://doi.org/10.1109/TVT.2020.2983569>).
- [19] P.K. Barik, A. Shukla, R. Datta, and Ch. Singhal, "A resource sharing scheme for intercell D2D communication in cellular networks: A repeated game theoretic approach", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7806–7820, 2020 (<https://doi.org/10.1109/TVT.2020.2991476>).

Sridhar Varadala

Research Scholar

 <http://orcid.org/0000-0003-3715-8197>

E-mail: varadalasri@gmail.com

Department of Electronics and Communication Engineering,
Sathyabama Institute of Science and Technology, Chennai,
India

S. Emalda Roslin, Ph.D.

 <http://orcid.org/0000-0001-8699-9510>

Department of Electronics and Communication Engineering,
Sathyabama Institute of Science and Technology, Chennai,
India

An Adaptive Richardson-Lucy Algorithm for Medical Image Restoration

Qunoot A. Yaqoub and Ayad A. Al-Ani

College of Information Engineering, Al-Nahrain University, Baghdad, Iraq

<https://doi.org/10.26636/jtit.2023.168222>

Abstract — Image restoration is the process of estimating the original image content from a degraded picture. In this paper, the Richardson-Lucy iterative algorithm was developed to improve the quality of degraded medical images. It has been assumed that medical images are exposed to two types of degradation. The first type is the blur function in the Gaussian form with different widths, i.e. $\sigma = 1, 2$, and 3 . The second type of degradation was assumed to be of the independent white Gaussian noise type with different signal-to-noise ratio values: $\text{SNR} = 10, 50$, and 100 . The results obtained from the adaptive filter are compared, quantitatively, with different conventional filters: inverse, Wiener, and constraint least square, by applying different measures, such as: power signal to noise ratio (PSNR), structural similarity index (SSID), and root mean square error (RMSE). The comparison showed that the adaptive recovery filter achieves better results.

Keywords — constraint least square filter, Gaussian blurring function, Gaussian noise function, inverse filter, non-blind deconvolution, point spread function, Richardson-Lucy algorithm, Wiener filter.

1. Introduction

In many applications, recorded or generated images are a degraded version of the original picture due to flaws in the imaging and capturing processes. These imperfections are very difficult to remove. The process of eliminating them is, therefore, an important task in image processing [1]. Flaws or degradation may be caused by noise or may result from blurring. Blurring can occur during image creation, transmission, and storage, and may result from transmission channel error, camera defects, atmospheric turbulence, relative motion between the object and the camera, as well as other factors [2]. The term “noise” refers to an unwanted random variation in the ideal signal and can be caused by a variety of factors, such as poor detector sensitivity, environmental factors, the discrete nature of radiation, transmission or quantization errors, and so on [3]. Additive noise and multiplicative noise are the two types of interference encountered. Additive noise is a linear image additive that is unaffected by the strength of the input signal. The Gaussian distribution with a mean of zero is used to represent the probability density function. Noise is also assumed to be white because its spectrum power distribution is nearly constant. Multiplicative noise differs depending on whether the input signal is multiplicative or correlated with the original signal. The Poisson distribution is used to

represent this type of noise [4]. The relationship between a degraded (blurry and noisy) image $g(x, y)$ and the object (original image) $f(x, y)$ is given by [1]:

$$g(x, y) = h(x, y) \otimes f(x, y) + n(x, y), \quad (1)$$

where $n(x, y)$ is the additive white Gaussian noise, $h(x, y)$ is the blurring function, also known as the point spread function (PSF), and $\otimes$ denotes the convolution process.

Medical imaging is nowadays essential for guiding disease diagnosis and treatment. One of the most basic and crucial aspects of medical imaging is image reconstruction. The main goal is to have high-quality clinical images used at the lowest possible cost and risk to the patients. Medical imaging devices may aid in early detection of diseases and, thus, in improving public health. MRI and CT scans, for example, are degraded by poor quality of equipment or patient movements during the image acquisition phase. This results in images that are mostly unusable. This shows that medical images must be of high quality to offer a quick and accurate diagnosis. To remove blurring and noise from degraded images, restoration processes must be performed. The goal of the restoration process is to estimate an image that is as close as possible to the original image $f(x, y)$, based on its degraded counterpart $g(x, y)$ [5]. There are two types of image deconvolution algorithms: blind (in which both PSF and noise are unknown) and non-blind (in which both PSF and noise are known). The blurring PSF and SNR from noise added are not known ahead of time in the case of blind restoration [6]. Figure 1 shows the degradation and restoration phases.

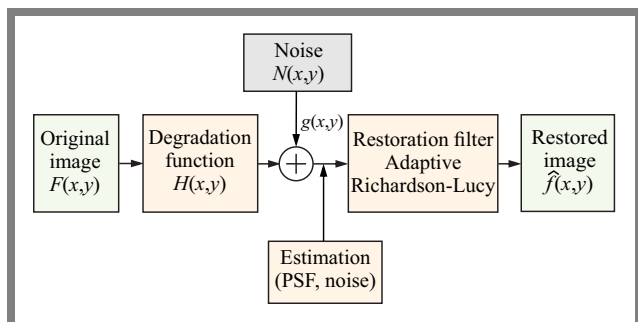


Fig. 1. Image degradation and restoration phases.

In this study, an iterative restoration technique known as the Richardson-Lucy algorithm (R-L) was adopted. A modified filter was created to restore CT scan medical images that have been blurred with a Gaussian function in different standard

deviation (STD) “width” of PSF = 1 and 3, and also corrupted by adding white Gaussian noise with different degrees of noise variance, i.e. with SNR = 10, 50, and 100.

The R-L algorithm has a problem with amplified noise and introduces the ringing effect as the number of iterations is increased, which also renders the computational process complex. The current research concept is based on the enhancement of the R-L algorithm in order to investigate the effect of strengthening the image’s edges during the restoration process, as well as the ability of the proposed algorithm to eliminate ringing and noise issues. This is accomplished by introducing two parameters, λ and β , into the equation with varying values, as well as calculating image edge detection parameters using the Sobel edge detection approach. Then, we compare the outcomes using various restoration filters, including inverse, Wiener, constraint least squares filter (CLSF), and the standard R-L, before modifying the image, in order to investigate how this development could affect the outcomes of the image restoration process. The effectiveness of this approach is determined using specific criteria, including mean square error (MSE), peak signal to noise ratio (PSNR), and structural similarity index (SSIM).

The main contributions of this study are:

- to represent the effect of adding to the priory restored image $f^{(k)}(x, y)$ with some β weight,
- to strengthen the image’s edges, adding the edges to the priory restored image with some λ weight.

The rest of this document is organized as follows. Section 2 offers an review of existing approaches that may be found in the literature. Section 3 describes the general recommended framework. Section 4 assesses our research experiment. Section 5 concludes the paper.

2. Related Works

Many research models focusing on this topic have been implemented to date. However, each strategy employed was tailored to a specific application area. Panfilova *et al.* [7] proposed a method of restoration consisting in extending the image beyond its original borders, reducing its brightness to zero at the new borders, and using an empirical criterion for defining the point at which the iterative process ends. To lessen the negative impact of the image restoration problem, linear blur compensation is applied. The Richardson-Lucy iterative method and its modifications are used to restore the image. The proposed modifications reduced image distortion by more than 50% in terms of the RMS metric, allowing the authors to estimate the number of iterations required to ensure better performance of the algorithm.

Anaconda-Mosquera *et al.* [8] presented hardware for implementing the R-L algorithm and for accomplishing the image restoration task, where the images are blurred due to the camera’s motion relative to the scene. The R-L was implemented in this case with the use of an FPGA-based platform relying on VHDL and with assumption that the capturing image sys-

tem is free of additive noise. The overall architecture scales from 3×3 to 9×9 mask sizes for the R-L convolution steps. The high digital signal processing complexity of this architecture, which depends on the number of iterations to be implemented, is a disadvantage.

Tselousov *et al.* [9] proposed a hybrid estimation method based on cepstral and gradient field analysis of the distorted image. The method offers better performance in terms of blind deconvolution image restoration based on the Bayes approach. The Richardson-Lucy engine revealed that, in most cases, images distorted by camera shake are characterized by better visual quality.

Aouinti *et al.* [10] introduced a genetic approach to optimize the iteration count of the R-L deconvolution to obtain a better restoration of degraded satellite images, enhancing image quality by increasing SNR. Panfilova *et al.* [11] proposed a control algorithm for R-L restoration to end its iterative procedure based on the criteria defined by the quality measure introduced. It was discovered that the restoration quality criteria based on the analysis of inter-line and inter-column correlation coefficients act as an alternative rather than a subordinate to other quality measures in use.

Liu *et al.* [12] improved the performance of the star sensor network working in a highly dynamic environment, using the R-L algorithm based on a radial basis function neural network (RBFNN). Firstly, using angular velocity information provided by a gyroscope, the point spread function (PSF) is calculated. RBFNN is then used to predict how many iterations the R-L algorithm will require to complete the process of deblurring the star image. The authors demonstrated that quality of the deblurred star image is excellent, even at high angular rates. Breykina *et al.* [13] developed an algorithm for automated image restoration based on the analysis of line-to-line and column-to-column correlation coefficients. The R-L method is used for image reconstruction in order to refine the distortion operator iteratively, and gradient descent is used to estimate the distorting operator. Experiments showed that images with smear and blur distortion yield correct restoration results, and that the method reduces the number of calculations required.

Zhao *et al.* [14] proposed a new image quality assessment index-based on the R-L algorithm. This method was designed for self-adaptive image de-blurring. Through experimental verification and comparison to other methods, the authors prove that the method achieves higher levels of efficiency, accuracy, and real-time performance. Bhonsle proposed, in [15], the R-L algorithm to correct the blurring problem, and the krill herd optimization to fine-tune the quality of the restored image, which results in a better restored image with higher PSNR, SNR and SSIM metrics.

Lyu *et al.* [16] attempted to improve ultrasonic c-scan image quality by combining the R-L algorithm and a faulty measurement model. Through the restoration of the simulated images, a connection between the optimal number of iterations and the flaw size-to-sound wavelength ratio is established, and the R-L iterative algorithm is used to restore the experimental

ultrasonic c-scan images with the optimal number of iterations. The proposed restoration method improves the accuracy of ultrasonic c-scan imaging significantly.

Li *et al.* [17] presented the Richardson-Lucy network (RLN), a fast and lightweight deep learning method for the deconvolution of three-dimensional fluorescence microscopy. The R-L is combined with a fully convolutional network structure in RLN, establishing a connection to the image formation process and thus improving network performance. The authors prove that RLN outperforms classical R-L deconvolution on volumes contaminated with severe out-of-focus fluorescence or noise, and it reconstructs large, cleared-tissue datasets four to six times faster than traditional multi-view pipelines.

3. Proposed Method

The purpose of this paper is to recover the quality of degraded medical images by using an inverse filter, Wiener filter, constrained least squares filter (CLSF), and the Richardson-Lucy algorithm to restore original images and to compare these results with adaptive Richardson-Lucy filter. Figure 2 shows the architecture of the proposed system.

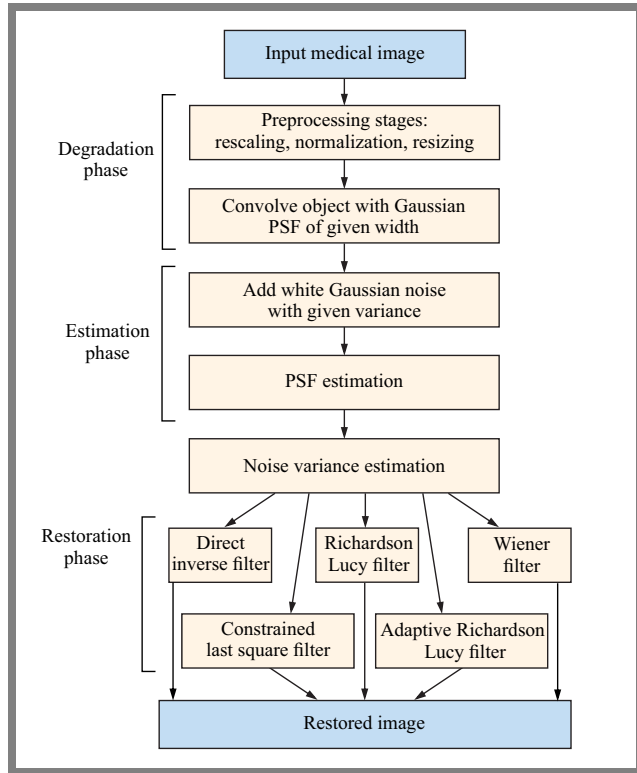


Fig. 2. Block diagram of the proposed system.

Direct inverse filtering is the simplest method of restoration. In this method, the Fourier transform (FT) of the degraded image $G(u, v)$ is divided by the FT of the degradation function $H(u, v)$ to estimate the FT of the image $F(u, v)$ [18]:

$$F(u, v) = \frac{G(u, v)}{H(u, v)}. \quad (2)$$

This method works fine when the degraded image contains no additive noise. The inverse filter has the advantage of

requiring only the blur PSF as a priori knowledge and allowing for perfect restoration in the absence of noise. As a result, it excessively dominates inverse-filtered images.

The Wiener filter is also known as the minimum mean square estimator. Some of the difficulties associated with inverse filtering were alleviated by attempting to recreate the error in the restored image using statistical methods. After modeling the error mathematically, the average error is minimized. As a result, the term minimum refers to a square estimator. The Wiener filter equation is [4]:

$$F(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{\delta_n(u, v)}{\delta_f(u, v)}} G(u, v), \quad (3)$$

where $H^*(u, v)$ denotes the complex conjugate of $H(u, v)$, $\delta_n(u, v)$ is the noise power spectrum, and $\delta_f(u, v)$ is the original image power spectrum.

The constrained least squares filter (CLSF) is another method for overcoming some of the Wiener filter's drawbacks, because only prior knowledge of the noise's mean and variance is required. This is accomplished by incorporating a smoothing criterion into the filter derivation, so that the result does not contain undesirable oscillations which affect the image by forming "waves". The constrained least squares filter is given by [8]:

$$F(u, v) = \frac{1}{H(u, v)} \left| \frac{|H^*(u, v)|^2}{|H(u, v)|^2 + \gamma |p(u, v)|^2} \right| G(u, v), \quad (4)$$

where γ is the adjustment factor, $p(u, v)$ is the smoothness criterion's Fourier transform (FT). The value is determined experimentally and is application dependent. A common function to use for $p(x, y)$ which is the inverse Fourier transform of $P(u, v)$ is the mask of the Laplacian filter [4]:

$$P(u, v) = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}, \quad (5)$$

3.1. Richardson-Lucy Deconvolution

The Richardson-Lucy deconvolution is a well-known image restoration method classified as a non-blind image restoration approach, because it necessitates prior knowledge of the blur and noise [13]. It is an iterative process that is dependent on the number of iterations, i.e. the most important factor in determining image quality. The R-L algorithm can be used efficiently when PSF is known or there is little or no noise data available [19]. The R-L algorithm is built on a maximum likelihood technique which uses Poisson statistics to model the image. Maximizing the model's likelihood function produces an equation that is satisfied when the next iteration converges [15]. The R-L algorithm was derived from Bayes' theorem. Because it is based on conditional probabilities, the algorithm considers statistical fluctuations in the signal and can thus reconstruct degraded images. Bayes' theorem is given by [20]:

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx}, \quad (6)$$

where $p(y|x)$ is the event's conditional probability y , given event x . $p(x)$ is the probability of event x , and $p(x|y)$ is the conditional probability in reverse, i.e. the probability of event x given event y . Probability $p(x)$ can be identified as the distribution of objects $f(x)$, while conditional probability $p(y|x)$ can be recognized as PSF-centered at x , i.e., $g(y, x)$. Probability $p(y)$ can be recognized as a degraded image or a convolution $c(y)$. The Richardson-Lucy algorithm equation is defined as [21]:

$$f^{(k+1)}(x, y) = \left[\frac{g(x, y)}{h(x, y) \otimes f^{(k)}(x, y)} \otimes h^T(x, y) \right] f^{(k)}(x, y), \quad (7)$$

where $f^{(k+1)}(x, y)$ is the new estimate of a blurred image from the previous one $f^{(i)}(x, y)$, $g(x, y)$, i is the iteration step, $h(x, y)$ is the blur filter (PSF) and $h^T(x, y)$ is the complex conjugate of PSF h , while $\otimes$ denotes convolution.

The R-L deconvolution algorithm has gained popularity in astronomy and medical imaging [20]. It is one of the most widely used deblurring algorithms in the field of image processing. It happens for a variety of reasons, including the fact that it operates efficiently regardless of the type of noise affecting the image. Furthermore, it does not require any information from the original clean image. This algorithm works in the presence of noise, but the noise increases along with the increase in the number of iterations [21]. Despite its benefits, the R-L method has several serious drawbacks, e.g. noise amplification.

3.2. Sobel Edge Detection

In an image, an edge is formed between the boundaries of two distinct regions [22]. The main goal of edge detection is to reduce the amount of data that needs to be processed by simplifying the pixels that make up the image's boundaries. The Sobel edge detection filter is one of the best edge detection algorithms, with a relatively low complexity [23]. The algorithm recognizes the boundaries of an image's horizontal and vertical axes separately. Typically, a pair of 3×3 horizontal and vertical convolution kernels is used by the Sobel operator to perform a 2D spatial gradient measurement on images [24]. To calculate approximate derivatives, the Sobel operator is given by two kernels (G_x, G_y) that are convolved with the original image. One of the kernels is for horizontal changes, and the other one is for vertical modifications, and at each location in the image, it can be written as [24]:

$$G = \sqrt{G_x^2 + G_y^2}. \quad (8)$$

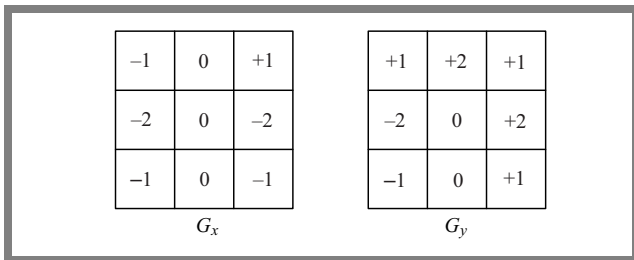


Fig. 3. Sobel masks with 3×3 dimensions.

As shown in Fig. 2, the first kernel G_x determines the incline in the x (columns) direction, while the other kernel G_y determines the tilt in the y (rows) direction [23].

4. Experimental Results and Discussions

The proposed work was implemented using the Matlab software. Before the restoration process, different preprocessing steps of raw data are completed.

Rescaling – the input image must be changed from an RGB color image to a 8-bit grayscale space picture in order to reduce the amount of data needed to represent the image. Consequently, computational efficiency and speed will be increased significantly. The conversion is given by a common formula from [25]:

$$\text{Gray} = 0.30 R + 0.59 G + 0.11 B. \quad (9)$$

Normalization – the process of data normalization makes sure that every input parameter has a consistent data distribution for each pixel. As a result, data is normalized by subtracting the mean from each pixel and dividing the result by the standard deviation (STD) value. The normalization process can be expressed as [26]:

$$x_{new} = a + \frac{(x - m_{min})(b - a)}{(x_{max} - x_{min})}, \quad (10)$$

where x represents a collection of the observed values, x_{min} represents the minimum values and x_{max} represents the maximum values in entire collection.

Resizing an image increases or decreases its resolution and is equivalent to scaling. The image size can be specified manually or by using a scaling factor. There are various algorithms for resizing an image that is stored in a compressed form, e.g. [27] which converts an image of size $m \times m$ into an image of size $k \times k$, where m is the number of pixels in the original image and k is the number of pixels in the compressed image.

4.1. Degradation Phase

A blurred image is produced by convolving the original image with Gaussian PSF. Two different values of PSF widths have been taken in consideration: $\sigma = 1$ and 3, as shown in Fig. 4b-c. To simulate a noisy image, we add white Gaussian noise (WGN) to the original one to achieve SNR = 10, 50, and 100, as shown in Fig. 4d-i.

4.2. Estimation Phase

In the case of blind restoration, the degradation parameters (PSF, noise) must be estimated. Inverse, Wiener, constraint least squares (CLSF) and R-L filters are non-blind restoration techniques. Known degradation parameters are applied to the degraded image during processing. Therefore, PSF and

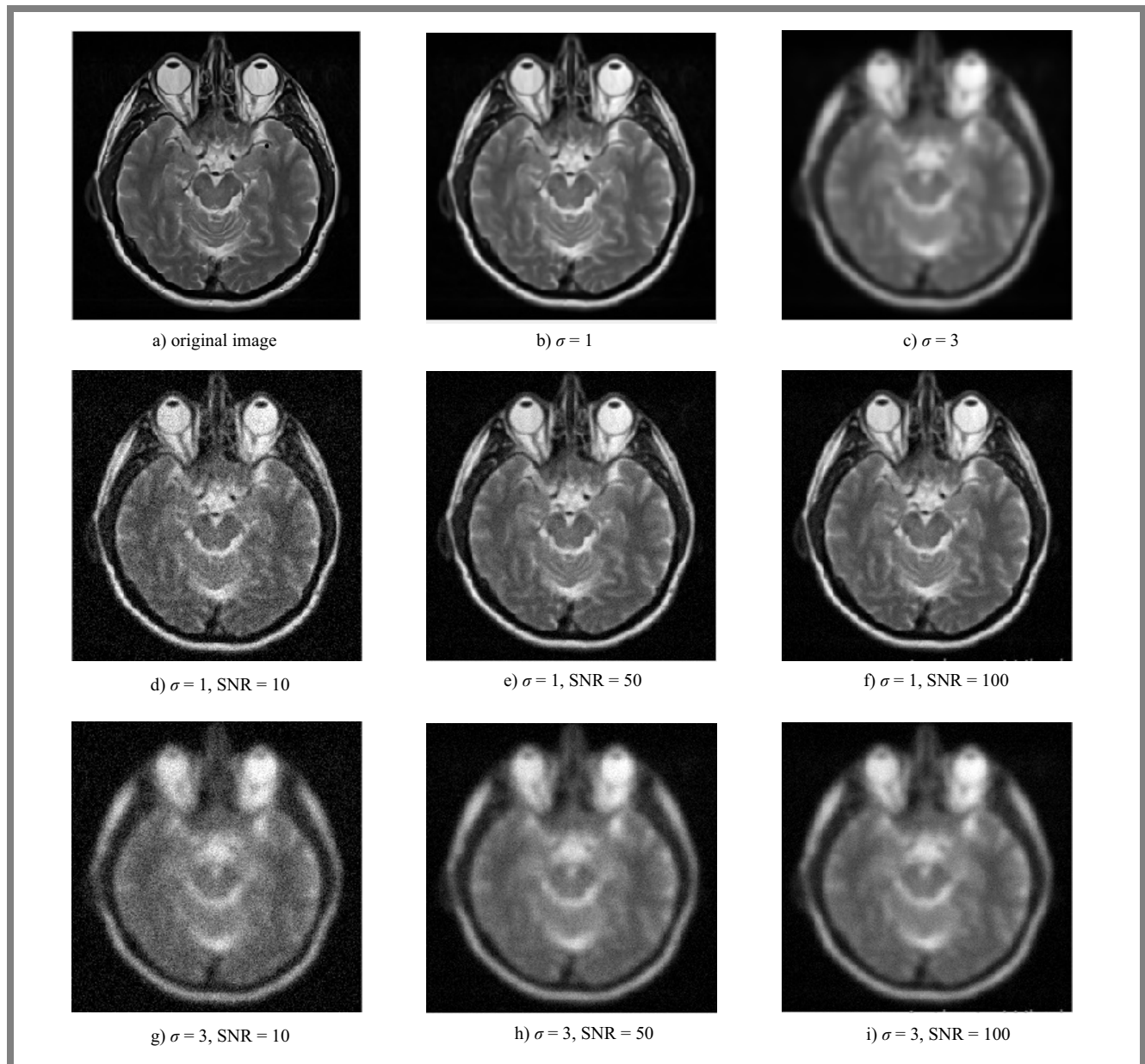


Fig. 4. Results of a degraded CT scan image for different values of σ and SNR.

noise must be estimated by using one of the estimations techniques.

4.3. Restoration Phase

To restore the original image $f(x, y)$, filters such as the inverse filter, Wiener filter, constraint least squares filter (CLSF) and R-L have been used. All filters were used to restore the degraded CT scan image, blurred with the Gaussian blurring function with various standard deviation values and distorted with Gaussian noise with different noise levels.

We note from Fig. 5, which represents the effect of a restored CT scan image for different values of σ and SNR by using the inverse filter, that the restored images are not good and the landmarks and details are not clear with or without noise and different values of PSF.

From Fig. 6, which represents the effect of a restored CT scan image for different σ and SNR values by using the Wiener filter, we note that the restored images look good for small PSF levels in the absence of noise, but quality is bad when the noise is present and PSF is increased.

From Fig. 7, which represents the effect of restoration by using CLSF, we see that the restored images are good for low PSF values in the absence of noise, but with high PSF levels, the resonance effect begins to appear on the edges. Moreover, when the noise is present, the image quality decreases with high PSF and low SNR.

4.4. Adaptive Richardson-Lucy Filter

Development of an RL algorithm to eliminate the problem of noise amplification and reduce the number of iterations

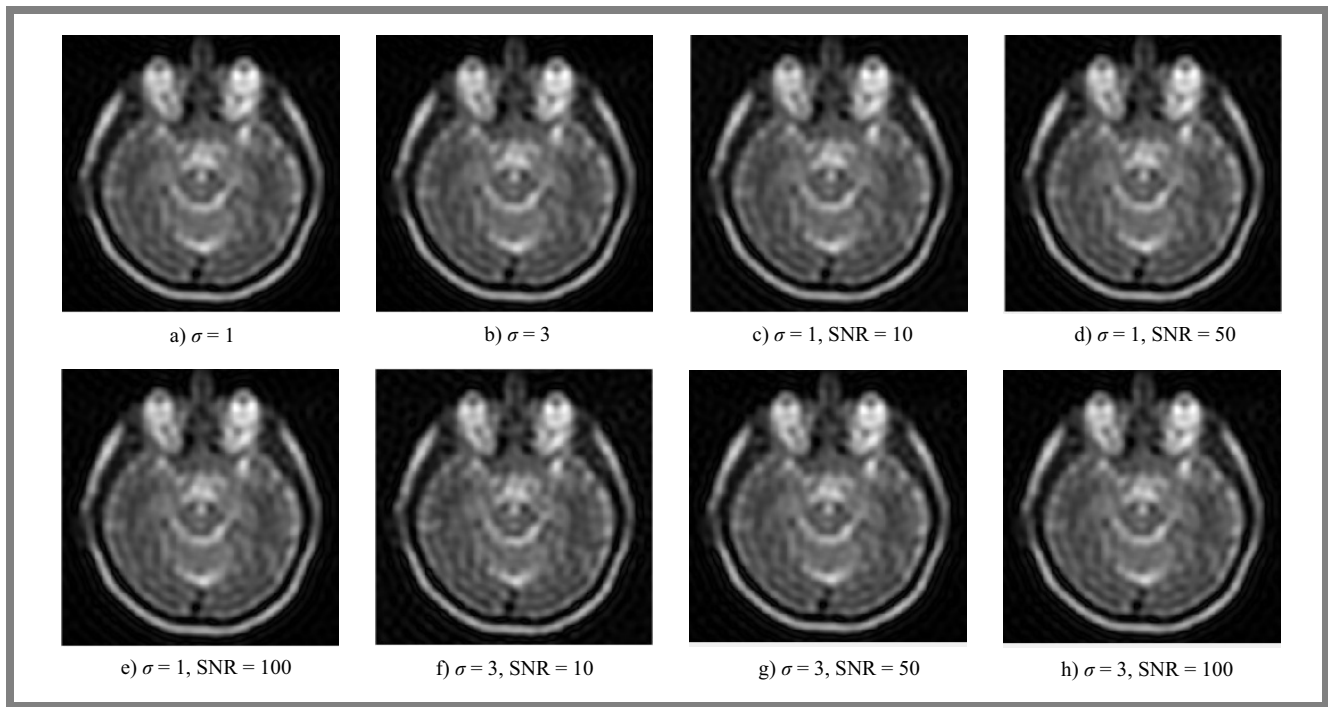


Fig. 5. Results of restoration for different σ and SNR values by using the inverse filter.

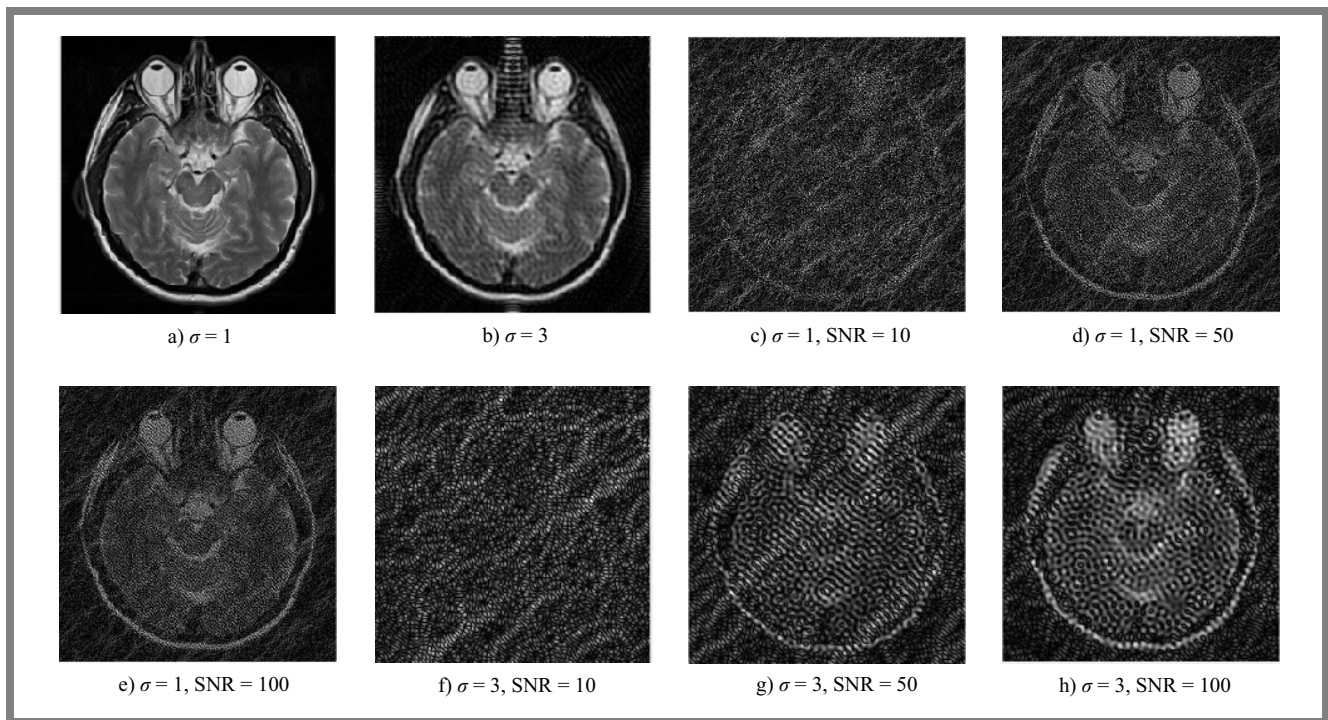


Fig. 6. Results of image restoration for different σ and SNR values using the Wiener filter.

since the increasing iteration causes increase the effect of resonance on the image. In our research, two new adaptations have been introduced. These two adaptations are done by introducing two parameters into the original RL algorithm:

- the first additive term represents the priori restored image (restored image from the previous iteration), through the

iteration process of specified weighting factor β , in order to reduce the number of iterations,

- second additive term represents the strengthening of the edges of the images, during the restoration process of the specified weighting factor λ .

The final RL filter after adding the above two new parameters can be described as:

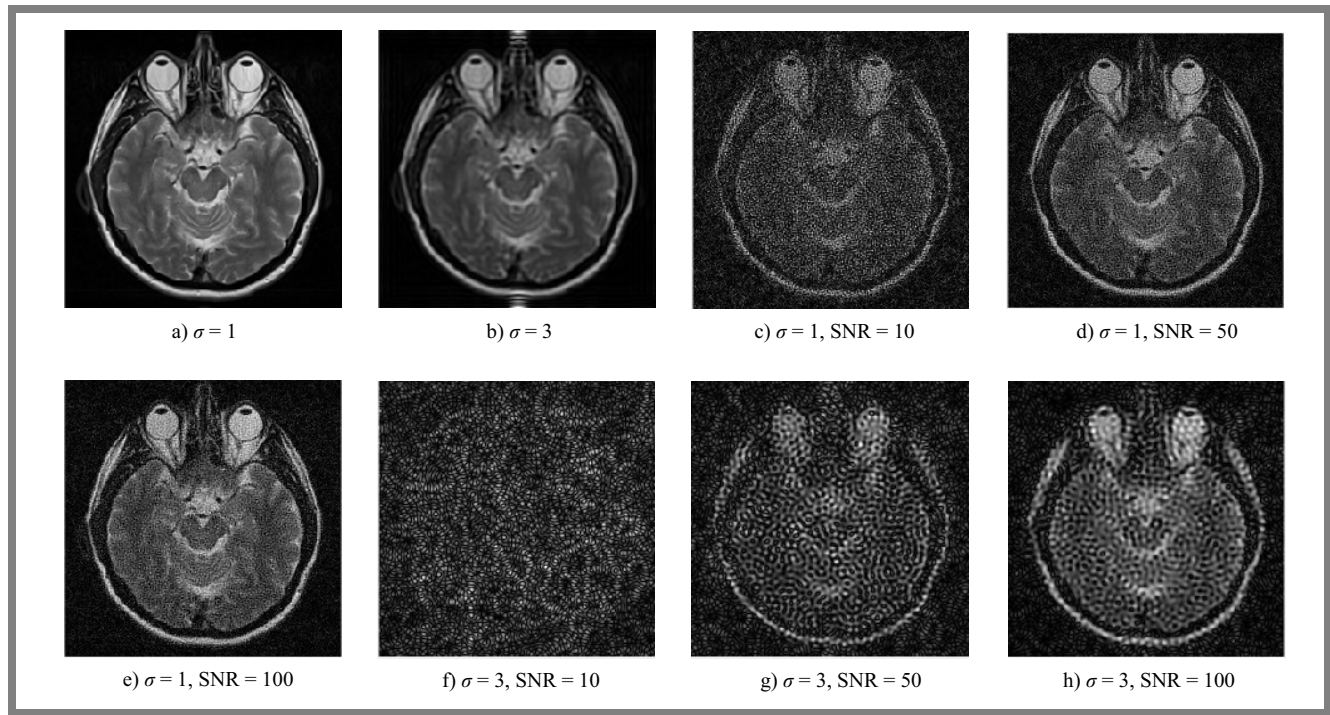


Fig. 7. Results of restoration of a CT scan image for different σ and SNR values by using the CLSF.

$$\begin{aligned} \hat{f}_{(i+1)}(x, y) = & \left[\frac{g(x, y)}{h(x, y) \otimes f_{(i)}(x, y)} \otimes h^T(x, y) \right] f_{(i)}(x, y) \\ & + \beta f_{(i)}(x, y) + \lambda (f_{(i)}(x, y) \otimes s(x, y)) \quad (11) \end{aligned}$$

where $s(x, y)$ represents the Sobel operator.

Algorithm 1 shows the steps of the restoration filter by using an adaptive RL filter.

Algorithm 1. Adaptive Richardson-Lucy filter.

Input: degraded image, estimated point spread function (PSF), number of iteration (I).

Output: restored image $\hat{f}_{(i+1)}(x, y)$

- 1: Set value of $f_{(i)}(x, y) = g(x, y)$ in the first iteration
- 2: Calculate the transpose of matrix $h^t(x, y)$
- 3: Choose γ and β within the range 0 to 1 with an increment of 0.25
- 4: For $i = 1, 2, 3, \dots, I$
- 5: Convolve the estimation PSF $h(x, y)$ with the previous estimate value of $f_{(i)}(x, y)$ to produce a reblurred image:

$$h(x, y) \otimes f_{(i)}(x, y)$$

- 6: Divide the blurred image $g(x, y)$ by the value from step 5
- 7: Convolve the value from step 6 with transpose of matrix $h^t(x, y)$
- 8: Multiply the value from step 7 with the previous estimated value of $f_{(i)}(x, y)$
- 9: Multiply β with the previous estimated value of $f_{(i)}(x, y)$:

$$\beta f_{(i)}(x, y)$$

- 10: Add the value from step 8 to the value from step 9

- 11: Apply Sobel edge detection to the previous estimated value of $f_{(i)}(x, y)$:

$$f_{(i)}(x, y) \otimes s(x, y)$$

- 12: Multiply λ with the value from step 11

- 13: Add the value from step 10 to the value from step 12 to $\hat{f}_{(i+1)}(x, y)$ calculate the new estimate of restored image

- 14: Repeat until iteration $>$ maximum

- 15: Return $\hat{f}_{(i+1)}(x, y)$

Figure 8 represents the effect of a restored CT scan image for different λ and β values using various filters in the following order: a) non-blurred, b) Gaussian blurred with $\sigma = 3$, SNR = 10, c) restored by inverse filter, d) restored by Wiener filter, e) restored by CLSF, f) restored using R-L method, g) restored using R-L $\lambda = 0.5$, h) restored using modified R-L method; $\lambda = 0.5$, $\beta = 0.25$, and i) restored using modified R-L method, $\lambda = 0.25$, $\beta = 1$.

4.5. Performance Evaluation

To demonstrate and evaluate the performance of the image restoration and enhancement techniques, several metrics have been used. First, we simulate the degraded images with different cases, i.e. Gaussian for blurring, different degrees of blurring, and different standard deviation “widths” of PSF, $\sigma = 1$, and 3, and so on. We also added white Gaussian noise of varying intensities to the noise function, i.e. different degrees of noise variance, different SNR = 5, 10, and 100. Figure 5 shows the original and degraded image using various degradation parameters.

Three different metrics have been used to validate the degree of similarity and loss of quality between the original image

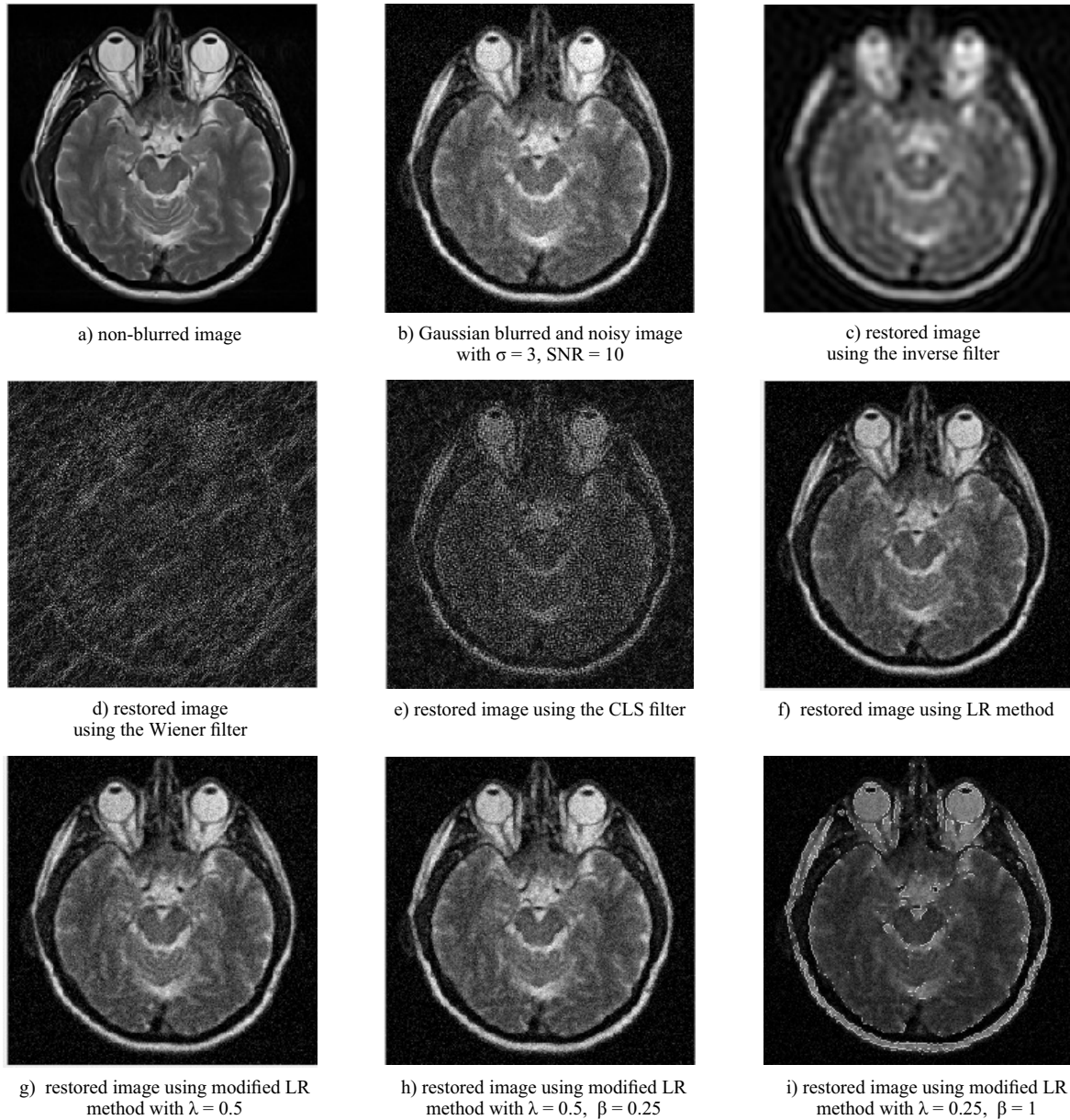


Fig. 8. Results of the restored CT scan image for $\sigma = 1$, $\text{SNR} = 10$.

and the degraded (blurred and noisy) image: peak signal to noise ratio (PSNR), structural similarity index (SSIM), and root mean square error (RMSE). A reference metric that requires both a reference image and a processed image to be acquired from the same image is called a full reference metric. These metrics are given by [28]:

$$\text{PSNR} = 20 \log \frac{\text{MAX}_f}{\sqrt{\text{MSE}}}, \quad (12)$$

where MAX_f is its highest possible pixel value, i.e. 255 for 8 bits per sample, and MSE stands for root mean square error:

$$\text{SSIM}_{(x,y)} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (13)$$

where: x and y are windows of similar size, μ_x , μ_y are the means of x and y , respectively, variances of x and y are σ_x^2 and σ_y^2 , respectively, and covariance between x and y is denoted by σ_{xy} . To keep the stability of computation, constants C_1 and C_2 are utilized. By default: $C_1 = (k_1l)^2$, $C_2 = (k_2l)^2$, where l is the pixel values' dynamic range (255 for 8-bit), $k_1 = 0.01$, $k_2 = 0.03$:

$$\text{RMSE} = \sqrt{\frac{1}{MN} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} [f(x,y) - g(x,y)]^2}, \quad (14)$$

where: $f(x,y)$ is the original image, $g(x,y)$ is the degraded image, while N and M represent the image dimensions.

Tab. 1. RMSE of restored images with a different type of filter.

Filter type	STD σ	RMSE, no noise	RMSE, SNR = 10	RMSE, SNR = 50	RMSE, SNR = 100
Direct inverse filter	1	0.511	0.511	0.511	0.511
	3	0.544	0.544	0.544	0.544
Wiener filter	1	0.508	0.495	0.506	0.507
	3	0.544	0.526	0.538	0.541
Constrained least squares filter	1	0.508	0.506	0.508	0.508
	3	0.544	0.526	0.538	0.540
R-L filter	1	0.196, 2 iter.	0.263, 1 iter.	0.237, 1 iter.	0.207, 2 iter.
	3	0.269, 15 iter.	0.293, 13 iter.	0.268, 15 iter.	0.267, 15 iter.
Adaptive R-L filter $\lambda = 0.5$	1	0.195	0.263	0.217	0.206
	3	0.256	0.295	0.262	0.261
Adaptive R-L filter $\lambda = 0.5, \beta = 0.25$	1	0.252	0.261	0.256	0.251
	3	0.306	0.290	0.311	0.315
Adaptive R-L filter $\lambda = 0.25, \beta = 1$	1	0.341	0.364	0.260	0.343
	3	0.356	0.380	0.310	0.371

Tab. 2. RMSE of images restored with the R-L filter, versus iteration number.

	$\sigma = 1$ no noise	$\sigma = 3$ no noise	$\sigma = 1$ SNR = 10	$\sigma = 1$ SNR = 50	$\sigma = 1$ SNR = 100	$\sigma = 3$ SNR = 10	$\sigma = 3$ SNR = 50	$\sigma = 3$ SNR = 100
1	0.212	0.304	0.263	0.237	0.216	0.323	0.304	0.304
2	0.196	0.295	0.267	0.244	0.207	0.313	0.294	0.294
3	0.207	0.289	0.281	0.262	0.212	0.308	0.288	0.288
4	0.221	0.285	0.293	0.280	0.228	0.304	0.284	0.284
5	0.231	0.282	0.295	0.295	0.249	0.301	0.281	0.281
6	0.246	0.279	0.316	0.307	0.264	0.300	0.279	0.278
7	0.266	0.277	0.333	0.320	0.289	0.298	0.277	0.276
8	0.282	0.275	0.356	0.331	0.305	0.297	0.275	0.274
9	0.293	0.274	0.366	0.341	0.315	0.295	0.274	0.272
10	0.299	0.273	0.376	0.350	0.308	0.294	0.273	0.271
11	0.313	0.272	0.384	0.358	0.317	0.294	0.272	0.270
12	0.322	0.271	0.389	0.365	0.326	0.293	0.271	0.269
13	0.336	0.270	0.395	0.372	0.341	0.293	0.270	0.268
14	0.349	0.269	0.402	0.378	0.352	0.293	0.269	0.267
15	0.356	0.269	0.407	0.383	0.358	0.293	0.268	0.267

Table 1 shows the performance evaluation results with varying width σ of the blurring function, for different values of real widths.

The variation of RMSE in filters presented in Tab. 1 can be summarized as:

- RMSE of degraded images in (Wiener, CLSF, R-L) filters decreases with increasing SNR and increases with increasing σ of the Gaussian blurring function,
- for an inverse filter, RMSE of degraded images increases with increasing σ of the Gaussian blurring function and is unaffected by SNR variation,

- RMSE of the restored image is lower than that of the inverse filter when using (Wiener and CLSF) filters,
- the R-L filter reduces RMSE of the restored image more than Wiener and CLSF,
- RMSE of the restored image is lower than RMSE of all other filters when using an adaptive R-L with $\lambda = 0.5$,
- RMSE of the image restored using an adaptive R-L filter with $\lambda = 0.5$ and $\beta = 0.25$ is lower than the R-L filter with SNR = 10.

Tab. 3. Metrics of images restored with the adaptive R-L filter with various λ, β .

$\lambda; \beta$	σ, SNR	$\sigma = 1$ no noise	$\sigma = 3$ no noise	$\sigma = 1$ SNR = 10	$\sigma = 1$ SNR = 100	$\sigma = 3$ SNR = 10	$\sigma = 3$ SNR = 100
0.25, 0	RMSE	0.19	0.25	0.26	0.20	0.29	0.25
	PSNR	28.39	23.51	23.21	27.24	21.17	23.45
	SSIM	0.89	0.76	0.48	0.76	0.47	0.66
0.5, 0	RMSE	0.19	0.25	0.26	0.20	0.29	0.26
	PSNR	28.39	23.61	23.15	27.40	21.20	23.32
	SSIM	0.89	0.77	0.49	0.76	0.46	0.66
1, 0	RMSE	0.19	0.26	0.26	0.20	0.29	0.26
	PSNR	28.51	23.20	23.25	27.35	21.24	22.85
	SSIM	0.89	0.76	0.48	0.76	0.47	0.66
0.25, 0.25	RMSE	0.25	0.30	0.28	0.25	0.32	0.30
	PSNR	23.92	20.46	21.97	23.63	19.70	20.55
	SSIM	0.86	0.70	0.43	0.71	0.38	0.60
0.5, 0.25	RMSE	0.25	0.30	0.26	0.25	0.29	0.31
	PSNR	23.91	20.54	22.01	23.97	19.51	20.01
	SSIM	0.86	0.69	0.43	0.71	0.37	0.59
1, 0.25	RMSE	0.26	0.31	0.28	0.25	0.33	0.32
	PSNR	23.76	19.96	21.80	23.81	19.08	19.76
	SSIM	0.86	0.70	0.42	0.71	0.36	0.58
0.25, 0.5	RMSE	0.19	0.26	0.28	0.20	0.32	0.31
	PSNR	22.91	20.24	21.59	23.20	19.29	20.17
	SSIM	0.83	0.69	0.44	0.70	0.37	0.60
0.5, 0.5	RMSE	0.26	0.31	0.29	0.26	0.33	0.31
	PSNR	23.37	20.20	21.28	23.33	19.25	20.00
	SSIM	0.84	0.69	10.73	0.71	0.39	0.60
1, 0.5	RMSE	0.26	0.32	0.28	0.26	0.32	0.31
	PSNR	23.40	19.73	21.79	23.14	19.40	19.90
	SSIM	0.85	0.69	0.44	0.70	0.37	0.59
0.25, 1	RMSE	0.34	0.35	0.36	0.34	0.38	0.37
	PSNR	18.67	17.93	17.53	18.56	16.79	17.20
	SSIM	0.76	0.66	0.48	0.69	0.44	0.60
0.5, 1	RMSE	0.31	0.34	0.34	0.32	0.36	0.35
	PSNR	19.83	18.72	18.51	19.71	17.43	18.0
	SSIM	0.79	0.67	0.47	0.70	0.43	0.61
1, 1	RMSE	0.28	0.32	0.31	0.29	0.34	0.33
	PSNR	21.57	19.69	20.20	21.47	18.32	19.04
	SSIM	0.81	0.68	0.49	0.72	0.43	0.60

RMSE summary for the R-L filter is shown in Tab. 2. The commentary is as follows:

- RMSE of degraded images increases along with increasing σ of the Gaussian blurring function and decreases with increasing SNR,
- the number of iterations increases with increasing the degradation parameters.

The restoration iteration is achieved when no change may be detected between the current restored image and the pre-

vious one. In such a situation, the repetition process is terminated.

5. Conclusion and Future Work

The results from simulation show that when using a small value of parameters λ and β , the adaptive R-L filter performs better in image quality restoration, as the Gaussian blurring function's standard deviation values and SNR increase with each iteration. The experimental results show a decrease in

the RMSE value of up to 0.25 when using $\lambda = 0.5$ in one iteration, compared to other existing methods, for example the Wiener filter, where RMSE values reach 0.5442. This demonstrates the effectiveness of our proposal in terms of improved accuracy of the medical image by adapting the R-L algorithm, but the results were not satisfactory when using large values of the parameters λ and β . We will continue work to improve the accuracy of the solution and to solve that problem in the future. It is also possible to use another type of the point spread function (PSF) and noise that can be applied in conjunction with the proposed framework in the near future.

References

- [1] P. Gupta and R. Mehra, "Blind restoration method for satellite images using memetic algorithm", *International Journal of Computer Applications*, vol. 130, no. 1, pp. 20–25, 2015 (<https://doi.org/10.5120/ijca2015906857>).
- [2] R. Mishra, N. Mittal, and S.K. Khatri, "Digital image restoration using image filtering techniques", *International Conference on Automation, Computational and Technology Management (ICACTM)*, London, UK, 2019, (<https://doi.org/10.1109/ICACTM.2019.8776813>).
- [3] P. Ganesan and V. Rajini, "Comparative study of denoising methods for satellite image restoration using Matlab", *International Journal of Advanced Research in Computer Science*, vol. 5, no. 4, pp. 74–77, 2013 (<http://www.ijarcs.info/index.php/Ijarcs/article/view/1685/1673>).
- [4] M.K. Kadhom, A.A. Al-Ani, and S.A. Saleh, "Image restoration using adaptive nonlinear techniques", M.Sc. thesis, Al-Nahrain University, Baghdad, Iraq, 2008.
- [5] M.A. Kadhim, Ed., "Restoration medical images from speckle noise using multifilters", *7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2021 (<https://doi.org/10.1109/ICACCS51430.2021.9441814>).
- [6] G.S. Karam, Z.M. Abood, H.H. Kareem, and H.G. Dowy, "Blurred image restoration with unknown point spread function", *Al-Mustansiriyah Journal of Science*, vol. 29, no. 1 pp. 189–194, 2018 (<https://doi.org/10.23851/mjs.v29i1.335>).
- [7] K. Panfilova and S. Umnyashkin, "Linear blur compensation in digital images using Lucy-Richardson method", *2016 IEEE NW Russian Young Researchers in Electrical and Electronic Engineering Conference (EIConRusNW)*, 2016 (<https://doi.org/10.1109/EIConRusNW.2016.7448179>).
- [8] O. Anaconda-Mosquera, J. Arias-García, D.M. Muñoz, and C.H. Llanos, "Efficient hardware implementation of the Richardson-Lucy algorithm for restoring motion-blurred image on reconfigurable digital system", *29th Symposium on Integrated Circuits and Systems Design (SBCCI)*, Belo Horizonte, Brazil, 2016 (<https://doi.org/10.1109/SBCCI.2016.7724056>).
- [9] A. Tselousov and S. Umnyashkin, "Kernel estimate for image restoration using blind deconvolution", *2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, St. Petersburg and Moscow, Russia, 2017 (<https://doi.org/10.1109/EIConRus.2017.7910667>).
- [10] F. Aouinti, N. M'barek, M.M. Nasri, M. Moussaoui, and B. Bouali, "An improved Richardson-Lucy algorithm based on genetic approach for satellite image restoration", *International Arab Journal of Information Technology*, vol. 15, no. 4, 2018 (<https://iajit.org/portals/PDF/July%202018,%20No.%204/11791.pdf>).
- [11] K. Panfilova and S. Umnyashkin, "Correlation-based quality measure for blind deconvolution restoration of blurred images based on Lucy-Richardson method", *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Saint Petersburg and Moscow, Russia, 2019 (<https://doi.org/10.1109/EIConRus.2019.8657324>).
- [12] D. Liu, X. Chen, and X. Liu, "An improved Richardson-Lucy algorithm for star image deblurring", *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Auckland, New Zealand, 2019 (<https://doi.org/10.1109/I2MTC.2019.8826905>).
- [13] K. Breykina and S. Umnyashkin, "Correlation-based iterative estimation of distortion kernel for image restoration", *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, St. Petersburg and Moscow, Russia, 2020 (<https://doi.org/10.1109/EIConRus49466.2020.9039119>).
- [14] B. Zhao, X. Chen, G. Feng, X. Zhao, and J. Jiang, Eds., "An improved LR algorithm for image deblurring", *2020 Chinese Control and Decision Conference (CCDC)*, Hefei, China, 2020 (<https://doi.org/10.1109/CCDC49329.2020.9164762>).
- [15] D. Bhonsle, "Quality improvement of Richardson Lucy Based deblurred images using krill herd optimization", *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, Bhilai, India, 2021 (<https://doi.org/10.1109/ICAECT49130.2021.9392469>).
- [16] D. Lyu, J. Tian, H. Hu, and X. He, "Ultrasonic c-scan image restoration method using the Richardson-Lucy algorithm and a flaw measurement model", *Applied Acoustics*, vol. 200, 109074, 2022 (<https://doi.org/10.1016/j.apacoust.2022.109074>).
- [17] Y. Li et al., "Incorporating the image formation process into deep learning improves network performance", *Nature Methods*, vol. 19, pp. 1427–1437, 2022 (<https://doi.org/10.1038/s41592-022-01652-7>).
- [18] H. Joshi and D.J. Sheetlani, "Image restoration techniques in image processing: An illustrative review", *International Journal of Advance Research in Science and Engineering*, vol. 6, no. 1, pp. 145–158, 2017 (http://www.ijarse.com/images/fullpdf/1512719410_1023_IJARSE.pdf).
- [19] N.K. El Abbadi, A.H. Abdulkhaleq, and S.A. Al Hassani, "A survey on blind de-blurring of digital image", *Iraqi Journal of Science*, vol. 63, no. 1, pp. 338–352, 2022 (<https://doi.org/10.24996/ijss.2022.63.1.32>).
- [20] D. Fish, A. Brinicombe, E. Pike, and J. Walker, "Blind deconvolution by means of the Richardson-Lucy algorithm", *Journal of the Optical Society of America A*, vol. 12, no. 1, pp. 58–65, 1995 (<https://doi.org/10.1364/JOSAA.12.000058>).
- [21] R.S.H. Al-Taweel and H.G. Daway, "Deblurring average blur by using adaptive Lucy Richardson", *Journal of College of Education*, no. 5, pp. 75–90, 2015 (<https://www.iasj.net/iasj/download/dd42a7cdacecf1d6>).
- [22] M. Juneja and P.S. Sandhu, "Performance evaluation of edge detection techniques for images in spatial domain", *International Journal of Computer Theory and Engineering*, vol. 5, no. 1, pp. 614–621, 2009 (<https://doi.org/10.7763/IJCTE.2009.V1.100>).
- [23] T. Sanida, A. Sideris, and M. Dasygenis, Eds., "A heterogeneous implementation of the Sobel edge detection filter using OpenCL", *2020 9th International Conference on Modern Circuits and Systems Technologies (MOCASST)*, Bremen, Germany, 2020 (<https://doi.org/10.1109/MOCASST49295.2020.9200249>).
- [24] A.H. Muhammad and H.S. Akbar, "Algorithms for edge detection by using fuzzy logic technique", *Kirkuk University Journal-Scientific Studies*, vol. 10, no. 1, pp. 173–190, 2015 (<https://doi.org/10.32894/kujss.2015.101966>).
- [25] C. Saravanan, "Color image to grayscale image conversion", *2010 Second International Conference on Computer Engineering and Applications*, Bali, Indonesia, 2010 (<https://doi.org/10.1109/ICCEA.2010.192>).
- [26] S. Patro and K.K. Sahu, "Normalization: A preprocessing stage", *International Advanced Research Journal in Science, Engineering and Technology*, vol. 2, no. 3, pp. 20–22, 2015 (<https://doi.org/10.17148/IARJSET.2015.2305>).
- [27] M.H. Muhson and A.A. Al-Ani, "Blind restoration using convolution neural network", *Iraqi Journal of Information and Communication Technology*, vol. 1, no. 1, pp. 25–32, 2021 (<https://doi.org/10.31987/ijict.1.1.178>).
- [28] M.A. Abdulrahman and A.H. Al-Fayadh, "Modifications with applications of some transforms in image processing", M.Sc. thesis, Al-Nahrain University, Baghdad, Iraq, 2018.

Qunoot A. Yaqoub

M.Sc. student at the Department of Information and Communication Engineering

E-mail: mohqunoot@gmail.com

College of Information Engineering, Al-Nahrain University, Baghdad, Iraq

Ayad A. Al-Ani, Prof.

Professor of Digital Image Processing at the Department of Information and Communication Engineering

E-mail: ayad.a@nahrainuniv.edu.iq

College of Information Engineering, Al-Nahrain University, Baghdad, Iraq

Preserving Integrity of Evidence with Blockchain Technology in Cloud Forensics for Immigration Management

Sahadev Maruti Shinde and Venkateswara Rao Gurralla

GITAM School of Engineering, Visakhapatnam, Andhra Pradesh, India

<https://doi.org/10.26636/jtit.2023.164522>

Abstract — As the popularity of cloud computing increases, safety concerns are growing as well. Cloud forensics (CF) is a smart adaptation of the digital forensics model that is used for fighting the related offenses. This paper proposes a new forensic method relying on a blockchain network. Here, the log files are accumulated and preserved in the blockchain using different peers. In order to protect the system against illegitimate users, an improved blowfish method is applied. In this particular instance, the system is made up of five distinct components: hypervisor (VMM), IPFS file storage, log ledger, node controller, and smart contract. The suggested approach includes six phases: creation of the log file, key setup and exchange, evidence setup and control, integrity assurance, agreement validation and confidential file release, as well as blockchain-based communication. To ensure efficient exchange of data exchange between the cloud provider and the client, the methodology comprises IPFS. The SSA (FOI-SSA) model, integrated with forensic operations, is used to select the keys in the best possible way. Finally, an analysis is conducted to prove the effectiveness of the proposed FOI-SSA technique.

Keywords — cloud computing, cloud forensics, FOI-SSA model, improved blowfish

1	2
HSO	Harmony search optimization
ECC	Elliptical curve cryptography
IPFS	Interplanetary file system
LGoE	Logical graph of evidence
LA	Lion algorithm
RSA	Rivest-Shamir-Adleman
SRVA	Secure ring verification based authentication
SIEM	Security data information and event management
SA-DECC	Sensitivity aware deep ECC
SSA	Sparrow search algorithm
SSO	Salp swarm optimization
SRVA	Secure ring verification based authentication
TPS	Transaction per second

Tab. 1. Abbreviations and terms used in this paper.

Abbreviation	Description
1	2
AES	Advanced encryption standard
BlockSLaaS	Blockchain assisted secure logging as-a-service
BES	Bald eagle search
CF	Cloud forensics
CC	Cloud computing
CSP	Cloud service provider
CFI	Cloud forensic investigator
CADF	Cloud auditing data federation
DMTF	Distributed management task force
DBO	Dynamic butterfly optimization
EB	Ethereum blockchain
FCS	Fuzzy based smart contracts
FIO	Forensic investigation optimization

1. Introduction

When individuals leave their countries and move to other states, we are dealing with migration [1]–[5]. Such persons go through immigration-related processes in order to become permanent residents of the their new country. Usually, the procedure is very complicated. The applicant needs to get a visa [1] and then apply for a permanent residency permit which may later be converted into citizenship [6]–[11]. This process becomes easier if the applicant is backed by a company or if their family member is already a resident of the country concerned [12]–[14]. Due to the strict immigration laws in effect in some countries, people revert to illegal practices and attempt to infiltrate states without permission [15], [16]. This leads to illegal immigration [17]–[19] – an issue faced by almost every other country in the world [20]–[22].

The proposed work keeps track of immigrants by storing several relevant pieces of information in the form of immutable [6] and unique blockchain records [23], [24]. When a person is suspected of illegal immigration, their official documents are

compared with the record stored using blockchain [25]–[27], i.e. a distributed ledger [28], [29].

Blockchain is a network of fault-tolerant and distributed servers that contain shared, duplicated, and distinct content [30]. The management of a blockchain is cheap and quick, since it is immutable and cannot contain false or duplicate information [31]–[33]. Blockchain can process fingerprints, facial recognition, and retinal scanning biometric data [17], [18]. Blockchain-based reactive data can be secured using protective confidentiality encryption, limiting its use to authorized entities only [34].

The novelty of this work lies in the fact that a novel blockchain-based CF scheme is proposed, where an improved blowfish mechanism is deployed for encryption purposes, and in exploiting the FOI-SSA algorithm for creating an optimal key. The paper is set up as follows. After the related works review given in Section 2, in Section 3 the project is presented and the model created is described. The FOI-SSA model recommended for generating the best key is described in Section 4. Sections 5 presents the conclusions.

2. Literature Review

This section surveys the eight existing blockchain-dependent evidence integrity preservation methods used in CF. Rane and *et al.* [1] proposed the forensic-aware BlockSLaaS model to steadily process and store logs by addressing multi-stakeholder collusion issues and facilitating confidentiality and integrity. CFI was capable of accessing logs for forensic purposes, using BlockSLaaS to protect the logs' privacy. However, the method failed to validate whether the service provider guaranteed precise logs. Jain *et al.* [2] presented a blockchain method for preserving the integrity of log files. IPFS and the blockchain technology were combined to transform a centralized storage system into a decentralized one. The integrity of log files was preserved by storing log files in blockchain. Thanks to such an approach, the system was used for storing huge log files at a minimal cost. However, the method failed to maximize CSP trust by minimizing CSP dependencies.

Pourvhab *et al.* [3] presented an SRVA scheme for protecting the system against unauthorized users. To ensure even better protection of the cloud platform, the secret keys were optimally produced by utilizing the HSO technique. In this case, the server stored the data after they had been encrypted with the use of the SA-DECC algorithm. By modifying FCS, such a strategy allowed the user to track down data and LGoE collected with the use of blockchain enabled the evidence to be studied. However, the proposed method failed to improve the digital forensics model. Dalezios *et al.* [4] proposed the DMTF with CADF standard for CF. The authors improved the Apache Cloud Stack platform by employing CADF activity tracking adopted in an Open Stack and made it more forensically reverberant. Stelly *et al.* [5] developed a method relying on automated container deployment and orchestration platforms to attain improved performance in digital forensics.

The results showed that the distributed container-based approach offered a workable technical foundation for addressing the increased data volumes in digital forensic investigations.

Park *et al.* in [6] presented a permission blockchain-based data integrity management system for CFs. Such a method was capable of certifying the integrity of data while processing more transactions. However, there is an issue that the evaluation of performance cannot be made on anticipated data dimension. However, the model can be utilized as one of the methods for addressing security-related issues in cloud platforms. It failed, however, to accumulate network data by performing simulations concerned with computing precise TPS. Dasaklis *et al.* [7] described a CF method relying on the available blockchain-based technologies. The approach provided a detailed review of the various advantages and shortcomings of the mutually beneficial relationship between blockchain technology and the current digital forensics approach. Unfortunately, the method failed to identify different research issues in digital forensics. Irfan *et al.* [8] presented a model using SIEM to address the problem of effective evidence collection in CFs. The method shared evidence with cloud users, whenever needed. The proposed method helped perform detailed CF by adapting evidence, but failed to improve the performance of the solution by applying advanced optimization techniques.

3. Blockchain-based Protocol Developed for Maintaining Integrity in CF

Anti-tampering and privacy protection are two critical security requirements in cloud computing environments. Figure 1 shows the outline of the proposed architecture. In judicial forensics, maintaining privacy is a top priority. The suggested technique adopts an appropriate mechanism for maintaining confidentiality and anonymity, ensuring that no private data is released during the derivation function of a blockchain-based process. The system incorporates eight elements, including hypervisor, virtual machine, node controller, log ledger, IPFS file storage, blockchain network, CFI, and smart contract.

3.1. Initialization Step

The start-up phase involves launching virtual machines, hypervisors, node controllers, IPFS cloud storage, smart contracts, blockchain networks, CFI, and log ledgers. The following is a more detailed depiction of each entity. A virtual machine (VM) is a computational source that runs programs maps using software, rather than a real computer. The hypervisor is software that creates and operates a collection of virtual machines, allowing one host to handle several guest VMs, by sharing resources virtually. The nodes controller gathers logs from all virtual platform sources via log libraries and creates log entries for each log. IPFS cloud storage is a file transfer mechanism depending on cryptography hashes, that can be readily stored on the blockchain and regulated to effectively store and transfer large files, while smart contract acts as a set

of applications that are kept on the blockchain and continue to run when specific conditions are satisfied.

The blockchain network, in turn, offers ledger and smart contract functions to varied apps, and if questionable actions on the cloud take place, the CFI is tasked with gathering and reviewing evidence. The last resource component is the log ledger which contains a set of recorded results with a timestamp. Therefore, it serves as a helpful proof for initiating legal action against a suspect. The ledger aids in the preservation of the chronology of created logs.

Table 1 contains all acronyms and abbreviations used in this paper, while Tab. 2 summarizes the symbols used.

Tab. 2. Symbols used in of proposed evidence integrity preservation mechanism.

Symbol	Description
M_{PWD}	Password of node controller
M_{ID}	User ID of node controller
K_M	Key of node controller
T	Time stamp
Hd	Host ID
L	Log file
P	Node controller program to record log file
pk	Public key
$\otimes$	Interpolation
$\oplus$	Ex-or operation
R_{Hd}	Requested ID
$\text{en}(\cdot)$	ECC encryption
S_M	Service name
$K(\cdot)$	Kernel transform
r	Random number
REQ	Integrity assurance request message
$h(\cdot)$	Hashing
S_{PWD}	Session password
K_L	Hypervisor key
A	Acknowledgement message
Q_{PWD}	CFI Password
Q_{Hd}	CFI ID
IA	Integrity assurance

3.2. Creation of Log File

At this stage, the user password M_{Hd} and user ID M_{PWD} are formed by the user, which accumulating logs from every each resource of the virtual podium. M_{Hd} and M_{PWD} are saved in the hypervisor as M_{Hd}^* and M_{PWD}^* . The node controller key is produced after obtaining the user's credentials. The key is created by XOR-ing the public key and the modulus of a random key. After combining the resulting product with the encrypted user ID, the modulus is used to generate the node controller key as:

$$K_M = |\text{en}(M_{Hd}^*) \bmod (r) \oplus pk|. \quad (1)$$

The Log L is formed in the hypervisor with the requested ID of user R_D , time stamp T , and service name S_M for Hd , as:

$$L = \langle Hd, T, R_D, S_M \rangle. \quad (2)$$

Similarly, the node controller key K_M , L , and encoded node controller programmes for recording the log files are provided to the node controller as:

$$K_M, L, \text{en}(P). \quad (3)$$

Thus, the node controller key is saved as K_M^* and the node controller program is set to trace the log file.

3.3. Key Setup and Exchange Process

Once the log file generation procedure is has been completed, three entities: the node controller, hypervisor, and log ledger, are used to start the key setup and exchange process. The stored node controller keys are given to the hypervisor and saved as K_M^R . The hypervisor key is created by adding the encrypted node controllers programs and the kernel transformation of an arbitrary integer with the stored key. The final hypervisor key is generated by XOR-ing the acquired result with the hash timestamp Ts , which is modelled as:

$$K_L = h(Ts) \oplus \text{en}(p) || K(r). \quad (4)$$

The hash of the finalized hypervisor key is proceeded here to generate the session password S_{PWD} , which is then supplied to the node controller and saved as S_{PWD}^* . As a result, the session password is:

$$S_{\text{PWD}} = h(K_L). \quad (5)$$

Next, the acquired hypervisor key is saved as K_L^* in the log ledger. The hash of the stored hypervisor key is saved in M_1 and passed to the node controller, where it is saved as M_1^* :

$$M_1 = h(K_L^*). \quad (6)$$

If $M_1^* = S_{\text{PWD}}$, the ledger is confirmed and the confirmation occurs in the node controller.

3.4. Evidence and Contract Phase

For the sake of achieving privacy, the CFI and smart contract are used. The CFI ID Q_{Hd} and password Q_{PWD} are created during this phase and the credentials are passed into a the smart contract that is saved as Q_{Hd}^* and Q_{PWD}^* . Furthermore, the acknowledgment packet containing an encoded node control program, an the encoded ID, the cloud hashing forensic researcher ID, and the header are given to CFI and saved as A^* :

$$A = \langle \text{en}(P), \text{en}(R), h(Q_{Hd}), \text{Header} \rangle. \quad (7)$$

The evidence is formed by XOR-ing the hash message and secure CFI ID, which is kept in the smart contract as EP^* and modelled as:

$$EP = h(A^*) \oplus \text{en}(Q, D). \quad (8)$$

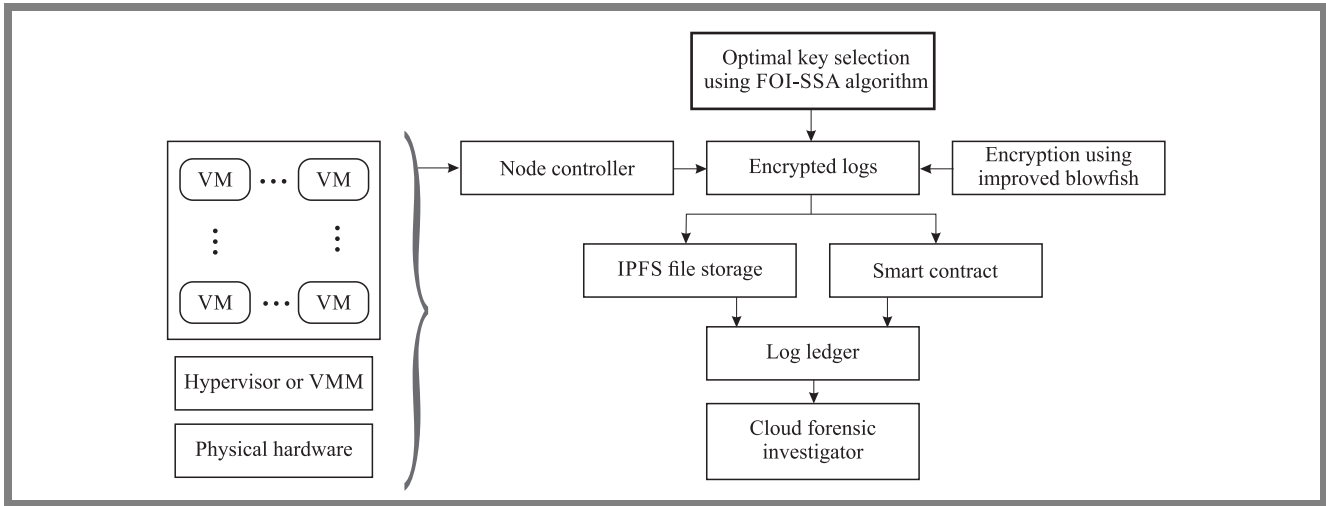


Fig. 1. Overall system model block diagram for evidence integrity preservation for cloud forensics.

The confirmation message is created by XOR-ing the saved evidence with the hashed message that is written as:

$$V = EP^* \oplus h(A). \quad (9)$$

If $V = \text{en}(Q_{Hd})$, then the validation was is considered accomplished and the information was is sent to CFI.

3.5. Integrity Assurance

Three entities are involved in this process: node controller, hypervisor, and CFI. The user ID, the request message and the stored key of the node controller are initially verified by the hypervisor that is expressed as:

$$M_{Hd}, REQ, K_M^*. \quad (10)$$

The user ID and the saved node controller key are verified by the hypervisor. Then, the message is created by XOR-ing the hash node controller key, the encoded shared key, and the hash hypervisor key, such as:

$$M_1 = h(K_N^*) \oplus \text{en}(\text{pk}) \oplus h(K_L). \quad (11)$$

The login password is formed by XOR-ing the hashed times-tamp with the encoded public key as:

$$S_{PWD} = h(Ts) \oplus \text{en}(\text{pk}). \quad (12)$$

The resultant message and session passcode are given into the CFI, which performs integrity assurance operations by XOR-ing the message and session passwords and storing the result in the hypervisor as:

$$IA = (M_1^* \oplus S_{PWD}). \quad (13)$$

Level 2 verification is calculated based on integrity security operations, using the saved integrity assurance and the saved message as:

$$V_2 = I_A^* \oplus M_1^*. \quad (14)$$

If $V_2 = S_{pwd}$, the guaranty is permitted. The required message and the CF ID are then sent to the node controller where the CF ID is saved.

3.6. Agreement and Confidential File Release

For agreement and private file release, the node controller, the CFI, and the blockchain network are considered. The CF ID and the request message are supplied to the node controller in this phase. Validation of the CF ID is carried out here in order to determine which one is genuine. Furthermore, two messages are produced, the first being formed by XOR-ing the encoded log and the encoded public key as:

$$R = \text{en}(\log) \oplus \text{en}(\text{pk}). \quad (15)$$

The other message is created by combining the public key with a value, and then combining the result with the hash Req message as:

$$R_G^* = H(\text{IREQ}) || (\text{pk} \Theta r). \quad (16)$$

The last message is created by XOR-ing the first and second messages, and then feeding them to the CFI and storing them as M_1^* , as shown in:

$$R^R = M_1^* \oplus R_G^Q. \quad (17)$$

The CFI creates two messages, the first of which is created by XOR-ing the last message with the second message. The second message is created by combining the public key with a randomized value, and then combining the result with the hash Req packet as:

$$R_G^Q = H(\text{IREQ}) || (\text{pk} \Theta r). \quad (18)$$

By XOR-ing the initial message with the encoded public key, the log is created. If $\langle S, P \rangle = I(\log)$, then it is found to be satisfactory and is sent again, and then a contact between the CFI and the blockchain network is created.

3.7. Improved Blowfish Algorithm

The blowfish scheme is highly efficient and is suitable for hardware implementation and related modeling [35]. However, to enhance the key management mechanism, a modified version of blowfish is introduced, as follows:

- the input includes 64 bit data,

- it includes 64-bit block ciphers with irregular key lengths,
- it includes four 32-bit S arrays and P boxes. The S array has 18 of 32-bit subkeys, while each P box comprising 256 entries,
- it comprises two elements: a key-expansion part and a data-encryption part.

The F operation employs four substitution boxes, each consisting 256 32-bit entries [36]. Conventionally, if block XL is divided to 8-bit blocks a, b, c, d , then the operation $F(XL)$ is shown as in Eq. (19). As per the modified blowfish model, $F(XL)$ is modelled as in Eq. (20):

$$F(XL) = [(P_{1,a} + P_{2,b} 2^{32}) \oplus P_{3,c}] + P_{4,d} 2^{32}, \quad (19)$$

$$F(XL) = [(P_{1,a} \oplus P_{3,c}) + (P_{2,b} \oplus P_{4,d}) 2^{32}]. \quad (20)$$

3.8. Blockchain-based Communication Phase

The blockchain-oriented communication stage is used to securely store and process logs, allowing for effective control of access to CFI and log integrity checking. An attack on the cloud can be carried out by a malevolent employee or an outside attacker. The functions that take place on the cloud platform generate logs for each VM operation, such as network interaction and VM setup logs. These logs are not power-independent, which means that if the VM is turned off, the data stored therein are lost. The suggested technique retrieves logs from Internet platforms and stores them in secondary memory storage to ensure data security and integrity.

The keys denoted by pk are optimally chosen via the FOI-SSA model. Figure 2 shows solutions in which nn indicates the overall count of keys. The objective Obj is to raise the key breaking time kbr as:

$$\text{Obj} = \max(kbr). \quad (21)$$

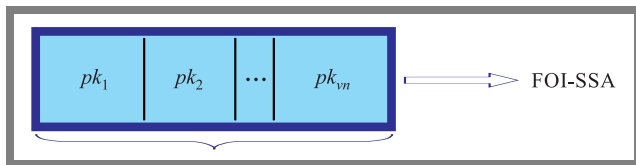


Fig. 2. Solution encoding scheme.

3.9. Proposed FOI-SSA Model for Optimal Key Generation

To achieve better convergence, FIO [37] is combined with the SSA model [38] to form FOI-SSA. Self-improvement of the optimization schemes results in better accuracy [39]–[42]. The behavior of the sparrows and formulated corresponding rules are described as:

- 1) The producers are typically highly energized. Assessment of each person's fitness values generates information about their energy reserves;
- 2) As individual sparrows start to chirp, the producers are required to direct all scavengers to the safe area when the alert value exceeds the safety level;
- 3) Every sparrow proceeds to production in accordance with how often it seeks out larger food sources, but the ratio of

scavengers to producers becomes higher across the board. The producers would act as he sparrows with maximum energy levels. Numerous starving scavengers are inclined to fly towards different locations in search for food, to gain energy;

- 4) Scroungers look for food by emulating a farmer who actually produces the healthiest food. To increase their predation rate, certain scavengers may keep a tight eye on the producers and engage in food wars;
- 5) Sparrows in the center of the group haphazardly walk to be close to others when the sparrows at the group's periphery are aware of danger and quickly go into the safe area to take a better position.

As per FOI-SSA, the chaotic-based OBL is performed to generate opposite solutions that ensure a good convergence rate.

The location of the sparrows is represented by:

$$Y = \begin{bmatrix} Y_{1,1} & Y_{1,2} & \dots & Y_{1,a} \\ Y_{2,1} & Y_{2,2} & \dots & Y_{2,a} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{s,1} & Y_{s,2} & \dots & Y_{s,a} \end{bmatrix}. \quad (22)$$

This implies the sparrow count and the size of the optimized variable. The fitness of the sparrow is defined by Eq. (23), which also addresses the fitness of the individuals.

$$F_Y = \begin{bmatrix} f([Y_{1,1} & Y_{1,2} & \dots & Y_{1,a}]) \\ f([Y_{2,1} & Y_{2,2} & \dots & Y_{2,a}]) \\ \vdots & \vdots & \ddots & \vdots \\ f([Y_{s,1} & Y_{s,2} & \dots & Y_{s,a}]) \end{bmatrix}. \quad (23)$$

The locations of producers are updated as per rules 1–2 and:

$$Y_{c,d}^{r+1} = \begin{cases} Y_{c,d}^r e^{\frac{-c}{\alpha \cdot \text{it}_{\max}}} & \text{if } C_2 < \text{st} \\ Y_{c,d}^r + P \cdot M & \text{if } C_2 \geq \text{st} \end{cases}. \quad (24)$$

In Eq. (24), r represents the iteration, max implies the maximum iteration, α is an arbitrary integer, st and C_2 are the safety threshold and the alarm value, and P is an arbitrary integer. M denotes a matrix of $1 \times d$ with element 1.

The scroungers follow rules 4–5. As stated earlier, various scroungers track producers. In FOI-SSA, the scrounger's position is updated using FIO as:

$$Y(i)_{\text{new}} = Y_{B_{ij}} + ra_{10}(Y_{B_{ij}} - Y_{B_{rj}}) + ra_{11}(Y_{\text{best}} - Y_{B_{ij}}), \quad (25)$$

where ra_{10} and ra_{11} are arbitrary integers (0 and 1), Y_{best} implies the best position, B_i denotes the agent. In addition, in FOI-SSA, Cauchy's mutation is performed as:

$$Y(i)_{\text{new}} = Y_{\text{best}} + Y_i^* \text{cauchy}(0, 1). \quad (26)$$

In addition, the model as per rule (6) is:

$$Y_{c,d}^{r+1} = \begin{cases} Y_{\text{best}}^r + \gamma \cdot |Y_{c,d}^r - Y_{\text{best}}^r| & \text{if } f_c < f_u \\ Y_{c,d}^r + Z \cdot \left(\frac{|Y_{c,d}^r - Y_{\text{worst}}^r|}{(f_c - f_w) + \varepsilon} \right) & \text{if } f_c = f_u \end{cases}, \quad (27)$$

where γ indicates the step size control parameter with a variance of 1 and a mean value of 0, R_{host} denotes the current global optimal location, $Z \in [-1, 1]$ denotes the route of the sparrow, f_c stands for the fitness value of the current sparrow, f_w and f_u are the worst fitness value and the current global best, ε is a small constant for avoiding the zero-division-error, and Y_{host} denotes the position at the center of the population.

4. Results and Discussion

The proposed CF integrity management plan has been created using Java and CloudSim. The performance of the FOI-SSA system was computed over EB [2], AES, ECC, RSA, El-Gamal, Signcryption, ECC + IPFS [43], DBO, BES, SSO, FIO, and SSA, taking into consideration such metrics as memory, detection rate, etc. In this case, the assessment was performed by altering the key size from 64 to 128 and the user count from 200 to 400.

4.1. Detection Rate Analysis

The detection accuracy of the proposed FOI-SSA algorithm is evaluated in comparison with traditional methods, for various key sizes of 64 and 128. Estimates concerning the FOI-SSA scheme, made over EB [2], AES, ECC, RSA, El-Gamal, Signcryption, ECC + IPFS [34], DBO, BES, SSO, FIO and SSA approaches are presented in Figs. 3 and 4 for user counts of 100, 200, 300 and 400. Here, the proposed FOI-SSA model showed an enhanced detection rate over EB, AES, ECC, RSA, El-Gamal, Signcryption, ECC + IPFS, DBO, BES, SSO, FIO and SSA. In Fig. 3, a higher detection rate is observed for

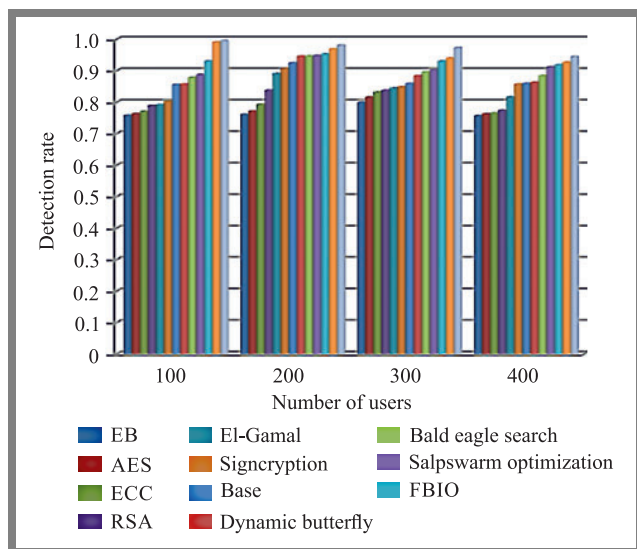


Fig. 3. Detection rate of FOI-SSA vs. other approaches, for a key size of 64.

FOI-SSA, with the user count of 100 for a key size of 64. With an increase in user count, the detection rates for FOI-SSA decreased for a key size of 64. This progression is the result of the enhanced blowfish concept and the integrated optimal key creation. Thus, the benefit of FOI-SSA is recognized over EB, AES, ECC, RSA, El-Gamal, Signcryption, ECC + IPFS, DBO, BES, SSO, FIO, and SSA.

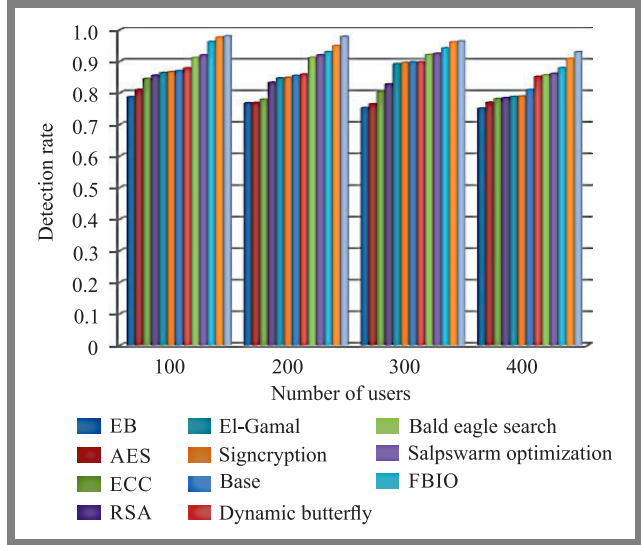


Fig. 4. Detection rate of FOI-SSA vs. other approaches, for a key size of 128.

4.2. Memory Usage Analysis

Figures 5–6 show an analysis of memory usage for FOI-SSA and other algorithms, for 128 and 64 key sizes. 8.5 MB of memory are used FOI-SSA for 100 users and a 64 key size, while other algorithms achieved higher utilization rates. Memory usage grows along with an increase in user count. These data help choose the optimal key and improve the blowfish concept.

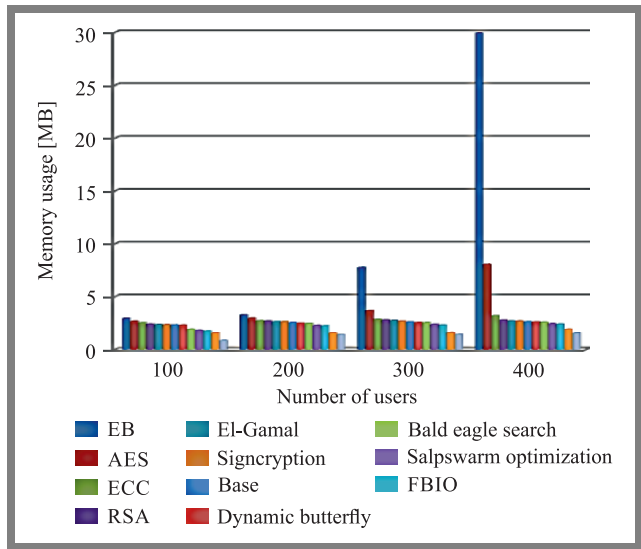


Fig. 5. Memory usage analysis of FOI-SSA vs. other approaches, for a key size of 64.

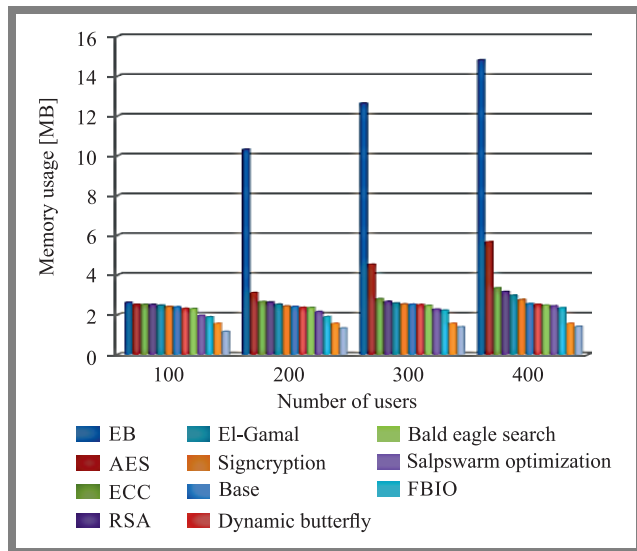


Fig. 6. Memory usage analysis of FOI-SSA vs. other approaches, for a key size of 128.

4.3. Time Analysis

Tables 3 and 4 show computational times for 128 and 64-bit keys. The analysis was performed for various user counts. For all key sizes, the time increases along with the user count. For 100 users, the computational time is shorter, but when the user count reaches 400, the time is longer for all other methods. However, FOI-SSA achieved a shorter time interval than its competitors. These advances are the result of using the blowfish concept and creating the optimal key.

4.4. Encryption and Decryption Time Analysis

The encryption time for various key sizes is summarized in Tables 5–6, while Tables 7–8 show the decoding time. For FOI-SSA, the encryption time is shorter for each key size. For user counts of 100 and 200, the decryption time is shorter with a 64-bit key which also requires less computational time for encryption. Thus, FOI-SSA achieves the shortest decryption and encryption times compared with its competitors, as shown in the tables.

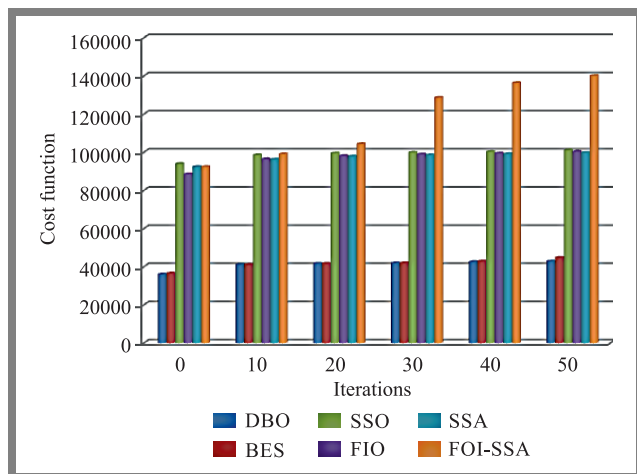


Fig. 7. Convergence analysis: FOI-SSA vs. other schemes.

4.5. Convergence Analysis

Convergence analysis of the proposed FOI-SSA system and the comparison of its performance with former models is shown in Fig. 7. FOI-SSA offers enhanced outcomes – the key break time at the 50-th iteration is 140004, meaning that it is higher than in the case of DBO, BES, SSO, FIO, and SSA. DBO's poor results were disclosed by obtaining a reduced key break time. Therefore, the goal is achieved, as shown in Eq. (21).

4.6. Attack Analysis

Figures 8 and 9 show the outcomes of research on various attack types, including inside and password spoofing attacks, for various key sizes and for varied user counts. Figure 8 shows the average key breakage time for inside and password spoofing attacks. While vulnerable to insider and password spoofing attacks, the proposed FOI-SSA approach has revealed a higher key break time. This is achieved due to a better blowfish concept and optimal key creation in FOI-SSA.

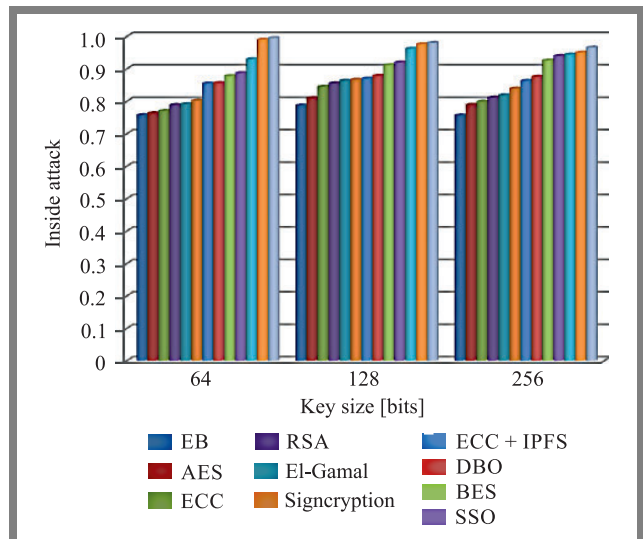


Fig. 8. Inside attack: FOI-SSA vs. other schemes.

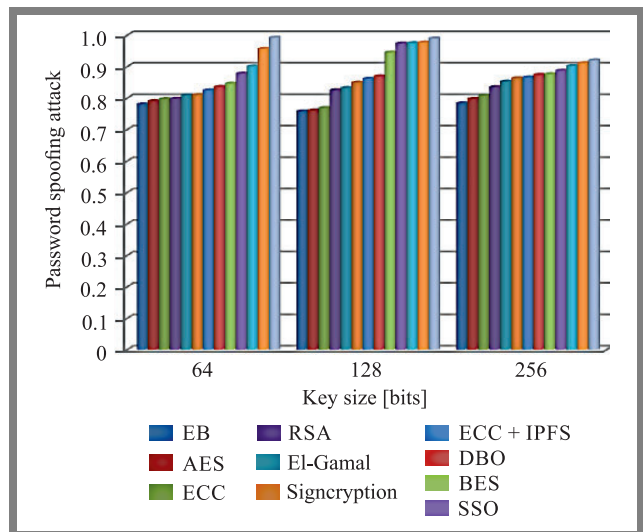


Fig. 9. Password spoofing attack: FOI-SSA vs. competitors.

Tab. 3. Time analysis using FOI-SSA over others algorithms for a key size of 64 [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	505	504	503	402	351	324	327	221	184	146	110	74	35
200	506	505	505	403	351	324	323	222	185	147	110	74	36
300	509	509	508	405	352	325	324	222	185	147	110	74	36
400	712	709	707	504	453	426	425	223	186	149	110	73	36

Tab. 4. Time analysis using FOI-SSA over others algorithms for a key size of 128 [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	613	612	612	407	355	327	326	217	181	145	109	73	35
200	643	642	641	472	414	377	375	223	187	149	110	73	35
300	702	701	701	487	420	393	391	224	187	149	112	73	37
400	790	789	787	598	546	519	518	226	188	150	113	74	37

Tab. 5. Encryption time for FOI-SSA vs. others algorithms for a key size of 64 [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	25	25	17	16	16	16	16	16	28	0.8	0.7	0.3	0.3
200	25	25	17	17	16	16	16	16	28	0.8	0.7	0.4	0.3
300	25	25	17	17	17	16	16	16	29	0.8	0.7	0.4	0.3
400	27	26	17	17	17	17	16	16	29	1	0.8	0.4	0.3

Tab. 6. Encryption time for FOI-SSA vs. others algorithms for a key size of 128 [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	25	25	16	16	16	17	16	16	3	0.8	0.6	0.4	0.3
200	25	25	17	17	17	17	16	16	3	0.9	0.8	0.4	0.3
300	35	26	17	17	17	17	16	16	3	0.9	0.8	0.5	0.5
400	55	27	18	17	17	17	17	16	4	1	1	0.8	0.7

Tab. 7. Decryption time for FOI-SSA vs. others algorithms for a key size of 64 [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	76	75	48	17	16	16	16	16	16	0.6	0.194	0.2	0.1
200	76	75	48	17	17	16	16	16	16	0.7	0.241	0.2	0.1
300	77	76	49	17	17	16	16	16	16	0.7	0.241	0.2	0.1
400	177	177	50	17	17	17	16	16	16	0.7	0.265	0.2	0.1

Tab. 8. Decryption time for FOI-SSA vs. others algorithms for a key size of [s].

User count	EB	AES	ECC	RSA	El-Gamal	Signcryption	ECC + IPFS	DBO	BES	SSO	FIO	SSA	FOI-SSA
100	124	75	49	16	17	16	16	16	15	0.6	0.2	0.2	0.1
200	179	77	49	17	16	16	16	16	16	0.7	0.3	0.2	0.1
300	259	117	49	17	17	16	16	16	16	0.7	0.4	0.3	0.1
400	275	142	68	17	17	17	17	16	16	1	0.5	0.3	0.2

5. Conclusion

This paper proposes a novel forensic method relying on a blockchain network. In order to protect the system against illegitimate users, an improved blowfish method is used. The system is made up of five distinct elements: hypervisor, node controller, log logger, IPFS file storage, and smart contract.

The suggested method entails six phases, including determination of the log file concept, key arrangement and exchange process, setup and control of evidence, assurance of integrity, agreement validation, release of the confidential file, and the blockchain-based communication phase. The proposed FOI-SSA approach offers an enhanced detection rate compared with EB, AES, ECC, RSA, El-Gamal, Signcryption,

ECC+IPFS, DBO, BES, SSO, FIO, and SSA algorithms. A higher detection rate was observed for FOI-SSA for a user count of 100 and a key size of 64. As the user count increases, the detection rate of FOI-SSA decreases for a key size of 64.

References

- [1] S. Rane and A. Dixit, "BlockSLaaS: Blockchain assisted secure logging-as-a-service for cloud forensics", *Proceedings of International Conference on Security & Privacy*, pp. 77–88, 2019 (https://doi.org/10.1007/978-981-13-7561-3_6).
- [2] P. Jain, "Decentralize log file storage and integrity preservation using blockchain", *International Journal of Computer Science and Information Technologies*, vol. 11, no. 2, pp. 21–30, 2020 (<https://ijcsit.com/docs/Volume%2011/vol11issue02/ijcsit2020110202.pdf>).
- [3] M. Pourvhab and G. Ekbatanifard, "Digital forensics architecture for evidence collection and provenance preservation in IaaS cloud environment using SDN and blockchain technology", *IEEE Access*, vol. 7, pp. 153349–153364, 2019 (<https://doi.org/10.1109/ACCESS.2019.2946978>).
- [4] N. Dalezios, S. Shiaeles, N. Kolokotronis, and B. Ghita, "Digital forensics cloud log unification: Implementing CADF in Apache CloudStack", *Journal of Information Security and Applications*, vol. 54, pp. 102555, 2020 (<https://doi.org/10.1016/j.jisa.2020.102555>).
- [5] C. Stelly and V. Roussev, "SCARF: A container-based approach to cloud-scale digital forensic processing", *Digital Investigation*, vol. 22, pp. S39–S47, 2017 (<https://doi.org/10.1016/j.diin.2017.06.008>).
- [6] J.H. Park, J.Y. Park, and E.N. Huh, "Blockchain based data logging and integrity management system for cloud forensics", *Computer Science & Information Technology*, vol. 149, 2017 (<https://doi.org/10.5121/csit.2017.71112>).
- [7] T.K. Dasaklis, F. Casino, and C. Patsakis, "SoK: Blockchain solutions for forensics", in *Technology Development for Security Practitioners*, B. Akhgar, D. Kavallieros, and E. Sdongos, Eds. Springer Cham, 2021, pp. 21–40 (https://doi.org/10.1007/978-3-030-69460-9_2).
- [8] M. Irfan, H. Abbas H, Y. Sun, A. Sajid, and M. Pasha, "A framework for cloud forensics evidence collection and analysis using security information and event management", *Security and Communication Networks*, vol. 9, no. 16, pp. 3790–3807, 2016 (<https://doi.org/10.1002/sec.1538>).
- [9] A.C. Kumar and R. Vimala, "Load balancing in cloud environment exploiting hybridization of chicken swarm and enhanced raven roosting optimization algorithm", *Multimedia Research*, vol. 3, no. 1, pp. 45–55, 2020 (<https://doi.org/10.46253/j.mr.v3i1.a5>).
- [10] R. Battistoni, R. Di Pietro, and F. Lombardi, "CURE-Towards enforcing a reliable timeline for cloud forensics: Model, architecture, and experiments", *Computer Communications*, vol. 91, pp. 29–43, 2016 (<https://doi.org/10.46253/j.mr.v3i1.a5>).
- [11] H. Mahdi, et al., "Vehicular Networks Performance Evaluation Based on Downlink Scheduling Algorithms for High-Speed Long Term Evolution-Vehicle", *International Journal of Interactive Mobile Technologies*, vol. 15, no. 21, 2021 (<https://doi.org/10.3991/ijim.v15i21.22475>).
- [12] Chao Lina, et al., "BSEn: A blockchain-based secure mutual authentication with fine-grained access control system for Industry 4.0", *Journal of Network and Computer Applications*, vol. 116, pp. 42–52, 2018 (<https://doi.org/10.1016/j.jnca.2018.05.005>).
- [13] D.F. Maesa, P. Mori, and L. Ricci, "A blockchain based approach for the definition of auditable access control systems", *Computers and Security*, vol. 84, pp. 93–119, 2019 (<https://doi.org/10.1016/j.cose.2019.03.016>).
- [14] P.T. Duy, H. Do Hoang, N.B. Khanh, and V.H. Pham, "Sdnlog-foren: Ensuring the integrity and tamper resistance of log files for SDN forensics using blockchain", *Proceedings of 2019 6th NAFOSTED Conference on Information and Computer Science (NICS)*, pp. 416–421, 2019 (<https://doi.org/10.1109/NICS48868.2019.9023852>).
- [15] G.G. Dagher, J. Mohler, M. Milojkovic, and P.B. Marell, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology", *Sustainable Cities and Society*, vol. 39, pp. 283–297, 2018 (<https://doi.org/10.1016/j.scs.2018.02.014>).
- [16] Hao Gu, L. Wanxin, Mark Nejad, and Chien-Chung Shen, "Access Control for Electronic Health Records with Hybrid Blockchain-Edge Architecture", *2019 IEEE International Conference on Blockchain*, 2019 (<https://doi.org/10.1109/Blockchain.2019.00015>).
- [17] B. Prasanalakshmi, A. Kannammal, and R. Sridevi, "Multimodal biometric cryptosystem involving face, fingerprint and palm vein", *International Journal of Computer Science Issues (IJCSI)*, vol. 8, no. 4, pp. 604, 2011 (URL: <http://ijcsi.org/papers/IJCSI-8-4-1-604-610.pdf>).
- [18] S. Kokilavani, M.D. Vignesh, E. Surya, A. Narendran, and B. Prasanalakshmi, "Enhanced biometric smart key scheme for smart card authentication", *2014 International Conference on Communication and Signal Processing*, pp. 1589–1592. IEEE, 2014 (<https://doi.org/10.1109/ICCSP.2014.6950116>).
- [19] V. Rupapara, F. Rustam, H.F. Shahzadehmood, A. Ashraf, and G.S. Choi, "Impact of SMOTE on imbalanced text features for toxic comments classification using RVVC model", *IEEE Access*, vol. 9, pp. 78621–78634 (<https://doi.org/10.1109/ACCESS.2021.3083638>).
- [20] M. Irfan, H. Abbas, and W. Iqbal, "Feasibility analysis for incorporating/deploying SIEM for forensics evidence collection in cloud environment", *Proceedings of 2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS)*, pp. 15–21, 2015 (<https://doi.org/10.1109/ICIS.2015.7166563>).
- [21] K. Kent, S. Chevalier, T. Grance, and H. Dang, "Guide to integrating forensic techniques into incident response", *NIST Special Publication*, vol. 10, no. 14, pp. 800–806, 2006 (<https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-86.pdf>).
- [22] S.K. Manoj and D.L. Bhaskari, "Cloud forensics – a framework for investigating cyber attacks in cloud environment", *Procedia Computer Science*, vol. 85, pp. 149–154, 2016 (<https://doi.org/10.1016/j.procs.2016.05.202>).
- [23] M. Ma, G. Shi, and F. Li, "Privacy-Oriented Blockchain-based Distributed Key Management Architecture for Hierarchical Access Control in the IoT Scenario", *IEEE Access*, vol. 7, pp. 34045–34059, 2019 (<https://doi.org/10.1109/ACCESS.2019.2904042>).
- [24] L. Pasquale, S. Hanvey, M. McGloin, and B. Nuseibeh, "Adaptive evidence collection in the cloud using attack scenarios", *Computers and Security*, vol. 59, pp. 236–54, 2016 (<https://doi.org/10.1016/j.cose.2016.03.001>).
- [25] V.K. Netajihole and G.P. Bhole, "Optimal Container Resource Allocation Using Hybrid SA-MFO Algorithm in Cloud Architecture", *Multimedia Research*, vol. 3, no. 1, pp. 11–20, 2020 (<https://doi.org/10.46253/j.mr.v3i1.a2>).
- [26] A.C. Kumar and R. Vimala, "Load Balancing in Cloud Environment Exploiting Hybridization of Chicken Swarm and Enhanced Raven Roosting Optimization Algorithm", *Multimedia Research*, vol. 3, no. 1, pp. 45–55, 2020 (<https://doi.org/10.46253/j.mr.v3i1.a5>).
- [27] M.K. Mahesh, "Workflow scheduling using Improved Moth Swarm Optimization Algorithm in Cloud Computing", *Multimedia Research*, vol. 3, no. 3, 2020 (<https://doi.org/10.46253/j.mr.v3i3.a5>).
- [28] M. Pourvhab and G. Ekbatanifard, "An efficient forensics architecture in software-defined networking-IoT using blockchain technology", *IEEE Access*, vol. 7, pp. 99573–99588, 2019 (DOI: 10.1109/ACCESS.2019.2930345).
- [29] P. Santra, P. Roy, D. Hazra, and P. Mahata, "Fuzzy data mining-based framework for forensic analysis and evidence generation in cloud environment", *Ambient Communications and Computer Systems*, pp. 119–129, 2018 (https://doi.org/10.1007/978-981-10-7386-1_10).

- [30] S. Aishwarya, R.S. Selvi, and R. Ilakkiya, "Secure public Auditing With Code-Regeneration In Multi Storage Cloud Architecture", 2006 (<https://ijartet.com/1370/v3s15ranganathancseit/conference>).
- [31] S. Wang, Y. Zhang, and Y. Zhang, "A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems", *IEEE Access*, vol. 6, pp. 38437–38450, 2018 (<https://doi.org/10.1109/ACCESS.2018.2851611>).
- [32] X. Qi, E.B. Sifah, K.O. Asamoah, J. Gao, X. Du, and M. Guizani, "MeDShare: Trust-less medical data sharing among cloud service providers via blockchain", *IEEE Access*, vol. 5, pp. 14757–14767, 2017 (<https://doi.org/10.1109/ACCESS.2017.2730843>).
- [33] S. Zawoad, A.K. Dutta, and R. Hasan, "Towards building forensics enabled cloud through secure logging-as-a-service", *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 2, pp. 148–62, 2015 (<https://doi.org/10.1109/TDSC.2015.2482484>).
- [34] P.P. Srivastava, S. Goyal, and A. Kumar, "Analysis of various NoSql database", *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, pp. 539–544, 2015 (<https://doi.org/10.1109/ICGCIoT.2015.7380523>).
- [35] M. Agrawal and P. Mishra, "A Modified Approach for Symmetric Key Cryptography Based on Blowfish Algorithm", *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 1, no. 6, 2012 (<https://www.ijeat.org/wp-content/uploads/papers/v1i6/F0610071612.pdf>).
- [36] R.K. Meyers and A.H. Desoky, "An Implementation of the Blowfish Cryptosystem", *2008 IEEE International Symposium on Signal Processing and Information Technology*, 2008 (<https://doi.org/10.1109/ISSPIT.2008.4775664>).
- [37] J-S. Chou and N-M. Nguyen, "FBI inspired meta-optimization", *Applied Soft Computing*, vol. 93, 2020 (<https://doi.org/10.1016/j.asoc.2020.106339>).
- [38] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm", *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020 (<https://doi.org/10.1080/21642583.2019.1708830>).
- [39] M.M. Beno, I.R. Valarmathi, S.M. Swamy, and B.R. Rajakumar, "Threshold prediction for segmenting tumour from brain MRI scans", *International Journal of Imaging Systems and Technolog*, vol. 24, no. 2, pp. 129–137, 2014 (<https://doi.org/10.1002/ima.22087>).
- [40] T. Renjith and M.J.S. Rangachar, "Hybrid Optimization based DBN for Face Recognition using Low-Resolution Images", *Multimedia Research*, vol. 1, no. 1, pp. 33–43, 2018 (<https://doi.org/10.46253/j.mr.v1i1.a5>).
- [41] J. Devagnanam, N.M. Elango, "Optimal Resource Allocation of Cluster using Hybrid Grey Wolf and Cuckoo Search Algorithm in Cloud Computing", *Journal of Networking and Communication Systems*, vol. 3, no. 1, pp. 31–40, 2020 (<https://doi.org/10.46253/jnacs.v3i1.a4>).
- [42] S.K.M. Shareef and R.S. Rao, "A Hybrid Learning Algorithm for Optimal Reactive Power Dispatch under Unbalanced Conditions", *Journal of Computational Mechanics Power System and Control*, vol. 1, no. 1, pp. 26–33, 2018 (<https://doi.org/10.46253/j.mr.v1i1.a5>).
- [43] P. Purnaye and V. Kulkarni, "BiSHM: Evidence detection and preservation model for cloud forensics", *Open Computer Science*, vol. 12, no. 1, 2022, pp. 154–170 (<https://doi.org/10.1515/comp-2022-0241>).

Sahadev Maruti Shinde

E-mail: sahadevmaruti@gmail.com

GITAM School of Engineering, Visakhapatnam, Andhra Pradesh, India

Venkateswara Rao Gurralla

Professor in Department of Computer Science and Engineering

GITAM School of Engineering, Visakhapatnam, Andhra Pradesh, India

Remedies to Thermal Radiation in Fused Silica Optical Fibers

Krzysztof Borzycki, Marek Jaworski, and Tomasz Kossek

National Institute of Telecommunications, Warsaw, Poland

<https://doi.org/10.26636/jtit.2023.166222>

Abstract — During fire incidents, optical fibers located within a fire-resistant cable are usually exposed to temperatures of 800°C to 1000°C. Hot fibers generate narrowband thermal (incandescent) radiation and collect broadband thermal radiation originating from the heated surroundings. The power of the second component, initially negligible, increases with time due to the rising number of fiber cracks and other defects acting as couplers for external radiation. Thermal radiation may interfere with fiber attenuation measurements performed during a fire test, but is rather unlikely to prevent data transmission with typical GbE and 10GbE transceivers during a fire. This problem may be remedied by combining the following methods: using single mode fibers instead of multimode fibers, using bandpass filters to block thermal radiation, and selecting proper transmitter power, wavelength and photodetector.

Keywords — *attenuation testing, fire resistant fiber optic cable, fire test, fused silica optical fiber, incandescent emission, interference filtering*

1. Introduction

Fire-resistant fiber optic cables temporarily retain optical continuity during a fire, at temperatures of up to 1000°C and over periods of 15–120 minutes [1]–[3], providing connectivity for fire safety and video surveillance systems, as well as for emergency communications. Such cables comprise standard telecom fibers made of fused silica.

Hot optical fibers either glow themselves or are exposed to radiation from their surroundings – usually from carbonized remains of polymer tubes and coatings. We have investigated both effects experimentally in [4], presenting a short review of fire test standards and conditions.

In this paper, we present the sources, spectral power distribution (SPD) and the power of thermal radiation (Section 2), based on results of the experiments conducted, specific methods relied upon for eliminating such phenomena (Section 3), the potential interference with Ethernet data links (Section 4), as well as methods for monitoring fiber attenuation in a cable in the course of a fire test (Section 5). Then, we present the results of an experiment in which bandpass filters have been used to reject thermal radiation (Section 6) and specify the activities recommended in connection with the fiber itself and with the wavelength in order to minimize such interference (Section 7). Section 8 contains a summary.

2. Thermal Radiation in Fibers Under Fire

Detailed characteristics of radiation generated by silica fibers at temperatures of up to 1000°C (such as power, spectral power distribution, variation with time, and physical mechanisms) are presented in paper [4]. Therefore, we only provide the summaries and representative examples thereof.

2.1. Incandescence of Fiber

Figures 1 and 2 show the measured SPDs of radiation emitted by a 50/125 μm multimode fiber OFS OM2 [7] and a single mode fiber conforming to the ITU-T G.652.A Recommendation [8], respectively. The 1 m samples were heated to 900°C before conducting the measurements. Due to the fact that the measurement resolution was limited (5 nm or 10 nm), the three strongest emission peaks at 1383 nm, 1393 nm and 1407 nm [9] are visible as a single peak. SPDs for all multimode and single-mode telecom fibers we have tested were similar, except for varying power levels. The power of radiation exiting each fiber was measured with an InGaAs sensor in the 800–1700 nm spectral range.

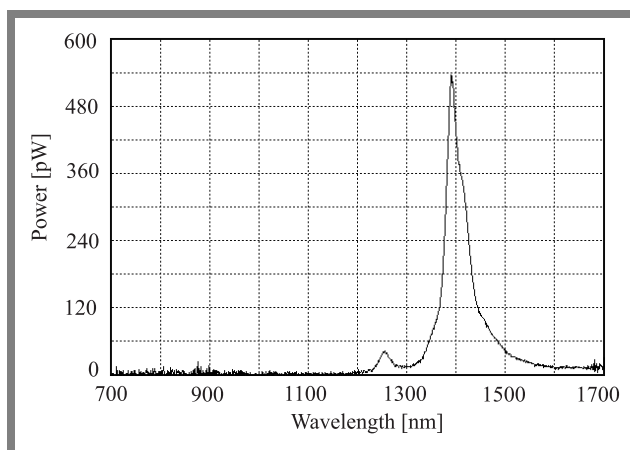


Fig. 1. Spectral power distribution of thermal emissions from an OM2 multimode fiber at 900°C. The total power is –49.39 dBm.

Incandescence of a single mode fiber, even exhibiting water peaks (ITU-T G.652.A), is substantially lower than of a 50/125 μm OM2/OM3/OM4 multimode fiber, due to differences in core diameter (8.3 vs. 50 μm) and numerical aperture (0.14 vs. 0.20). For the same OH^- content and temperature, power that is approx. 72 times lower (18.6 dB) may be ex-

pected, but real values measured in our experiments varied considerably.

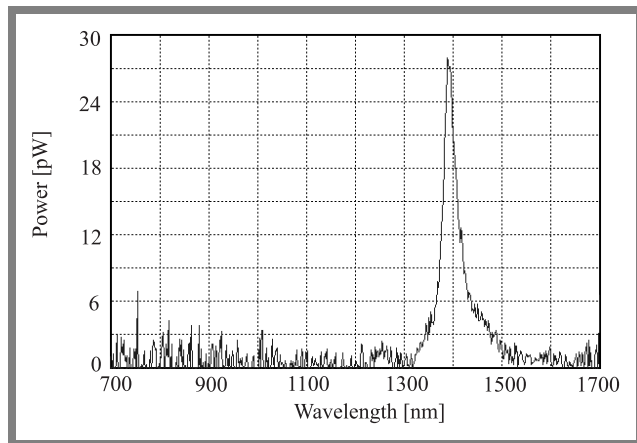


Fig. 2. Spectral power distribution of thermal emission from a G.652.A single-mode fiber at 900°C. Total power was -66.71 dBm.

Fiber incandescence decreases with time due to the ‘drying’ of the fiber’s core and diffusion of hydrogen through the cladding [5]. A 50% intensity reduction was observed in our experiments after 1–2 hours at 1000°C [4], which corresponds to a typical duration of a fire incident or a fire test. The fading phenomenon is slower when fiber cladding is doped with fluorine.

2.2. Thermal Radiation Coupled into the Fiber

The coupling of external radiation into the fiber is caused by radiation being reflected through cracks penetrating the fiber’s cladding and core, created when microscopic inclusions of cristobalite (crystalline form of silica) appear on the fiber’s surface [4], [6], [10]. Cracks in heat-degraded fibers have complex shapes and various depths.

A crack reflects a small portion, i.e. approx. 4% of incoming radiation at an angle dependent on the crack’s direction. A small portion of external radiation is directed into the fiber’s acceptance angle, while a similar portion of radiation transmitted through the fiber is lost. Other defects in fiber core or at the core/cladding interface, like inclusions of cristobalite, tiny gas bubbles or defects introduced by gamma radiation [11], also facilitate unwanted coupling of external radiation.

This coupling phenomenon occurs more efficiently in multimode fibers due to their higher numerical aperture, i.e. 0.20 in a 50/125 μm OM2/OM3/OM4 graded index multimode fiber, versus 0.14 in a ITU-T G.652 single-mode fiber.

We have observed that the power of coupled radiation and the number of defects in the fiber section affected by high temperature grow as the heating duration increases. Light may be seen escaping from the fiber at locations with cracks or inclusions in the core. The strongest interference is experienced when the fiber is surrounded by black char, as its emissivity coefficient is close to 1. For temperature and wavelength ranges of interest, thermal radiation is characterized by continuous SPD and its intensity is rising along with wavelength.

An example of such an SPD we have measured is shown in Fig. 3.

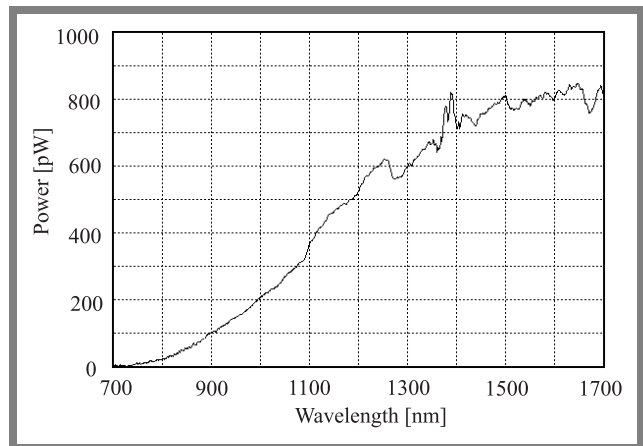


Fig. 3. Spectral power distribution of coupled thermal emissions of an OM2 multimode fiber at 1000°C. Localized light escapes were noticed due to cracks, but optical continuity of the fiber was retained. Total power equaled -39.79 dBm.

2.3. Changes of Thermal Radiation Spectrum over Time

When fused silica fiber is exposed to high temperatures (over 800°C) for a duration typical for a fire incident, defined in standards [1], [2] as 30–120 minutes, two concurrent changes take place over time:

- fading of thermal emissions due to the destruction of OH^- ions [5],
- progressive degradation of the fiber by cristobalite inclusions and cracking [4], [6].

We have observed a gradual decrease of (narrowband) fiber incandescence and an increase in (broadband) radiation coupled into the fiber.

Initially, the SPD is composed of narrow peaks (see Figs. 1 and 2), but a continuous long-wavelength component rises steadily with time and becomes dominant in a degraded fiber (Fig. 3). Lowered effectiveness of the bandpass filter in blocking related interference is one of the consequences of such a phenomenon.

A fast and total fiber failure is possible at temperatures above 950°C. The fiber loses its optical continuity, the power of coupled thermal radiation increases to approx. -33 dBm and -50 dBm for a 1 m section of an OM2 multimode and single-mode fiber at 1000°C, respectively. The SPD is similar to that illustrated in Fig. 3. In our experiments, almost 30% of multimode samples and only one sample of a single mode fiber failed under such conditions.

2.4. Power of Incandescent and Coupled Radiation

During the experiments, the power of incandescent radiation measured in the 800–1700 nm band was relatively low, i.e. amounted to -60 to -49 dBm over a 1 m section an OM2 or OM3 multimode fiber at 900°C and -85 to -66 dBm in a single-mode fiber under the same conditions, including old G.652.A fibers with high OH^- content.

The radiation exiting at the end of the fiber may exhibit a dip at “water peak” wavelengths of approx. 1390 nm, indicating high attenuation in hot fiber at this range i.e. 5–8 dB/m at 900–1000°C. Consequently, the accumulated power of incandescent radiation is self-limiting (saturating) in fiber samples longer than 0.5–1 m.

The power of coupled radiation measured during laboratory tests on 1 m fiber sections heated to 900°C and covered by carbon char equaled approx. –42 dBm for OM2 and OM3 multimode fibers, –56 dBm for a G.652.A single mode fiber, and –60 dBm for a G.657.A2 single mode fiber.

Thermal radiation coupled into the fiber is also self-limiting. Increasing density of cracks and inclusions in deteriorated fibers raises the coupled power (per unit of length). However, attenuation grows as well, due to the escape of guided radiation from the fiber. Measurements performed during a fire test in accordance with the DIN 4102-12 standard suggest that saturation length of an OM2 multimode fiber at 1000°C is 2–3 m (test duration: 90 minutes). The power of coupled radiation measured under such conditions was –38 to –35 dBm for a 32 m section of an OM2 fiber placed in the fire zone, meaning that power saturation conditions were met. For a 3 m effective saturation length, the power of coupled radiation accumulated in a long fiber sample is approx. 5 dB higher compared to 1 m sections tested in the laboratory.

We estimate the maximum expected power of thermal radiation in a fire according to the DIN 4102-12 [2] standard as approx. –37 dBm, –51 dBm and –55 dBm, respectively, for undamaged fibers indicated above, using an InGaAs photodiode (800–1700 nm) as sensor. For a 850 nm sensor based on a silicon photodiode, with sensitivity up to 1100 nm and a multimode fiber, this power is reduced to approx. –45 dBm.

3. Limitation of External Thermal Radiation

In this section, we present two potential methods for modifying the design of fire resistant cables in order to prevent coupling of undesirable thermal radiation into optical fibers.

3.1. Elimination of Black Soot

If the material surrounding the glass fiber is characterized by low emissivity and high opacity at the wavelength of interest, e.g. 1300 nm, coupled radiation will be eliminated. Titanium dioxide (TiO₂) with the grain size of 2 μm is an example of such material. The black carbon soot left after pyrolysis of fiber coatings, the loose tube and the gel must be fully burned. However, since typical fire resistant cables include a fire barrier made of mica tape and often also a steel armor [4] layer, no oxygen source is available to ensure such a full burn inside the cable. Although oxygen may be provided by incorporation of a solid oxidizer, a rapid burn process would produce undesirable heat and flue gases, such as CO₂, CO, and H₂O vapor, making this idea impractical.

3.2. Fibers with Photonic Barrier

Another concept is to use fibers with a photonic barrier surrounding the core, made of microscopic gas-filled holes – like hole-assisted fibers (HAF) [12], [13] or ‘needles’ existing in Corning NanoStructures fibers [14]. Such fibers were originally developed for bending-insensitive cables in FTTH networks. The photonic barrier prevents both radiation escape from bent fibers (this was the original purpose) and entry of external radiation. Unfortunately, fibers with a photonic barrier are inconvenient to splice and expensive. After 2010, such fibers were replaced by other designs. Their reintroduction in fire resistant cables would be valuable.

4. Interference in Operation of Digital Fiber Optic Links

Transmission of data over fibers in a cable affected by fire may subject to:

- increase in fiber attenuation due to physical deterioration and light escape,
- addition of unmodulated thermal radiation to the data carrying signal at the receiver.

Most low- and medium-speed data links [15] use non-return to zero (NRZ) modulation, with the signal transmitted either at full power (logical ‘1’) or low power (logical ‘0’). There is a minimum value of the EXT power ratio in ‘1’ and ‘0’ states (in dB) of the signal at the receiver input that is required for error-free transmission. Without optical interference, the extinction signal ratio at the receiver port is the same as at the transmitter port:

$$\text{EXT}_{\text{TX}} = 10 \log \left(\frac{P_1}{P_0} \right).$$

Additional optical interference power P_I is defined as a sum of incandescent and coupled radiation accumulated in the hot section of a fiber and attenuated in the section of between the fire zone and the receiver. The extinction ratio, as seen by the receiver, is then:

$$\text{EXT}_{\text{RX}} = 10 \log \left(\frac{P_1 + P_I}{P_0 + P_I} \right).$$

In the receiver, a fixed difference between “1” and “0” levels at the decision circuit is maintained by automatic gain control responding to the difference between power received in “0” and “1” states, designated as OMA (Optical Modulation Amplitude) [15], but noise level increases, as PI produces an equivalent of dark current in photodiode.

How much interference may an Ethernet optical link tolerate? The EXT values required in IEEE 802.3 [15] for 100 Mb/s, 1 Gb/s and 10 Gb/s transmitters vary between 3 and 9 dB – see Table 1. This corresponds to power emitted in the “0” state being equal to 50.0% and 12.6% of power emitted in the “1” state, respectively. Therefore, a rough limit for tolerable thermal radiation power may be set at 12.5% of signal power at “1” level, or 6 dB below receiver sensitivity.

Tab. 1. Parameters of Ethernet interfaces defined in IEEE 802.3 [15], calculated tolerated interference limit and estimated power of thermal radiation at 1000°C.

Interface designation by IEEE	Max. link length [m]	Fiber type	Nominal wavelength [nm]	Transmitter EXT min. [dB]	Receiver sensitivity [dBm]	Interference limit [dBm]	Thermal radiation max. [dBm]
100BASE-BX10*	10 000	SMF	1310/1550	6.6	−28.2 (−23.3)	−34.2	−51.0
1000BASE-SX	550	MMF	850	9.0	−17.0 (−13.0)	−23.0	−45.0
1000BASE-LX	550	MMF	1300	9.0	−19.0 (−14.4)	−25.0	−37.0
1000BASE-LX	5 000	SMF	1310	9.0	−19.0 (−14.4)	−25.0	−55.0
10GBASE-SR	400	MMF	850	3.0	−11.1 (−7.5)	−16.1	−45.0
10GBASE-LR	10 000	SMF	1310	3.5	−14.4 (−10.3)	−20.4	−51.0
10GBASE-ER	30 000	SMF	1550	3.0	−14.1 (−11.3)	−20.1	−55.0
10GBASE-LRM	220	MMF	1300	3.5	−6.5 (−6.0)	−12.5	−37.0
10GBASE-PR-D4*	40 000	SMF	1270/1578	6.0	−29.0 (−27.0)	−35.0	−55.0

Abbreviations: MMF – 50/125 μm multimode fiber (OM2 – OM5), SMF – single mode fiber.

* Duplex transceiver for access networks incorporating a WDM bandpass filter.

Such parameters as thermal radiation power, receiver sensitivity and interference limit for Ethernet transceivers, are presented in Table 1. IEEE 802.3 includes different requirements for receiver sensitivity, measured when the input signal has the lowest extinction ratio permitted for the transmitter (stressed receiver sensitivity) – as shown in parentheses. However, the limit of interfering thermal radiation presented in Tab. 1 is calculated for the receiver sensitivity measured under optimal conditions, including high EXT.

The use of a G.652.A fiber is assumed for single mode links that are up to 10 km long, as such a transmission medium is still permitted in the ISO/IEC 11801 standard for structural cabling systems [16].

Data from Tab. 1 prove that thermal radiation will not disturb a 100 Mb/s, 1 Gb/s or 10 Gb/s Ethernet data link, but compliance must be verified in the case of very sensitive receivers or analog transmission equipment.

If the estimated power of optical interference may be an issue, the use of a high-power transmitter can keep the signal-to-noise ratio high enough for an error-free transmission. The receiver must be fitted with an attenuator to reduce input power to the upper level of the specified range, to ensure best tolerance to an increased fiber loss.

The other method is to use a bandpass filter, as described in Section 6. Duplex transceivers designed for access networks have built-in bandpass filters separating signals transmitted in opposite directions at different wavelengths, which also reduces interference originating from thermal radiation.

5. Interference During Fiber Attenuation Measurements

5.1. Comparison of LSPM and OTDR Methods

Changes in fiber attenuation during cable fire testing are normally measured using a light source and a power meter

(LSPM) setup as shown in Fig. 4. There is the “Method B” from the EN 60793-1-40 standard [17] and it offers several advantages:

- wide dynamic range, up to 80 dB with commercially available instruments,
- high resolution: 0.01 dB or 0.001 dB with typical measurement equipment,
- no need for launch and tail fibers to test a short fiber,
- short measurement time of up to 1 ms.

The disadvantage of the LSPM method is its sensitivity to thermal radiation. It adds to light source radiation, meaning that indicated fiber attenuation (or its change vs. initial conditions) is lower than real one.

Measurements performed with an optical time domain reflectometer (OTDR), called the “Method C” in EN 60793-1-40, are a better idea for testing long fibers, connectors and splices, and for locating faults. Both methods are approved in the EN 60793-1-46 standard [18] for testing fibers and cables when the impact of external factors (mechanical force applied to the cable or fiber, temperature, aging, etc.) on fiber attenuation are to be established. In OTDR measurements, interference from thermal radiation increases the noise floor at the receiver and reduces dynamic range. OTDR does not indicate a hot zone as an upward shift of fiber trace, but can show the loss resulting from fiber deterioration or bending. Distance measurements are not affected.

5.2. Interference During Measurements with LSPM

The optical path between the light source and the optical power meter includes a relatively short (1–30 m) fiber enclosed in a cable subjected to fire (Fig. 4). Its initial loss IL_F may change during the test. The combined loss of all other components along the optical path, i.e. fibers outside the fire zone, splices, connectors and in multi-channel, multi-wavelength or filter-equipped setups, including switches, splitters or filters (IL_C), is significantly higher than IL_F and equals up to

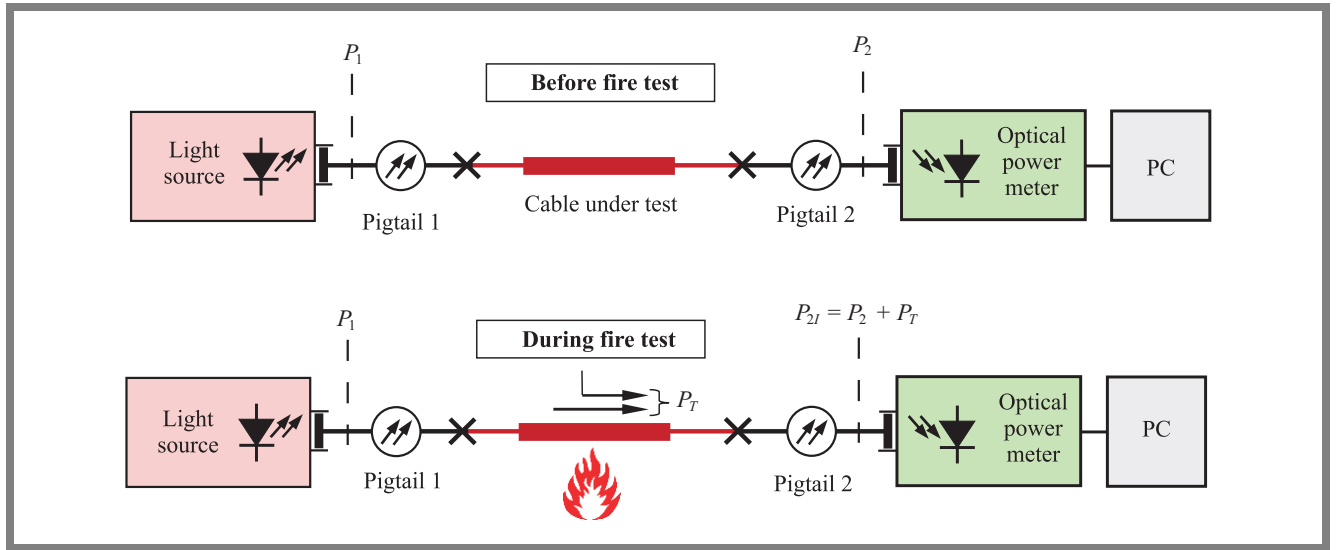


Fig. 4. LSPM setup for measuring and recording variations in attenuation (loss) of a single optical fiber in a cable during a fire test. The fiber tested is fusion spliced to pigtails to provide connections to test instruments.

20 dB in some cases. It is assumed that IL_C is constant during the test and that variations of the measured power are caused by changes in IL_F only. As all pieces of measurement equipment are affected by specific level of uncertainty, IEC standards for fiber cables [19] define an indicated loss change of up to 0.05 dB for a single-mode fiber and 0.20 dB for a multimode fiber, as a ‘no loss change’ condition.

Unfortunately, the standards relate only to testing fibers that are free of any optical interference. It is assumed that all radiation present in the fiber being tested originates from the light source which generates radiation characterized by a stable power (P_1) and SPD. Therefore, variations in output optical power (P_2) measured with an optical power meter can be directly converted to fiber attenuation (IL_F) changes.

Using logarithmic scale of power (in dBm) and loss (in dB), the following formulas for power reaching the input of the optical power meter (P_2) and the fiber loss variations measured are used:

$$P_2 = P_1 - (IL_F + IL_C),$$

$$\Delta IL_C = -\Delta P_2,$$

where ΔIL_C and ΔP_2 are variations of parameters from values before the start of the fire test.

As presented in Section 2 and in paper [4], a hot fiber both generates and collects radiation from its surroundings, which adds to radiation emitted by the light source, as shown in Fig. 4. With classical LSPM setup, the optical power meter (OPM) displays a sum of:

- real test power P_2 originating from the source and attenuated of variation to the power meter,
- power of interfering thermal radiation P_T .

Let us identify a value of P_2 affected by optical interference as P_{2I} :

$$P_{2I} = P_2 + P_T \quad \text{for power values in [W]}$$

or

$$P_{2I} = 10 \log \left[10^{\frac{P_2}{10}} + 10^{\frac{P_T}{10}} \right] \quad \text{for power values in [dBm]}.$$

The error in loss indication $Err IL$ (in the logarithmic scale) depends on P_T and P_2 ratio:

$$Err IL = 10 \log \left(\frac{P_2 + P_T}{P_2} \right) = 10 \log \left(\frac{P_1}{P_2} + 1 \right) \quad \text{in linear scale}$$

or:

$$Err IL = 10 \log \left(10^{\frac{P_T - P_2}{10}} + 1 \right) \quad \text{in log scale}.$$

P_T measured by OPM when the light source is off is not accurate, because:

- sensitivity of the photodetector in the power meter is wavelength-dependent,
- the power meter is calibrated for the central wavelength of the light source used, and not for thermal radiation,
- a portion of thermal emissions lies outside of photodetector’s sensitivity range.

However, this effective value of interference power is useful for error calculations.

The power and loss evaluation error resulting from P_T is shown in Table 2.

Tab. 2. Fiber loss indication error ($Err IL$) as a test signal P_2 to interference P_T ratio.

$Err IL$ [dB]	P_T / P_2	P_2 / P_T	$P_2 - P_T$ [dB]
1.00	0.2589	3.86	5.87
0.50	0.1220	8.20	9.14
0.20	0.0471	21.23	13.27
0.10	0.0233	42.92	16.33
0.05	0.0116	86.21	19.36

To be compliant with ‘no loss change’ limits, the thermal radiation error shall be below 0.05 dB for single-mode fiber cables or 0.20 dB for multimode fibers. This condition is

ensured when interference power P_T at the input to the power meter is 13.5 dB or 19.5 dB lower, respectively, than the power indicated with an active light source. This condition shall be met at the maximum increase of loss of the cable under test.

During fire testing, intense thermal radiation can mask even a large increase in fiber attenuation or a fiber failure, producing erroneous decrease in fiber loss. Such a phenomenon is explained in Fig. 5, showing the measurement data from a DIN 4102-12 [2] fire test [4].

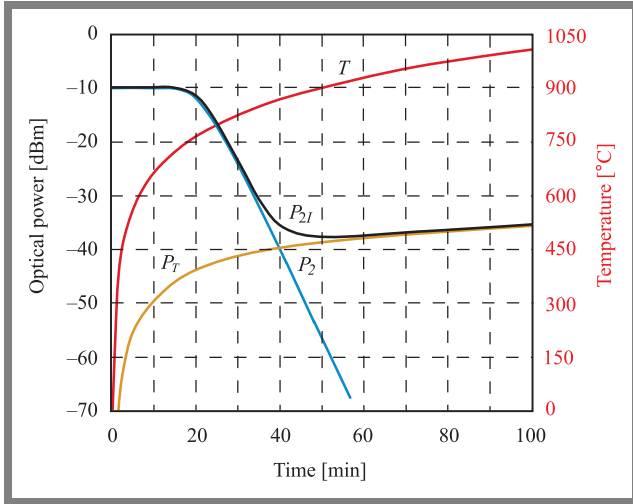


Fig. 5. Temperature T , power P_2 , P_T and P_{2I} versus time in a 33 m section of an OM2 multimode fiber fire resistant cable during a DIN 4102-12 fire test. Interference caused by thermal radiation may result in a false test pass status, remaining unnoticed.

A similar problem exists when the LSPM test setup measures loss variations at several wavelengths, e.g. 1310 nm, 1550 nm, and 1625 nm for single mode fibers, i.e. with multiple light sources and a WDM demultiplexer with limited selectivity (typically 25 dB). To minimize errors in multi-wavelength measurements, the power of all light sources shall be adjusted, e.g. by using attenuators, to obtain similar readouts on the power meter with each source used solo. This compensates for the loss differences in WDM multiplexers or fibers at all wavelengths.

5.3. Limitations Imposed by Optical Interference

The simplest way of overcoming interference from thermal radiation is to employ a light source with minimum output power, dependent on:

- maximum measurable loss (IL_{MAX}),
- maximum expected power of thermal radiation (P_T).

Table 3 shows a data for fire test performed in accordance with DIN 4102-12 [2], and with the peak temperature of 1000°C. The setup allows to measure the increase in attenuation up to 120% of the limit defined in EN 50582 [20] i.e. 1 dB/m for a single mode fiber at 1550 nm and 2 dB/m for a multimode fiber at 1300 nm, with Err_{IL} limited to 0.05 dB and 0.20 dB, respectively. In both cases, the power meter incorporates an InGaAs photodetector, with a sensitivity window of 800–1700 nm.

The DIN 4102-12 standard requires the cable placed in the fire chamber to be at least 3 m long, hence the accumulated power of (predominantly coupled) thermal radiation is close to saturation and reaches values presented in Subsection 2.4.

For the sake of simplicity, a uniform distribution of attenuation increase along the fire-affected zone is assumed.

Typical light sources with an FP laser generate power between -10 dBm and $+3$ dBm. For such a source, the maximum length of fiber in the fire chamber is limited to 15–25 m for single mode fiber and 7–10 m for multimode fibers. Because of these limits, concatenation of multiple fibers in the cable under test shall be avoided. Fortunately, the increase of attenuation allowed for fire tests [19] is high enough to ensure good measurement accuracy with samples of just 0.8–1 m.

5.4. Bandpass Filtering

Signal filtering improves the P_2/P_T ratio by 15–25 dB, and also (alternatively):

- reduces measurement error Err_{IL} , as shown in Tab. 2,
- increases maximum loss which can be measured with the same Err_{IL} by 15–25 dB,
- allows to use a less powerful light source or to increase the connection loss (IL_C) by 15–25 dB.

However, the improvement from filtering can be insufficient, e.g. while testing multimode fibers longer than 20 m (Tab. 3).

5.5. Correction of Thermal Radiation Background

Another method used for eliminating interference caused by thermal radiation, as researched by the authors, involves the following:

- periodic disconnection by switching fibers or deactivating of light source,
- two power measurements performed in each cycle: one with the source on (P_{2I}) and one with the source off (P_T),
- subtraction of P_T from P_{2I} to obtain a true power of the test signal (P_2):

$$P_2 = P_{2I} - P_T \quad \text{power values in [W].}$$

Such a method may be automated. Assuming the power of the light source in its off state is zero, the improvement in signal-to-interference ratio is limited by the resolution of power meter used (typically 0.01 or 0.001 dB) and by the combined internal noise of the light source and power meter for the time interval between measurements of P_T and P_{2I} .

Assuming that combined instrument resolution and noise equals 0.002 dB, the interference caused by thermal radiation is reduced by 33.36 dB. Unfortunately, the noise level in the corrected signal (with P_2 at the lowest level) will be approx. 3 dB. The real improvement corresponding to uncertainty of 0.05–0.20 dB, limited by LSPM resolution and noise, is 20–25 dB. Such a value is comparable to the improvement achieved by including an optical bandpass filter in front of the power meter. Combined use of both techniques may reduce interference by approx. 40 dB.

Tab. 3. Power of thermal radiation and required output power of the light source for several types of fibers and their lengths in a cable subjected to fire. Connection loss IL_C is 1 dB in all cases. Signal-to-interference ratio is: 19.5 dB for single-mode fibers and 13.5 dB for multimode fibers.

Fiber type	Fiber length [m]	1.2 × max. fiber loss [dB]	Max. total loss [dB]	P_T [dBm]	P_2 [dBm]	P_1 [dBm]
Single mode G.652.D, G.657.A1, G.657.A2	4	4.8	5.8	−55.5	−36.0	−30.2
	10	12.0	13.0	−55.0	−35.5	−22.5
	20	24.0	25.0	−55.0	−35.5	−10.5
	30	36.0	37.0	−55.0	−35.5	+1.5
Single mode G.652.A	4	4.8	5.8	−51.5	−32.0	−26.2
	10	12.0	13.0	−51.0	−31.5	−18.5
	20	24.0	25.0	−51.0	−31.5	−6.5
	30	36.0	37.0	−51.0	−31.5	+5.5
Multimode 50/125 μm OM2, OM3, OM4, OM5	4	9.6	10.6	−37.5	−24.5	−13.9
	8	19.2	20.2	−37.0	−24.0	−3.6
	10	24.0	25.0	−37.0	−24.0	+1.0
	12	28.8	29.8	−37.0	−24.0	+5.8

6. Elimination of Thermal Radiation with Bandpass Filters

In theory, passive bandpass filtering performed at the receiver's input in the case of a data link or optical power meter in the case of an attenuation test system eliminates unwanted radiation having wavelengths outside of the transmission/test band. However, effectiveness of such an approach is limited by the following factors:

- Insertion of filter adds 2–4 dB to signal loss;
- Wavelength tolerance of transmitters and light sources (typically ± 20 nm) and filters (typ. ± 10 nm), plus the spectral width of radiation emitted by FP lasers (typically 5–10 nm) requires a 30–50 nm bandpass, unless the components are individually matched;
- Fiber test wavelengths are defined in applicable standards, most of them require testing of multimode cables loss at 1300 nm, i.e. near emission peaks (Tab. 1), and require a wideband InGaAs detector, while all applications of OM2/3/4/5 fibers are at wavelengths from 820 to 940 nm.
- A low-cost bandpass filter often has too narrow a stop band to effectively filter out thermal radiation across the entire sensitivity range of InGaAs photodetectors. Designers of data links or test systems must carefully review filter specifications. When a silicon photodetector is used in 850 nm multimode system, the bandpass filter needs to block wavelengths of up to 1100–1150 nm only.

6.1. Experiment

We have measured attenuation of unwanted thermal radiation with commercially available bandpass filters for both blackbody radiation coupled into the fiber (Tab. 4) and thermal emission from a multimode fiber (Tab. 5). Filters were installed in a Thorlabs FOMF/M fiber-to-fiber coupler with

FC interfaces. 50/125 μm multimode (OM2) type fibers were placed on both sides. The optical power meter (HP 81532A) was equipped with an InGaAs photodetector. The improvement in signal-to-interference (S/I) ratio was calculated as a reduction of power observed after insertion of the filter into the optical path, corrected by the insertion loss of the same coupler with a 1304 nm filter, measured separately. This parameter reflects an improvement in the S/I ratio of the data link or in attenuation of test setup operating at 1300 nm or 1310 nm wavelengths.

As expected, bandpass filtering was more effective against thermal emissions (illustrated as a few narrow peaks in Fig. 1) than against coupled wideband blackbody radiation (Fig. 3), as a portion of blackbody radiation always falls within the filter's passband.

SPD shown in Fig. 3 suggests that bandpass filtering will be less effective at 1550 nm due to higher spectral density of the coupled thermal radiation encountered, with an estimated S/I gain being approx. 2.5 dB lower than at 1300/1310 nm. The evolution of thermal radiation spectrum with the increasing duration of heat exposure slowly reduces S/I improvement provided by filtering, because of the growing wideband coupled radiation.

7. Selection of Fiber and Wavelength to Minimize Interference from Thermal Radiation

The recommendations below summarize how to mitigate the negative impact of thermal radiation:

- single mode fibers generate low thermal emissions,
- wavelength bands with low spectral density of thermal emissions should be used,

Tab. 4. Measured improvement in S/I ratio achieved by insertion of a bandpass filter [dB]. The transmission wavelength was 1304 nm and the radiation source was carbon soot heated to 800°C and 1000°C. The emissions spectrum was as presented in Fig. 3.

Manufacturer and type of filter	Central wavelength [nm]	Passband width [nm]	Blocking band [nm]	S/I improvement at 800°C [dB]	S/I improvement at 1000°C [dB]
Thorlabs FB1300-30	1300	30	200 – 2600	16.65	15.99
Edmund Optics FB1300-25	1300	25	200 – 1800	18.12	17.45
Thorlabs FB1300-12	1300	12	200 – 2600	20.94	20.34

Tab. 5. Improvement in S/I ratio achieved by insertion of a bandpass filter [dB]. The transmission wavelength was 1304 nm and the radiation source was an OM2 fiber heated to 800°C and 900°C. Emission SPD was as shown in Figs. 1 and 2. The test could not be performed at 1000°C due to the failure of the fiber at 950°C.

Manufacturer and type of filter	Central wavelength [nm]	Passband width [nm]	Blocking band [nm]	S/I improvement at 800°C [dB]	S/I improvement at 900°C [dB]
Thorlabs FB1300-30	1300	30	200 – 2600	22.28	20.86
Edmund Optics FB1300-25	1300	25	200 – 1800	24.70	23.20
Thorlabs FB1300-12	1300	12	200 – 2600	25.34	25.76

– Si photodetectors are less sensitive to thermal emission,
– bandpass filters improve the signal-to-interference ratio.
Using a cable with an OH-free single mode fiber compliant with ITU-T G.652.D or G.657.A1/A2 [21] is the best approach. With 50/125 μm multimode fibers [21], the best option is to use the 850 nm band, where spectral density of thermal radiation of either type is low (Figs. 2, 3). The use of silicon photodetectors at the 850 nm single wavelength or 850–940 nm CWDM links operating over an OM5 fiber is a great idea due to insensitivity to radiation with wavelengths longer than 1100 nm, where the bulk of thermal radiation is located (Fig. 3).

8. Summary

The data obtained in the course of the experiments with telecom-grade fused silica fibers and fire tests of fiber optic cables indicate that optical fibers heated to temperatures above 850°C become a source of thermal radiation of non-negligible power. The resulting interference cannot be ignored in some applications, in particular during attenuation measurements.

Fortunately, there are several effective methods of preventing or mitigating this phenomenon, provided that the designer of fiber optic system is aware of the issue. Unfortunately, current standards for testing fiber optic cables do not include a warning or any advice how to reduce the negative impact of thermal radiation.

References

- [1] –, EN 50200, “Method of test for resistance to fire of unprotected small cables for use in emergency circuits”.
- [2] –, DIN 4102-12, “Fire behavior of building materials and elements – Part 12: Fire resistance of electric cable systems required to maintain circuit integrity – Requirements and testing”.
- [3] –, EN 50582: “Procedure to assess the circuit integrity of optical fibres in a cable under resistance to fire testing”.
- [4] K. Borzycki, M. Jaworski, and T. Kossek, “Some effects of high temperature in fused silica optical fibers”, *J. Telecommunications and Inform. Technol.*, no. 3, pp. 56–71, 2021 (DOI: 10.26636/jtit.2021.153521).
- [5] A.H. Rose and T.J. Bruno, “The observation of OH in annealed optical fiber”, *J. Non-Cryst. Solids*, vol. 231, no. 3, pp. 280–285, 1998 (DOI: 10.1016/S0022-3093(98)00676-0).
- [6] A.H. Rose, “Devitrification in annealed optical fiber”, *J. Light-wave Technol.*, vol. 15, no. 5, pp. 808–814, 1997 (DOI: 10.1109/50.580819).
- [7] OFS Fitel datasheet, Fiber-151, „50 μm graded-index OM2 – bend-insensitive multimode optical fiber”, 4/2018 (URL: <https://fiber-optic-catalog.ofsoptics.com/documents/pdf/Graded-Index-50-B0-MMF-fiber-151-web.pdf>).
- [8] –, ITU-T G.652, “Characteristics of a single-mode optical fibre and cable”, 2016 (URL: <https://www.itu.int/rec/T-REC-G.652-201611-I/en>).
- [9] O. Humbach, H. Fabian, U. Grzesik, U. Haken, and W. Heitmann, “Analysis of OH absorption bands in synthetic silica”, *J. Non-Crystalline Solids*, vol. 203, pp. 19–26, 1996 (DOI: 10.1016/0022-3093(96)00329-8).
- [10] –, “Investigation of high temperature silica based fiber optic materials”, *Final scientific/technical report*, DOE Award no. DE-FE0027891. Virginia Polytechnic Institute & State University, 2018 (URL: <https://www.osti.gov/servlets/purl/1489125>).
- [11] A. Honda, K. Toh, S. Nagata, B. Tsuchiya, and T. Shikama, “Effect of temperature and irradiation on fused silica optical fiber for temperature measurement”, *J. of Nuclear Materials*, vol. 367, pp. 1117–1121, 2007 (DOI: 10.1016/j.jnucmat.2007.03.193).
- [12] K. Saitoh, Y. Tsuchida, and M. Koshiba, “Bending-insensitive single-mode hole-assisted fibers with reduced splice loss”, *Opt. Lett.* 30, pp. 1779–1781, 2005 (DOI: 10.1364/OL.30.00177).
- [13] W. Luo, S. Li, W. Chen, D. Wang, and Q. Mo, “Low-loss bending-insensitive micro-structured optical fiber for FTTH”, *Proc. 61-st IWCS*, 2012, pp. 454–457 (URL: <https://www.yumpu.com/en/document/read/30323166/low-loss-bending-insensitive-micro-structured-optical-fiber-for-ftth>).
- [14] M.-J. Li, *et al.*, “Ultra-low bending loss single-mode fiber for FTTH”, *J. Lightwave Technol.*, vol. 27, no. 3, pp. 376–382, 2009 (DOI: 10.1109/JLT.2008.2010413).
- [15] IEEE Computer Society, IEEE Standard for Ethernet – IEEE Std. 802.3-2018 (URL: <https://standards.ieee.org/ieee/802.3/7071/>).
- [16] ISO/IEC 11801, Information technology – Generic cabling for customer premises, 2017 (URL: <https://www.iso.org/standard/66182.html>).

- [17] –, EN-IEC 60793-1-40, “Optical fibres – Part 1–40: Attenuation measurement methods”.
- [18] –, EN 60793-1-46, “Optical fibres – Part 1–46: Measurement methods and test procedures – Monitoring of changes in optical transmittance”.
- [19] –, EN-IEC 60794-1-20, “Optical fibre cables – Part 1–20: Generic specification – Basic optical cable test procedures – General and definitions”.
- [20] –, EN 50582, “Procedure to assess the circuit integrity of optical fibres in a cable under resistance to fire testing”.
- [21] –, Recommendation ITU-T G.657, “Characteristics of a bending-loss insensitive single-mode optical fibre and cable”, 2016 (URL: <https://www.itu.int/rec/T-REC-G.657-201611-I/en>).
- [22] –, IEC 60793-2-10, “Optical fibres – Part 2–10: Product specifications – Sectional specification for category A1 multimode fibres”.
- [23] –, EN-IEC 60793-2-50, “Optical fibres – Part 2–50: Product specifications – Sectional specification for class B single-mode fibres”.

Krzysztof Borzycki, Ph.D.

Assistant Professor at the NIT Central Chamber for Telecommunication Metrology

 <https://orcid.org/0000-0001-6066-6590>

E-mail: k.borzycki@il-pib.pl

National Institute of Telecommunications, Warsaw, Poland

<https://www.gov.pl/instytut-lacznosci>

Marek Jaworski, Ph.D.

Assistant Professor at the NIT Central Chamber for Telecommunication Metrology

 <https://orcid.org/0000-0002-6742-4874>

E-mail: m.jaworski@il-pib.pl

National Institute of Telecommunications, Warsaw, Poland

<https://www.gov.pl/instytut-lacznosci>

Tomasz Kossek, Ph.D.

Assistant Professor at the NIT Central Chamber for Telecommunication Metrology

 <https://orcid.org/0000-0001-6670-2871>

E-mail: t.kossek@il-pib.pl

National Institute of Telecommunications, Warsaw, Poland

<https://www.gov.pl/instytut-lacznosci>

Information for Authors

Journal of Telecommunications and Information Technology (JTIT) is published quarterly since 1999. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review.** Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

We encourage submissions from a diverse range of authors from across all countries and backgrounds.

Manuscript

Latex files are preferred and Editorial Office provides a style to prepare the material along with the documentation. We also accept Microsoft Word and PDF files. A typical article is 10 pages long (approximately 6,000 words) and must include the following contents:

- Authors' names and affiliations in the following format:
Name and surname, academic title,
Position held,
ORCID number,
E-mail address from the University's domain,
Faculty and name of the University,
Link to University website.
- Abstract (150-200 words). The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.
- Keywords related to the content of the article. About four keywords or phrases in alphabetical order should be used, separated by commas.
- The content of the article in a typical structure, i.e.: introduction, related work, conducted research, conclusions, references.

Figures, tables and photos

Together with the article, please send files with graphics with the highest resolution available, 150 dpi or more in bitmap resolution (jpg, png) and vector (cdr, svg, ps, pdf) formats are welcomed.

References

We use four main citation styles for a journal article, for an Internet article, for a conference paper, and for a book. Below are examples of citations. In each item, the DOI number or link to the PDF of the cited article should be provided.

- [1] R.K. Meyers and A.H. Desoky, "An implementation of the blowfish cryptosystem", *2008 IEEE International Symposium on Signal Processing and Information Technology*, 2008 (<https://doi.org/10.1109/IS-SPIT.2008.4775664>).
- [2] K. Nowicki and T. Uhl, *Ethernet End-to-End*, 1st ed. Germany, Shaker-Publisher, 2008 (ISBN: 978383832271404).
- [3] C. Shorten and T.M. Khoshgoftaar, "A survey on image data augmentation for deep learning", *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019 (<https://doi.org/10.1186/s40537-019-0197-0>).

- [4] S. Wong *et al.*, "Traffic forecasting using vehicle-to-vehicle communication", *3rd Annual Conference on Learning for Dynamics and Control*, pp. 917–929, 2021 (<https://arxiv.org/pdf/2104.05528>).

The paper with full version and anonymous (blind) PDF version for the blind review process should be submitted by e-mail: redakcja@il-pib.pl or editorial.office@il-pib.pl.

Reviewing Process

The article is initially approved by the Editor-In-Chief and if the decision is positive, is then sent to the reviewers. Depending on the subject of the article, it takes few weeks. In the next step, reviews are showed to authors who have 2 weeks to correct the article. Finally, the corrected text can be re-presented to the reviewer for reevaluation, which will take another 2 weeks.

As a result, after about 3 months, we are able to send the text for publication in the upcoming issue of Journal.

When the reviews are inconsistent, additional corrections are necessary, or the reviewer expects additional verification because the corrections ordered by the author are insufficient or additional problems arise, the review of the article may be extended by another month or more.

Editorial Work

Positively reviewed and corrected article is next prepared by the editorial office for publication. At the end of this process the author receives an edited version for approval.

Copyright formular

Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the authors agree to automatically transfer the copyright for their article to the publisher, if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and all translation rights. No part of the present JTIT should not be reproduced in any form nor transmitted or translated into a machine language without prior written consent of the publisher.

Ghostwriting declaration

We require formal declaration that the process of writing the paper was not influenced by any third party. In the article, all the contributions of other people are clearly indicated. The theories presented, methods used, analysis and research, as well as the copyrights to the drawings, photographs and other figures belong to the authors or are clearly credited in the text. The author must also indicate whether his work has received financial support and if the realization of the whole project was possible thanks to the permission and cooperation with scientific institutions, associations and others.

Other information

- The JTIT being an Open Access Journal (OAJ) has no article processing charges (APCs). The published articles can be downloaded freely without any payment.
- Journal of Telecommunications and Information Technology has entered into an electronic licensing relationship with EBSCO Publishing, the world's most prolific aggregator of full text journals, magazines and other sources. The text of Journal of Telecommunications and Information Technology can be found on EBSCO Publishing databases. For more information on EBSCO Publishing, please visit www.epnet.com.

An Adaptive Richardson-Lucy Algorithm for Medical Image Restoration

Q.A. Yaqoub and A.A. Al-Ani

Paper

66

Preserving Integrity of Evidence with Blockchain Technology in Cloud Forensics for Immigration Management

S.M. Shinde and V.R. Gurrula

Paper

78

Remedies to Thermal Radiation in Fused Silica Optical Fibers

K. Borzycki, M. Jaworski, and T. Kossek

Paper

88



National Institute
of Telecommunications

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

phone +48 22 512 81 83
fax +48 22 512 84 00
e-mail: redakcja@il-pib.pl
editorial.office@il-pib.pl
<https://www.gov.pl/web/instytut-lacznosci>