JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

4/2022



(*Contents Continued on Back Cover*)

on-line: ISSN 1899-8852

Paper

# Modeling the Geometry of an Underwater Channel for Acoustic Communication

Hala A. Naman[1,2] and A.E. Abdelkareem[1]

[1] *College of Information Engineering. Al-Nahrain University, Iraq,*
[2] *College of Engineering. University of Wasit, Iraq*

https://doi.org/10.26636/jtit.2022.164822

**Abstract — The achievement of efficient data transmissions via underwater acoustic channels, while dealing with large data packets and real-time data fed by underwater sensors, requires a high data rate. However, diffraction, refraction, and reflection phenomena, as well as phase and amplitude variations, are common problems experienced in underwater acoustic (UWA) channels. These factors make it difficult to achieve high-speed and long-range underwater acoustic communications. Due to multipath interference caused by surface and ocean floor reflections, the process of modeling acoustic channels under the water's surface is of key importance. This work proposes a simple geometry-based channel model for underwater communication. The impact that varying numbers of reflections, low water depth values, and distances between the transmitter and the receiver exert on channel impulse response and transmission loss is examined. The high degree of similarity between numerical simulations and actual results demonstrates that the proposed model is suitable for describing shallow underwater acoustic communication environments.**

***Keywords — channel impulse response, geometry channel model, multipath propagation, underwater acoustic communication.***

## 1. Introduction

Due to the increasing number of civilian and military applications of ocean monitoring solutions, underwater wireless communication has become, in recent years, a research domain with great prospects. Underwater acoustic channels are a difficult environment for ensuring reliable communications, due to the significant influence of multipath propagation, Doppler shifts, and high signal attenuation [1], [2]. Such parameters as efficiency and capacity of communication channels are affected by the abovementioned factors, as the time dispersion of the transmitted signal caused by multipath propagation leads to frequency-selective fading in the channel's frequency response. High frequencies are largely absorbed at large distances at the sea level, meaning they cannot be used in underwater channels. The usable band for underwater communications ranges from tens of Hz to 1 MHz, and this is why the vast majority of underwater communication signals are of the acoustic variety, since only such a solution is capable of working efficiently at distances of many kilometers [3]. Models highlighting the influence of real underwater environment parameters on underwater acoustic channels serve as an efficient channel analysis method. A signal traveling from the source to the receiver does not always follow the shortest path available. Usually, it reflects from the bottom and the surface of the water, meaning that it suffers from refraction caused by propagation speed differences [4]. Therefore, the received signal is influenced by the presence of its multiple delayed replicas, a phenomenon known as inter-symbol interference (ISI). In addition, significant multipath signal amplitude causes serious signal corruption-related issues in the case of signals propagating in shallow water [5], [6]. Consequently, the modeling an underwater transmission system's channel is necessary for analyzing the multipath phenomena and for coming up with effective solutions facilitating underwater acoustic communication.

## 2. Related Work

Paper [7] proposes a simple, effective, geometry-based channel model for time-reversal communication by analyzing transmitter and receiver height at different distances. This research also proposes time reversal and OFDM communication, which work together rather well. According to simulations, TR-OFDM can produce a BER of less than 0.001 at SNR exceeding 10 dB. In [8], a geometry-based model is developed to represent multipath scattering between the transmitter and the receiver. Such an approach allows to explore the impact of scattering environments on propagation characteristics with minimum complexity, by using a rectangle to characterize the communication environments of the ocean's vertical cross section, where scatterers are randomly distributed on the surface and on the bottom. In [9], the angle of arrival distribution has been calculated for the propagation path model in a UW channel that experiences incoherent scattering from the surface and the bottom. The scattering would be directed for the medium frequency range and an acoustically flat surface/bottom. Incoherent scattering significantly contributes to the scattered field with high-frequency acoustic waves or rough boundaries. This approach has the potential to aid the development of precise and cost-effective systems for high-frequency communication between mobile UW and surface platforms. Paper [10] proposed a non-stationary wide-

band channel model for UWA in shallow water. Multiple motion effects, time-varying angles, distances, cluster locations with channel geometry, and ultra-wideband properties are incorporated in this geometry-based stochastic model, making it more realistic and capable of sustaining long time/distance simulations.

This paper proposes a simple time-varying geometry model using a triangle concept to define the characteristics of a shallow underwater acoustic channel environment. Assuming that the multipath is randomly distributed on the sea's surface and bottom, the direct path components propagate by reflections on the sea surface, and multipath components propagate by reflections on the sea floor. Finally, all of them come together to form the proposed model. Channel impulse response and transmission loss are considered in the proposed model as well.

This paper is organized as follows: Section 3 describes the theory of channel characteristics used in UWA communication. The proposed geometry channel model is shown in Section 4. Section 5 compares the proposed model with the Bellhop simulator result and shows simulation results. Section 6 presents the paper's final findings and discussions.

## 3. Channel Characteristics in UWA Communication

To introduce the basic channel model, it is necessary to understand its properties, such as delay of propagation, absorption, spreading loss, and transmission loss. The Doppler effect and inter-symbol interference (ISI) take place, since the channel has a time-varying multipath, delays disperse over 100 symbol times, and the acoustic speed is slow [11]. Figure 1 depicts the architecture of this propagation model. In this case, it is assumed that the propagation vibrations are a combination of direct path reflections and multipath reflections. Because of bandwidth limitations, the parameters of the signal's sensitivity to such a multipath propagation environment change with time and vary at transmitter and receiver positions. Vertical and horizontal connection configurations are the primary factors taken into consideration while establishing the rate of multipath propagation [12]. On the other hand, multipath spreading in horizontal channels might be substantially longer than those in vertical ones. This initial sea surface reflection is shown in green in Fig. 1. The red line represents the first shallow water bottom reflection of two forms of reflection.

Transmission loss (TL) characterizes the decay of signal intensity with increasing distance. Such parameters as total spreading and attenuation loss are used to determine this value, as demonstrated in Eq. (1) in [13]. During propagation, the inevitable frictional conversion of sound converts it into heat.

$$\mathrm{TL}(f) = 2\alpha_z + k\log(r) + a(f)r. \tag{1}$$

As both the transmitter and receiver are toroidal beam-shaped transducers, attenuation $\alpha_z$ in the $z-y$ and $z-x$ planes is fixed at 15 dB, $r$ is the propagation range in meters and $k$

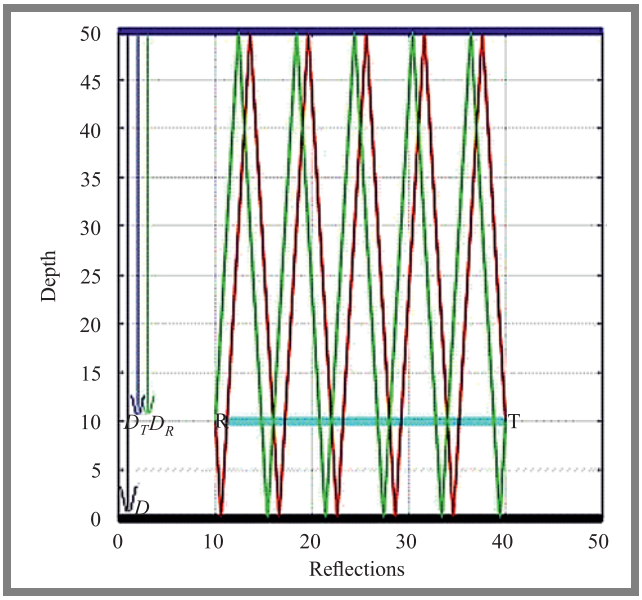



**Fig. 1.** Multipath reflection phenomena scheme in the communication channel.

represents the spreading factor selected describing the spreading loss. Spherical or cylindrical spreading determines the spreading loss in the second part of Eq. (1). When the boundaries of all reflected waves are far away from the transmitter and the receiver, meaning that no channeling of acoustic energy can occur, we say that the sound has spread spherically. In shallow water, this impact may be higher at higher frequencies because of the frequency-dependent attenuation per unit of distance. At $k = 20$, the fundamental loss law for spherical spreading is the "inverse-square law" which defines intensity $I(r)$ at range $r$ as a percentage of the intensity at the standard reference range of 1 m. When the sea acts as an acoustic waveguide due to surface and seafloor reflections, spherical propagation conditions no longer apply. In this case, propagation may be defined according to the cylindrical spreading law, with $k = 10$. Considering that the boundary reflections are highly dependent on both the state of the sea and the characteristics of the bottom material, and that low-frequency sound travels very well through the sea floor, the cylindrical law is designed to be conservative in estimating losses. A useful rule is do adopt $k = 15$, i.e. a value a which an equidistant point between the spherical and cylindrical laws exists.

The third part of Eq. (1) represents attenuation in the sea – a phenomenon caused mainly by viscous friction, occurring at frequencies exceeding 1 MHz. At lower frequencies, molecular resonance results in the fact that attenuation of sea water is lower than that of purified water. Magnesium sulfate present in the sea water solution causes additional attenuation below 500 kHz, beyond the loss in pure water, finally increasing its level by a factor of approx. 18 for frequencies lower than 70 kHz. In spite of its negligible concentration in water, boric acid reliably boosts the loss at frequencies of less than 700 Hz, by a factor of 16. Based on those relations, Eqs.

(2)–(10) are used to determine $\alpha(f)$ [13]:

$$\alpha(f) = \alpha_1 + \alpha_\alpha + \alpha_3, \tag{2}$$

where:

$$\alpha_1 = af^2 \quad \text{(fresh water attenuation)}, \tag{3}$$

$$\alpha_2 = \frac{bf_0}{(1+(f_0/f)^2)} \quad \text{(magnesium sulfate relaxation)}, \tag{4}$$

$$\alpha_3 = \frac{cf_1}{(1+(f_1/f)^2)} \quad \text{(boric acid relaxation)}, \tag{5}$$

$$a = 1.3\times10^{-7} + 2.1\times10^{-10}(T-38^\circ)^2, \tag{6}$$

$$b = 2S\times10^{-5}, \tag{7}$$

$$f_0 = 50\times(T+1), \tag{8}$$

$$c = 1.2\times10^{-4}, \tag{9}$$

$$f_1 = (10)^{(T-4)/100}. \tag{10}$$

Salinity is denoted by $S$, temperature by $T$ [°C], and frequency is denoted by $f$ [kHz].

## 4. Proposed Geometry Channel Model

The proposed geometry-based propagation model uses a triangle concept to represent the various contexts of communication pertaining to shallow water propagation channels. It assumes that multipath reflection occurs randomly at the sea surface and at various depths. Direct path propagation delay $T_D$ is determined by using the velocity formula $v = d/t$ where $v$ denotes the velocity, $d$ indicates the displacement, and $t$ refers to the time:

$$T_D = (D_T - D_R)/c, \tag{11}$$

where the depth of the transmitter is $D_T$, the depth of the receiver is $D_R$, and $c$ denotes the speed of sound. This model assumes that the $D_R$ parameter is not constant and its value is relative to $D_T$. The angle $\theta = 90^\circ$ is between a column of water and the horizontal direction between the transmitter and receiver, as shown in Fig. 2.

First, by connecting T and R points with $L$, we get a right-sided triangle. Next, using the definition of cosine, $D_R$ can by determined:

$$D_R = D_T - L\cdot\cos(\theta), \tag{12}$$

where $L$ is the distance between the transmitter and the receiver.

According to Fig. 3, the delay in transmission between the transmitter and the receiver may be calculated using the Pythagorean theorem of right triangles, whether the signal is traveling over a straight path or a multipath. In this case, the hypotenuse reflects the time it takes for each reflection to complete its propagation. There are two types reverberations: surface reflections, in which the initial reflection occurs on the surface, and bottom reflections, in which the initial reflection occurs on the sea floor.

In the constructed triangle, $d_1$ is the propagation range created by the initial reflection between the transmitter and the surface. The magnitude of the vertical motion on the first reflection line segment is equal to $D_T$ and next to receiver R via $d_2$.

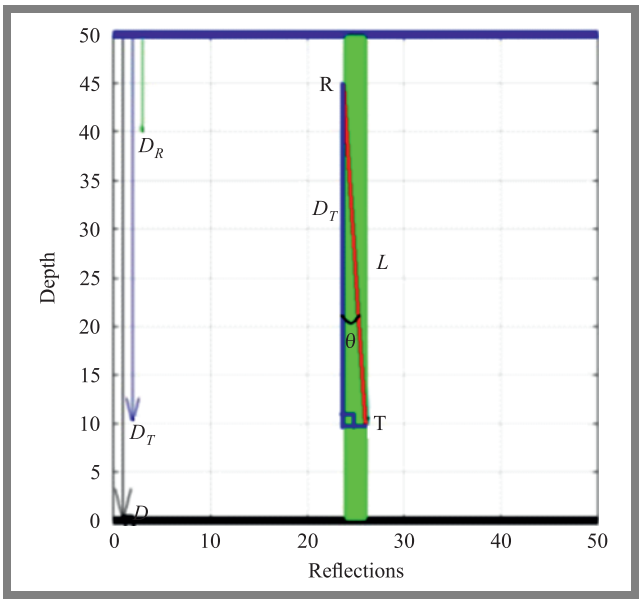



**Fig. 2.** Method for determining receiver depth.

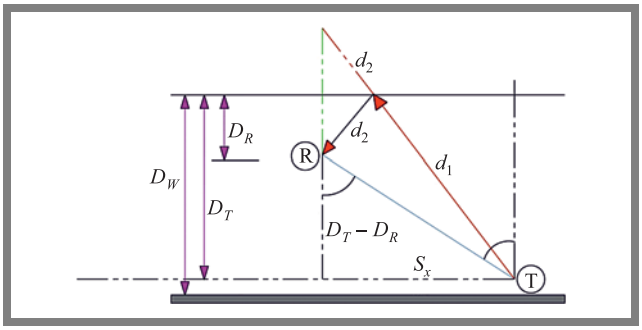



**Fig. 3.** Pythagorean theorem used for determining multipath surface reflections.

The magnitude of the vertical motion on the second reflection line segment is $D_R$. After drawing the reflected distance, $d_2$ is supplied to $d_1$ as an extension. So, the sum of the vertical motion on these two segments $d_1$ and $d_2$ is $D_T + D_R$. The side designated by $S_y$ is perpendicular to the hypotenuse, while the side designated by $S_x$ is next to it:

$$S = \sqrt{S_x^2 + S_y^2}, \tag{13}$$

$$S_y = D_T + D_R. \tag{14}$$

The method of calculating $S_x$ is depicted in Fig 3. Let us consider two transient vertical line segments from points T and R that are parallel to each other and are intersected by a blue diagonal line. The resulting angles will be equal in pairs (alternate angles), i.e. $\theta = \theta_1$ and:

$$\begin{aligned}
\theta_1 + \theta_2 &= 90^\circ \Longrightarrow \theta_1 = 90^\circ - \theta_2,\\
\mathrm{tg}(\theta - 1) &= \frac{\sin(\theta-1)}{\cos(\theta_1)} = \frac{\cos(90^\circ = \theta_2)}{\sin(90^\circ - \theta_2)}\\
&= \mathrm{ctg}(90^\circ - \theta_2) = \mathrm{ctg}(\theta_2),\\
\mathrm{ctg}(\theta_2) &= \frac{S_x}{D_t - D_r},\\
\theta &= \theta_1 \Longrightarrow \mathrm{tg}(\theta) = \frac{S_x}{D_t - D_r},\\
S_x &= (D_T - D_R)\,\mathrm{tg}(\theta).
\end{aligned} \tag{15}$$

For the two reflections ($N = 2$) , the multipath propagation is represented in Fig. 4.

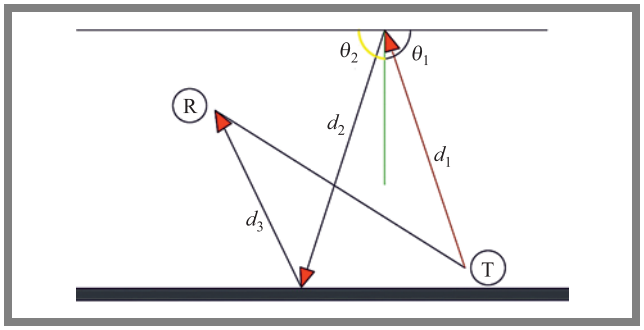



**Fig. 4.** Multipath propagation for $N = 2$.

During multipath propagation, $S_x$ is unaffected by reflections, whereas $S_y$ varies, which results in an increase of $d_1$, $d_2$ and $d_3$. Suppose $\theta_1$ and $\theta_2$ are two consecutive reflection angles, $d_1$ with $\theta_1$ and $d_2$ with $\theta_2$ are trigonometric pairs, such that:

$$\begin{aligned} \theta_2 &= 180^\circ - \theta_1 \Longrightarrow \sin(\theta_2) = \sin(\theta_1) \text{ and } \cos(\theta_2) \\ &= -\cos(\theta_1), \\ \operatorname{tg}(\theta_2) &= -\operatorname{tg}(\theta_1) \Longrightarrow m(d_2) = -m(d_1), \end{aligned}$$

where $\theta_1$ and $\theta_2$ are two arbitrary consecutive reflection angles. Then:

$$\forall i \in \mathbb{N} m(d_{i+1}) = -m(d_i).$$

Both consecutive reflection lines have the same slope as the opposite sign. Suppose $m(d_i)$ represents the slope of $d_i$ and:

$$\begin{aligned} \forall i \in \mathbb{N} |m(d_{i+1})| &= |m(d_i)|, \\ \forall i \in \mathbb{N} m(d_{i+2}) &= m(d_i), \\ \forall i \in \mathbb{N} \forall j \in \mathbb{N} |m(d_i)| &= |m(d_j)|, \\ \forall i \in \mathbb{N} |m(d_i)| &= |m(d_1)|. \end{aligned}$$

All the reflection lines are along the first reflection line, thus allowing to create vector **S** that moves horizontally and vertically, as shown in Fig. 5.

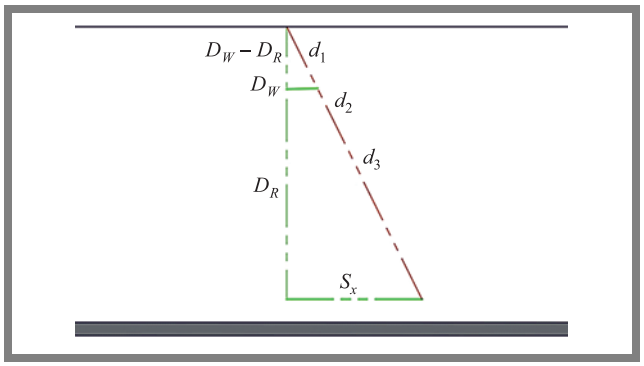



**Fig. 5.** Equivalent model for multipath propagation for $N = 2$.

For any number of surface reflections $n$ (between 1 and $N$) , the propagation delay is:

$$S_{y_n} = D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + -1^{n+1} D_R, \tag{16}$$

where $D_W$ is the water depth.

Suppose $n$ increases by two units. Therefore, two consecutive reflection vectors with opposite slopes are added to the previous set of reflections. The first reflection vector coming from the sea level to the depth and the second reflection vector going from the depth to the sea level. Each of these two shifts by $D_W$ on the vertical axis. Hence, total displacement increases by $2D_W$, as shown in Fig. 6.

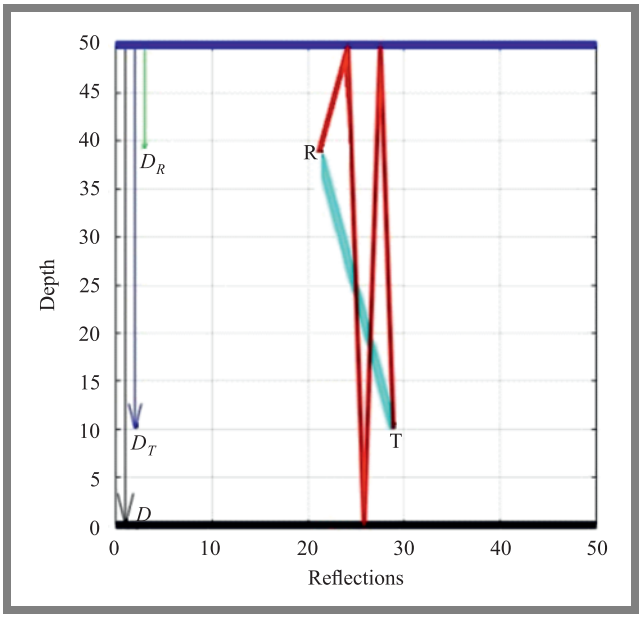



**Fig. 6.** Surface reflections when $N = 3$.

$$\begin{aligned} S_{y_{n+2}} &= S_{y_n} + 2D_W \quad (N = 3), \\ S_{y_{n+2}} &= D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + -1^{n+1} D_R + 2D_W, \\ S_{y_{n+2}} &= D_T + \left( \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) 2D_W + \left( -1^{n+1} D_R \right), \\ S_{y_{n+2}} &= D_T + \left\lfloor \frac{n+2}{2} \right\rfloor 2D_W + \left( -1^{n+1} D_R \right). \end{aligned}$$

Using induction on a natural odd number:

$$\forall n \in \mathbb{N}_O, S_{y_n} = D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + \left( -1^{n+1} D_R \right).$$

Suppose $n = 2$, as shown in Fig. 7. From Eq. (16):

$$\begin{aligned} S_{y_2} &= D_T + \left\lfloor \frac{2}{2} \right\rfloor 2D_W + \left( -1^{2+1} D_R \right), \\ S_{y_2} &= D_T + 2D_W + \left( -1^3 D_R \right), \\ S_{y_2} &= D_T + 2D_W - D_R. \end{aligned}$$

Assume $n$ is an even natural number, i.e. $n \in \{2, 4, 6, \dots\}$. If $S_{y_n}$ is correct, $S_{y_{n+2}}$ also is a correct formula.

$$S_{y_n} = D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + \left( -1^{n+1} D_R \right).$$

Now suppose n increases by another two units. Consequently, two consecutive reflection vectors with opposite slopes are added to the previous set of reflections. The first reflection vector coming from the sea level to the depth and the second reflection vector going from the depth to the sea level. Each of these two shifts by $D_W$ on the vertical axis and total displacement increases by $2D_W$ on the vertical axis, as shown in Fig. 8.

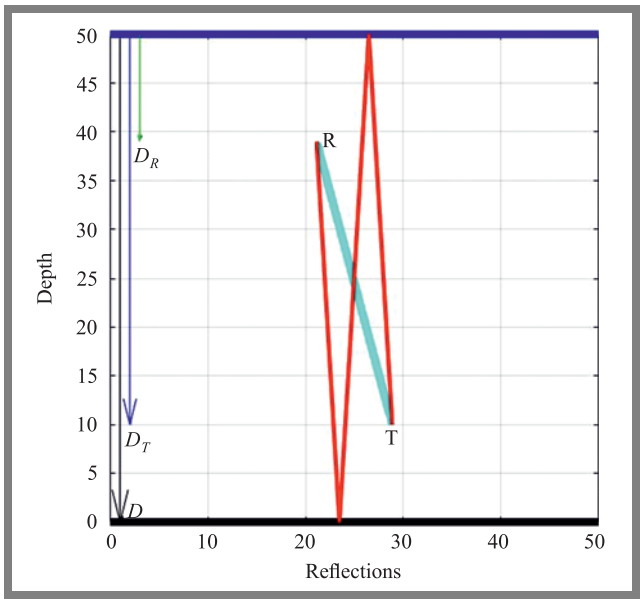



**Fig. 7.** Surface reflections when $N = 2$.

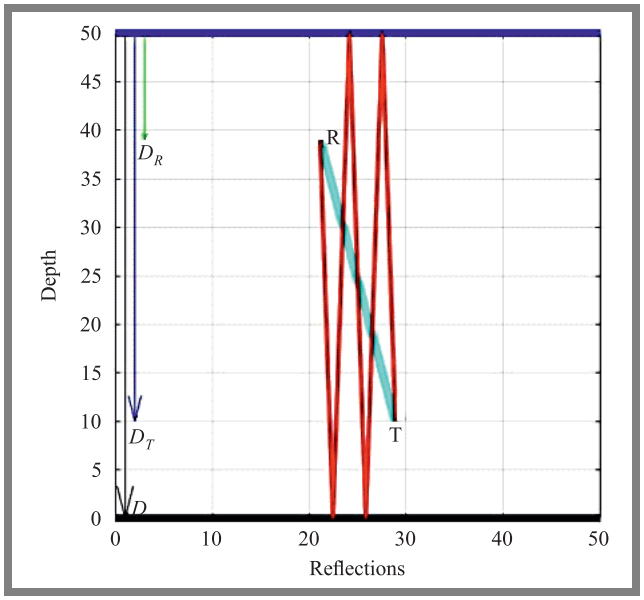



**Fig. 8.** Surface reflections when $N = 4$.

$$\begin{aligned} S_{y_{n+2}} &= S_{y_n} + 2D_W \\ S_{y_{n+2}} &= D_T + \left(\left\lfloor \frac{n}{2} \right\rfloor 2D_W\right) - (D_R) + 2D_W. \end{aligned}$$

Using induction on a natural even number:

$$\forall n \in \mathbb{N}_E; \mathrm{BS}_{y_n} = D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + D_R,$$

we conclude that Eq. (16) is valid for all natural numbers.

The propagation delay for any number of reflections is calculated using the velocity formula and Eq. (13):

$$\begin{aligned} S &= \sqrt{S_x^2 + S_y^2} = \qquad (17) \\ &= \sqrt{[(D_T - D_R)\,\mathrm{tg}(\theta)]^2 + \left[D_T + \left\lceil \tfrac{n}{2} \right\rceil 2D_W + (-1^{n+1} D_R)\right]^2} \end{aligned}$$

$$T_{Sn} = \frac{\sqrt{(D_T - D_R)^2\,\mathrm{tg}^2(\theta) + \left(D_T + \left\lfloor \frac{n}{2} \right\rfloor 2D_W + -1^{n+1} D_R\right)^2}}{c}$$

### *4.1. Bottom Reflections Delay Propagation*

The same method is used to calculate the propagation time for a single bottom reflection. As illustrated in Fig. 9, $B_x$ is unaffected by the number of reflections, whereas $B_y$ varies.

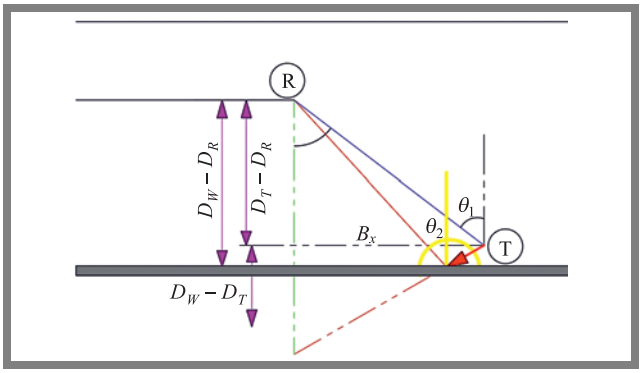



**Fig. 9.** Pythagorean theorem applied to multipath bottom reflections.

Assume $\theta_1$ and $\theta_2$ are two consecutive reflection angles

$$\begin{aligned} \theta_2 &= 180^\circ - \theta_1 \Longrightarrow \sin(\theta_2) \\ &= \sin(\theta_1) \text{ and } \cos(\theta_2) = -\cos(\theta_1), \\ \mathrm{tg}(\theta_2) &= -\mathrm{tg}(\theta_1) \Longrightarrow m(d_2) = -m(d_1), \\ \forall i \in \mathbb{N} m(d_{i+1}) &= -m(d_i). \end{aligned}$$

Both consecutive reflection lines have the same slope as the opposite sign. $m(d_i)$ represents the slope of $d_i$ then:

$$\begin{aligned} \forall i \in \mathbb{N} |m(d_{i+1})| &= |m(d_i)|, \\ \forall i \in \mathbb{N} m(d_{i+2}) &= m(d_i), \\ \forall i \in \mathbb{N} \forall j \in \mathbb{N} |m(d_i)| &= |m(d_j)|, \\ \forall i \in \mathbb{N} |m(d_i)| &= |m(d_1)|. \end{aligned}$$

All the reflection lines lie along the first reflection line allowing to create vector $\mathbf{B}$ that moves horizontally and vertically:

$$B = \sqrt{B_x 2 + B_y 2}, \qquad (18)$$

$$B_y = D_T + D_R. \qquad (19)$$

To calculate $B_x$, let us consider two transient vertical line segments from points T and R that are parallel to each other and are intersected by a blue diagonal line, as shown in Fig. 10. The resulting angles will be equal in pairs and $\theta = \theta_1$.

$$\begin{aligned} \theta_1 + \theta_2 &= 90^\circ \Longrightarrow \theta_1 = 90^\circ - \theta_2, \\ \mathrm{tg}(\theta_1) &= \frac{\sin(\theta_1)}{\cos(\theta_1)} = \frac{\cos(90^\circ - \theta_2)}{\sin(90^\circ - \theta_2)} \\ &= \mathrm{ctg}(90^\circ - \theta_2) = \mathrm{ctg}(\theta_2), \\ \mathrm{ctg}(\theta_2) &= \frac{B_x}{D_T - D_R}, \\ \mathrm{tg}(\theta) &= \frac{B_x}{D_T - D_R} \Rightarrow B_x = (D_T - D_R)\,\mathrm{tg}(\theta). \qquad (20) \end{aligned}$$

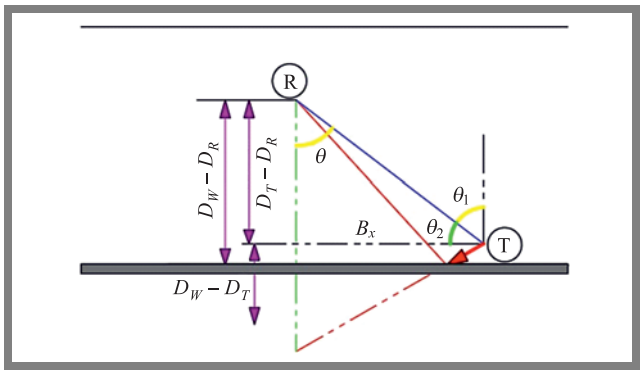



**Fig. 10.** Alternate angles for $B_x$ calculation.

For any number of bottom reflections $n$, the propagation delay is:

$$B_{y_n} = \left\lceil \frac{n}{2} \right\rceil 2D_W - \left(D_T + -1^{n+1} D_R\right). \tag{21}$$

If $n$ increases by two units, two consecutive reflection vectors with opposite slopes are added, as shown in Fig. 11. The first reflection vector is coming from the sea level to the depth of the sea and the other reflection vector is going from the depth to the sea level. Each of these two shifts by $D_W$ on the vertical axis. So, total displacement increases by $2D_W$ on the vertical axis.

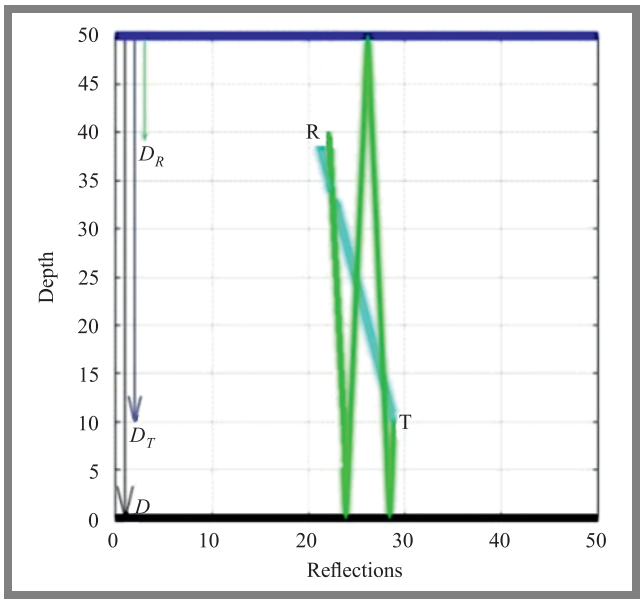



**Fig. 11.** Bottom reflection for $N = 3$.

$$\begin{aligned} B_{y_{n+2}} &= B_{y_n} + 2D_W, \\ B_{y_{n+2}} &= \left(\left\lceil \frac{n}{2} \right\rceil 2D_W\right) - \left(D_T + -1^{n+2+1} D_R\right) + 2D_W, \\ B_{y_{n+2}} &= \left\lceil \frac{n+2}{2} \right\rceil 2D_W - \left(D_T + -1^{n+1} D_R\right), \end{aligned}$$

Suppose $n$ increases by another two units (Fig. 12).

Similarly, each of these two moves by $D_W$ on the vertical axis and total displacement increases by $2D_W$ on the vertical axis.

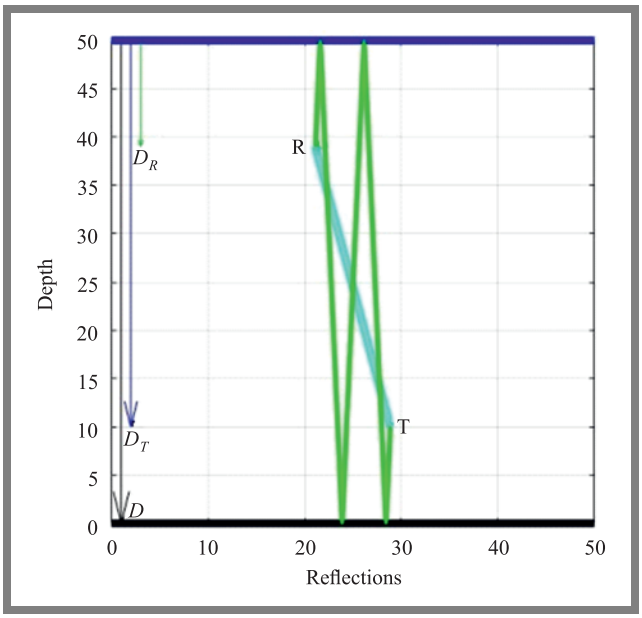



**Fig. 12.** Bottom reflection when $N = 4$.

$$\begin{aligned} B_{y_{n+2}} &= B_{y_n} + 2D_W, \\ B_{y_{n+2}} &= \left(\left\lceil \frac{n}{2} \right\rceil 2D_W\right) - \left(D_T + -1^{n+2+1} D_R\right) + 2D_W, \\ B_{y_{n+2}} &= \left\lceil \frac{n+2}{2} \right\rceil 2D_W - \left(D_T + -1^{n+1} D_R\right). \end{aligned}$$

The propagation delay for any number of reflections is:

$$B = \sqrt{B_x^2 + B_y^2} = \tag{22}$$
$$= \sqrt{\left((D_T - D_R)\operatorname{tg}\theta\right)^2 + \left(\left\lceil \tfrac{n}{2} \right\rceil 2D_W - \left(D_T + -1^{n+1} D_R\right)\right)^2}$$
$$T_{Bn} =$$
$$= \frac{\sqrt{(D_T - D_R)^2 \operatorname{tg}^2\theta + \left(\left\lceil \tfrac{n}{2} \right\rceil 2D_W - \left(D_T + -1^{n+1} D_R\right)\right)^2}}{c}$$

The main equation of channel impulse response (CIR) representing a multipath channel can be modeled as:

$$h(\tau, t) = \sum_{l=1}^{L} h_l(t)\delta[t - \tau_1(t)], \tag{23}$$

where $h_l(t)$ represents the path amplitudes, $\tau_1(t)$ denotes the time-varying path delays based on the proposed model and $L$ is the total number of paths.

## 5. Simulation Results

This section presents the results of a simulation performed with the use of Matlab for the proposed channel model. Table 1 presents the simulation parameters.

Normal incident waves at the sea surface border create a pressure reflection coefficient of $-1$, while waves reflected from the sea floor have a pressure reflection coefficient of approx. 1 [14]. Even surface reflections add up constructively at the receiver side, while odd surface and bottom reflections add up destructively.

**Tab. 1.** Simulation parameters.

| Parameter | Value |
|---|---|
| Transmitter depth | 40 m |
| Receiver depth | Related to transmitter depth – Eq. (12) |
| Temperature | 14°C |
| Salinity | 35 |
| Frequency | 12 kHz |
| Spreading factor | 15 |

The multipath channel between the source and the receiver is shown in Fig. 13 as a communication channel response. The delay spread of the reverberations in the communication channel equaled up to 79 ms.

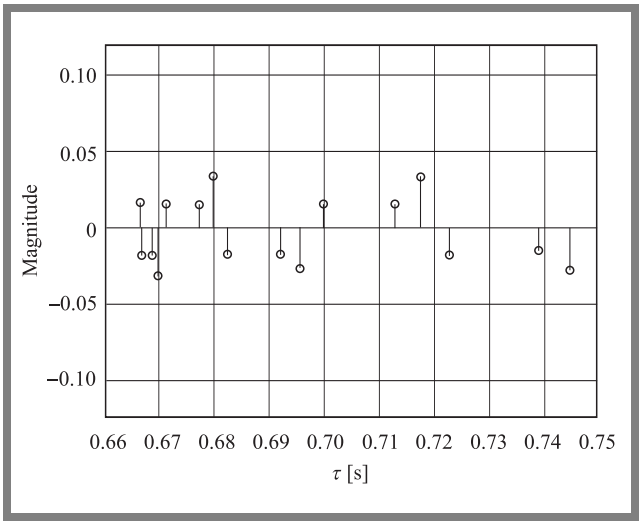



**Fig. 13.** Communication CIR with $D_W = 50$ and $L = 1000$.

To evaluate the performance of the proposed model, the same configuration parameters $(D_W, D_R, D_T, L, f)$ were implemented in the Bellhop simulator [15] as indicated in Fig. 14. Here, the red path shows the direct path or a signal that begins at the sea floor and gradually refracts upwards or diverges upwards until a steep positive sound speed gradient is met, at which point it refracts downward. The green path of the signal reaches the receiver via a reflection from the water surface and/or bottom. Black paths refer to multiple multipath components arriving at the receiver due to the increased variety of potential paths reflected off the sea surface and an uneven sea bottom. Comparison of the delay spread achieved by the newly proposed model (79 ms) and the results of the Bellhop simulator [15] (78 ms) shows that the results are very close. The difference stems from the number of reflections used to represent the propagation delay.

When simulating the proposed model at different water depths, it was found that the multipath propagation effect increases with depth. As shown in Fig. 15, the delay spread for depths of 50, 100, 150, and 200 m equals 79, 279, 535, and 824 ms, respectively. Reflection of sound at the surface and bottom and its refraction caused by spatial variation in the speed of sound in water is responsible for multipath propagation. The speed of sound varies with temperature and pressure.

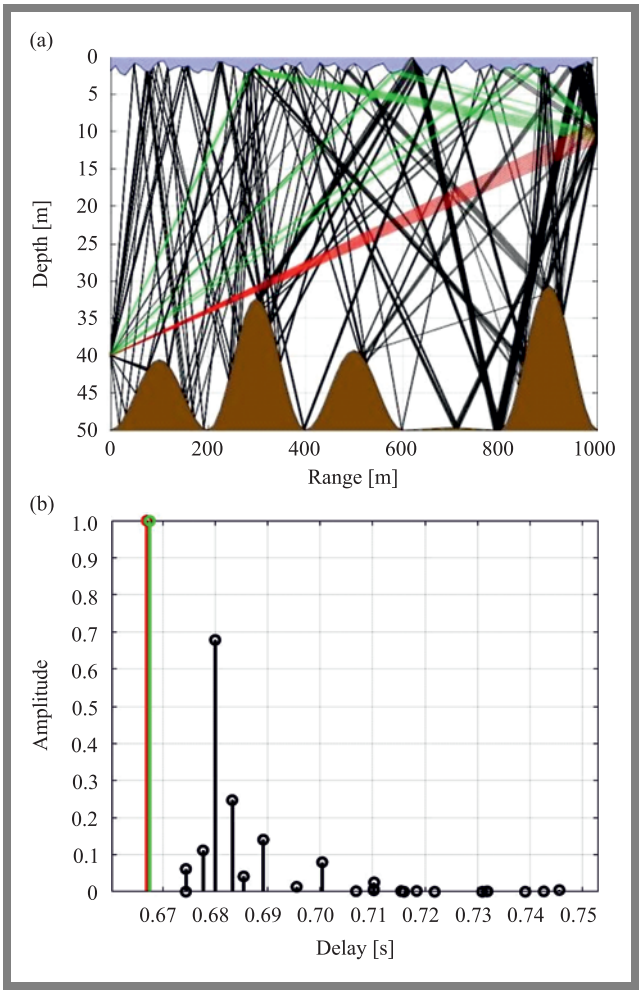



**Fig. 14.** Underwater multipath propagation environment obtained with the Bellhop simulator: a) ray trace, b) CIR.

For a depth of 50-100 m, the speed of sound remains constant, while at a depth of 150–200 m, the water temperature is lower, and therefore the speed of sound decreases.

Next, the simulations verified different distances between the transmitter and the receiver. Here, the highest delay spread for channel lengths of 500 and 1000 m was 138 and 79 ms, respectively. However, the 500 m CIR appears to be more severe, because some of its paths have a similar amplitude to the direct path, and the delay spread is longer, as shown in Fig. 16.

The transmission and reflection losses in sea water were then researched in a geometry model to describe CIR taps. It was shown that the model adequately captured the characteristics of the communication channel. In Figs. 17–18, the transmission loss increased with more multipath reflections and the growing communication distance, reaching more than 81 dB and 78 dB, respectively. The transmission loss depends on how the propagation of sound changes from spherical to cylindrical propagation. Moreover, attenuation increases with the communication distance which, in turn, increases the power loss. In shallow waters with a boundary formed by the sea floor, the grazing angle is affected as well. Figures 17–18 show that the transmission loss of the communication chan-

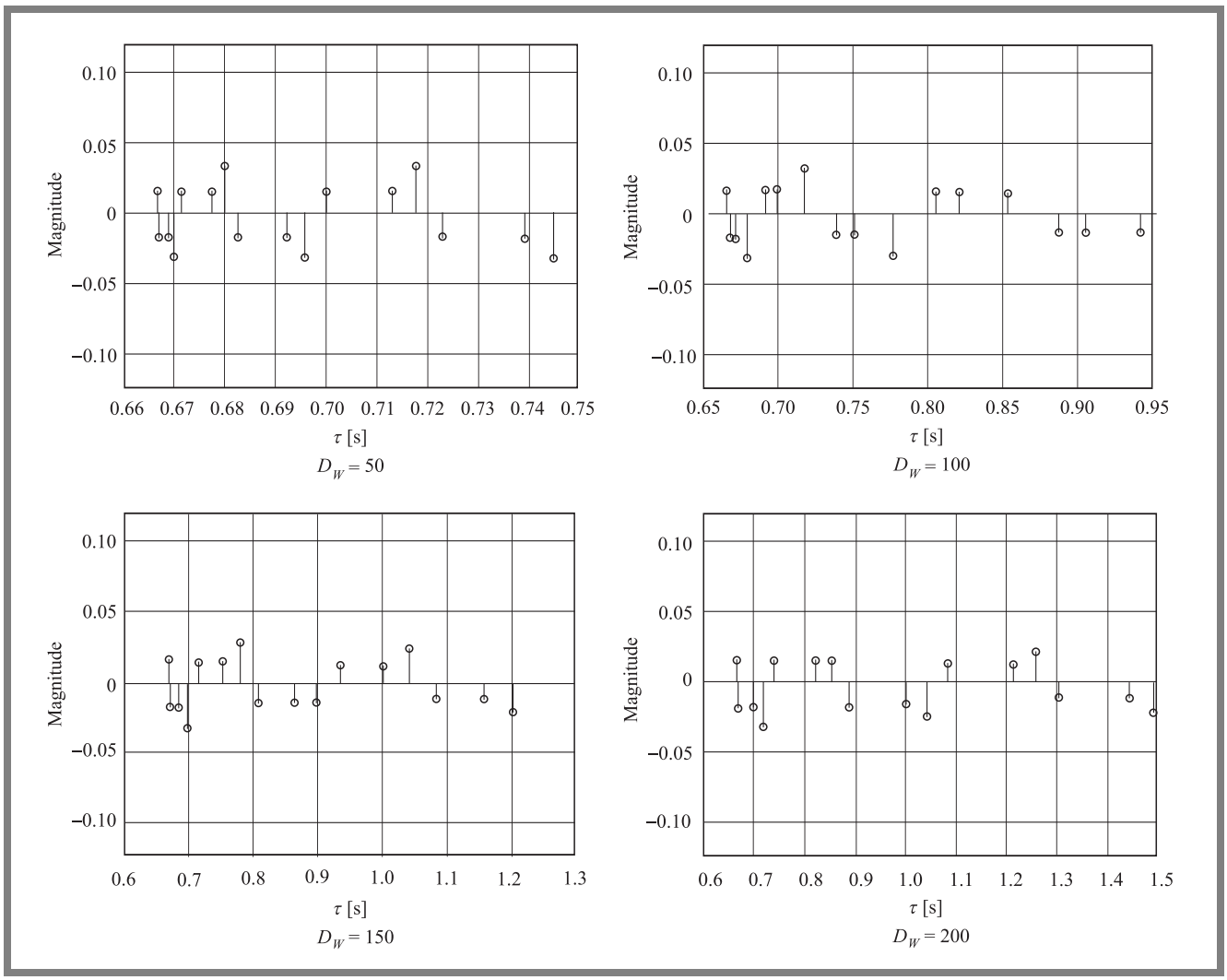



**Fig. 15.** Communication channel impulse response for $N = 20$ and $L = 1000$ with different water depths.

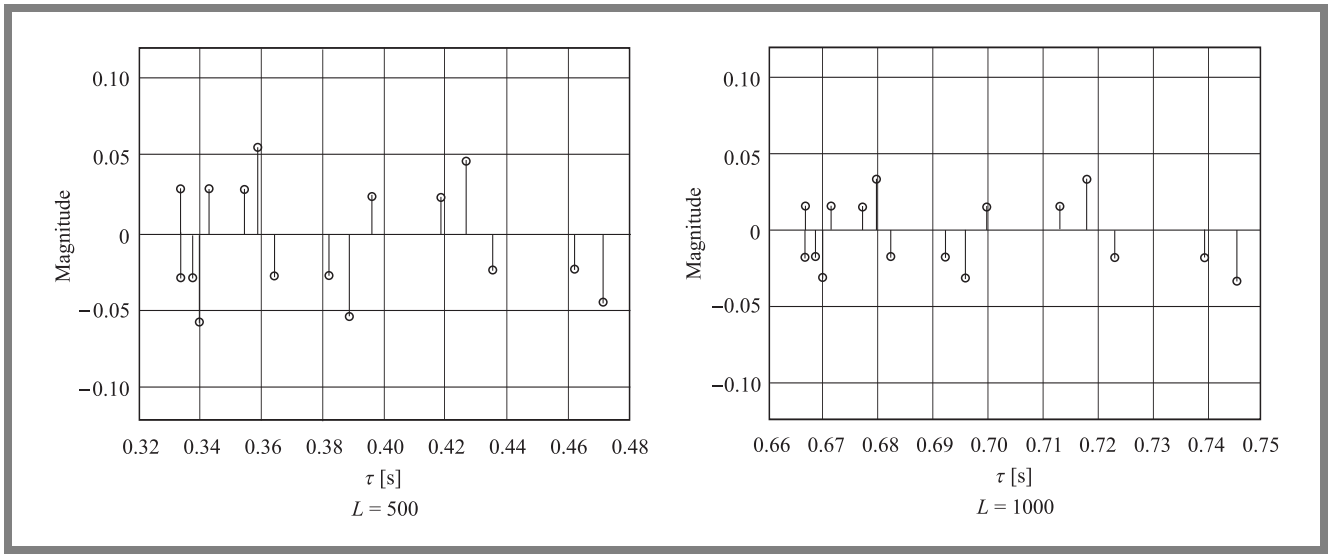



**Fig. 16.** Communication channel impulse response with $D_W = 50$ and different $L$ values.

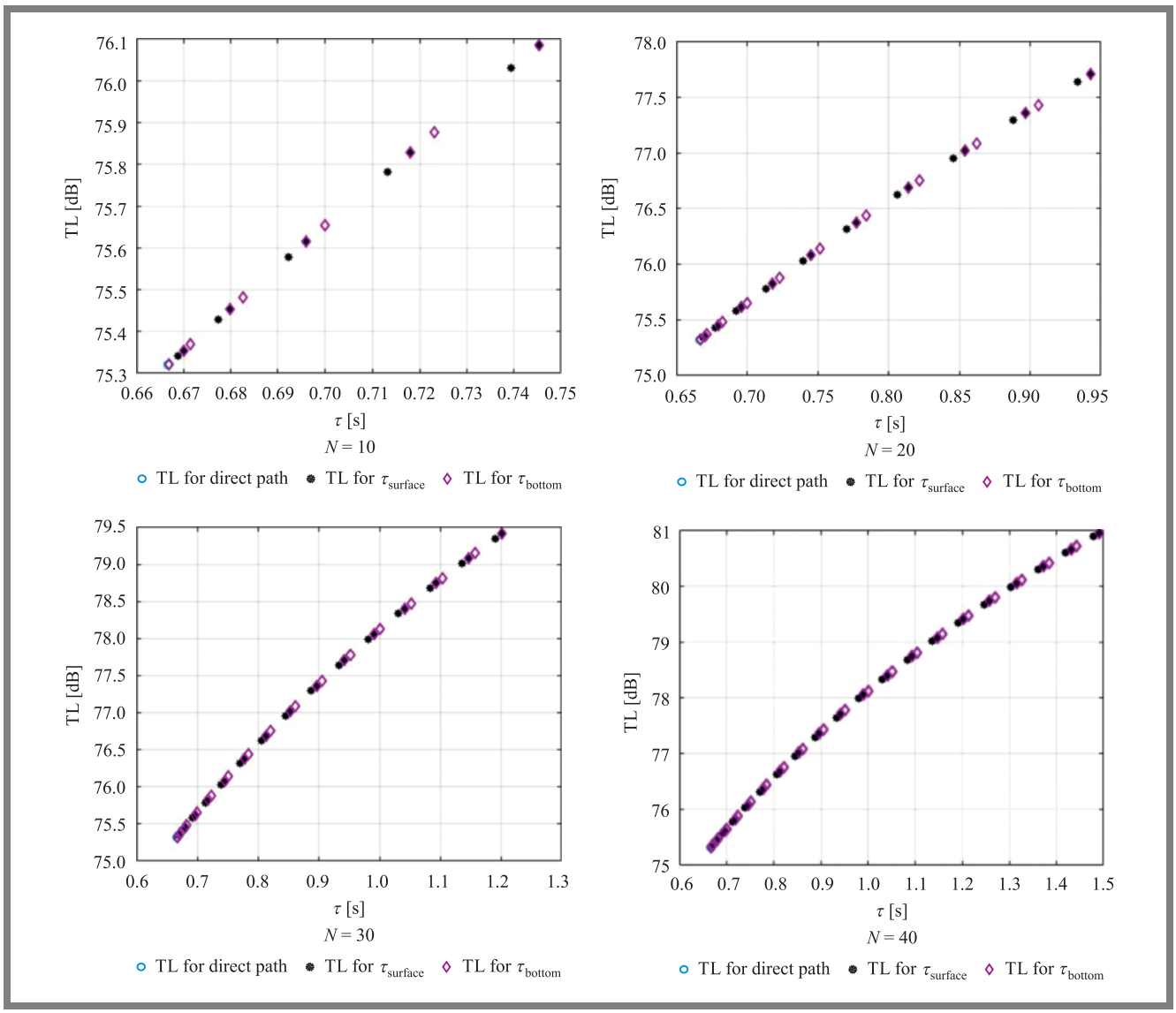



**Fig. 17.** Transmission loss of the communication channel with $L = 1000$, $D_W = 50$ with varying number of reflections.

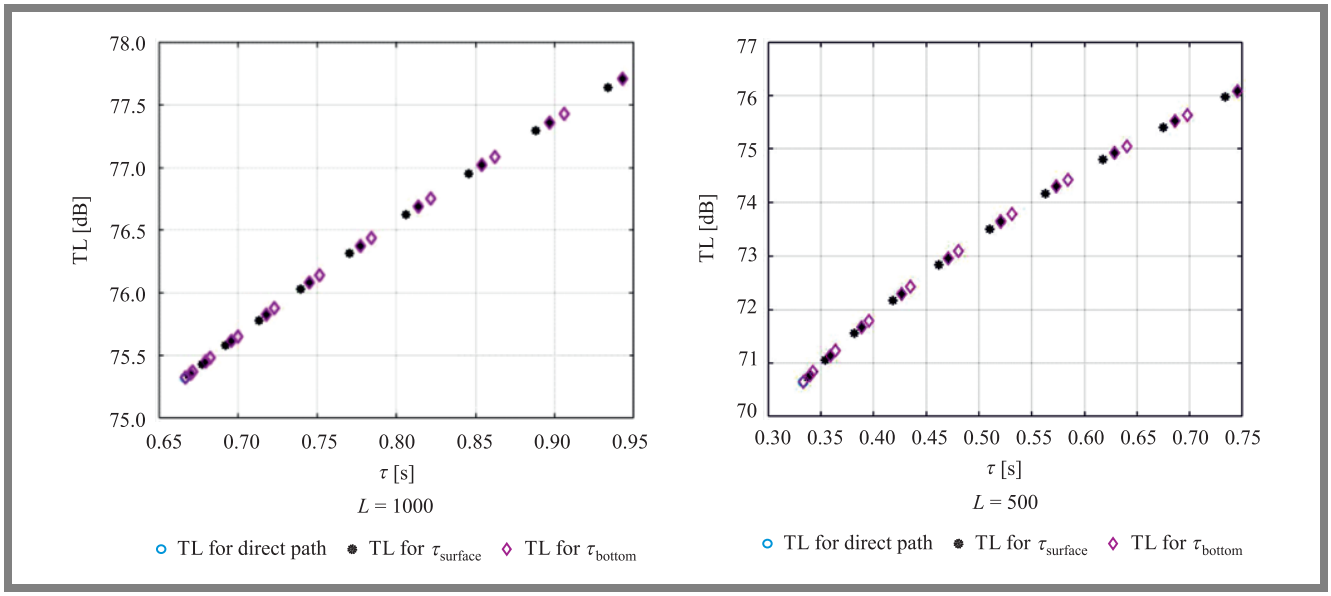



**Fig. 18.** Transmission loss of communication channel with $D_W = 50$ and different $L$.

nel increases with the growing number of reverberations and the communication distance.

## 6. Conclusions

This paper presents a geometric model of an underwater communication channel environment for multipath propagation delay simulations based on the triangle concept. It describes the multipath effect caused by surface and bottom reverberations and determines the communication channel's impulse response by changing the underwater depth and communication distance. The simulation results show that when the communication distance and the number of reflections from the surface and the bottom increase, the propagation losses increase as well. Comparisons of the propagation delay identified in the course of the simulations and the traditional results confirm that the proposed model successfully describes the communication channel's behavior in a shallow water scenario.

## References


[1] M. Stojanovic and J. Preisig, "Underwater Acoustic Communication Channels: Propagation Models and Statistical Characterization", *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, 2009 (DOI: 10.1109/MCOM.2009.4752682).

[2] A. E. Abdelkareem, B.S. Sharif, C.C. Tsimenidis, and J.A. Neasham, "Time Varying Doppler-Shift Compensation for OFDM-Based Shallow Underwater Acoustic Communication Systems", *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, 2011, pp. 885–891 (DOI: 10.1109/MASS.2011.105).

[3] B.J. Heidemann, M. Stojanovic, and M. Zorzi, "Underwater sensor networks: applications, advances and challenges", *Trans. R. Soc. A*, vol. 370, pp. 158–175, 2012 (DOI: 10.1098/rsta.2011.0214).

[4] J. Hui and X. Sheng, "Reverberation Channel", *Underwater Acoustic Channel, Singapore: Springer Nature Singapore*, 2022, pp. 175–191 (DOI: 10.1007/978-981-19-0774-6).

[5] R.F. Coates, "Underwater acoustic communication", *Sea Technol.*, vol. 35, no. 7, pp. 41–47, 1994 (URL: https://archive.org/details/sim_sea-technology_1994-07_35_7/page/n3/mode/2up).

[6] M. Elamassie, F. Miramirkhani, and M. Uysal, "Channel modeling and performance characterization of underwater visible light communications", *2018 IEEE Int. Conf. Commun. Work. ICC Work, 2018 – Proc.*, pp. 1–5, 2018 (DOI: 10.1109/ICCW.2018.8403731).

[7] Y. Widiarti, Suwadi, Wirawan, and T. Suryani, "A Geometry-Based Underwater Acoustic Channel Model for Time Reversal Acoustic Communication", *Proceeding – 2018 Int. Semin. Intell. Technol. Its Appl. ISITIA 2018*, pp. 345–350, 2018 (DOI: 10.1109/ISITIA.2018.8711067).

[8] J. Zhou, H. Jiang, P. Wu, and Q. Chen, "Study of Propagation Channel Characteristics for Underwater Acoustic Communication Environments", *IEEE Access*, vol. 7, pp. 79438–79445, 2019 (DOI: 10.1109/ACCESS.2019.2921808).

[9] M. Rawat, B. Lall, and S. Srirangarajan, "Angle of Arrival Distribution in an Underwater Acoustic Communication Channel with Incoherent Scattering", *IEEE Access*, vol. 8, pp. 133204–133211, 2020 (DOI: 10.1109/ACCESS.2020.3008602).

[10] X. Zhu, C. X. Wang, and R. Ma, "A 2D Non-Stationary Channel Model for Underwater Acoustic Communication Systems", *IEEE Veh. Technol. Conf.*, vol. 2021-April, pp. 0–5, 2021 (DOI: 10.1109/VTC2021-Spring51267.2021.9448976).

[11] A.E. Abdelkareem, B.S. Sharif, and C. C. Tsimenidis, "Adaptive time varying Doppler shift compensation algorithm for OFDM-based underwater acoustic communication systems", *Ad Hoc Networks*, vol. 45, pp. 104–119, 2016 (DOI: 10.1016/j.adhoc.2015.05.011).

[12] Y. Lou and N. Ahmed, "Basic Principles of Underwater Acoustic Communication", *Underwater Communications and Networks, Cham: Springer International Publishing*, pp. 3–33, 2022 (DOI: 10.1007/978-3-030-86649-5_1).

[13] R.F. W. Coates, "Underwater Acoustic Systems", *Wiley*, 1990 (ISBN: 9780470215449).

[14] D. Porta, "Underwater acoustic communications", *Sea Technol.*, vol. 39, no. 2, pp. 49–55, 1998 (URL: https://archive.org/details/sim_sea-technology_1998-02_39_2/page/n3/mode/2up).

[15] N. Morozs *et al.*, "Channel Modeling for Underwater Acoustic Network Simulation", no. June, pp. 1–25, 2020 (DOI: 10.1109/ACCESS.2020.3011620).



**Hala A. Naman** received the B.Sc. and M.Sc. degrees in computer engineering from the University of Technology, Baghdad, Iraq, in 2001 and 2014, respectively. Currently, she is a Ph.D. candidate in Information and Communication Engineering at Al-Nahrain University, Iraq, and is a working lecturer college of engineering at the University of Wasit, Iraq. Her research interests include image processing, computer network, and signal processing for underwater communications.
E-mail: haltaee@uowasit.edu.iq
College of Information Engineering. Al-Nahrain University, Iraq
College of Engineering. University of Wasit, Iraq



**A.E. Abdelkareem** received the B.Sc. and M.Sc. degrees in computer engineering from the University of Technology, Baghdad, Iraq, in 1992 and 2002, respectively. He received the Ph.D. degree from Newcatle University, England, 2012. Currently, he is working lecturer at Al-Nahrain University, Iraq. His research interests include computer architectures, signal processing for underwater communications, digital signal processors, WSN and Internet of underwater things and embedded systems using SHARC processor.
E-mail: Ammar.algassab@nahrainuniv.edu.iq
College of Information Engineering. Al-Nahrain University, Iraq

Paper

# Design Low Complexity SCMA Codebook Using Arnold's Cat Map

Sura S. Mohamed and Hikmat N. Abdullah

*Al-Nahrain University, College of Information Engineering*

https://doi.org/10.26636/jtit.2022.164422

**Abstract — In 5G wireless communications, sparse code multiple access (SCMA) – a multi-dimensional codebook based on a specific category of the non-orthogonal multiple access (NOMA) technique - enables many users to share non-orthogonal resource components with a low level of detection complexity. The multi-dimensional SCMA (MD-SCMA) codebook design presented in this study is based on the constellation rotation and interleaving method. Initially, a subset of the lattice $Z^2$ is used to form the mother constellation's initial dimension. The first dimension is then rotated to produce other dimensions. Additionally, interleaving is employed for even dimensions to enhance fading channel performance. Arnold's chaotic cat map is proposed as the interleaving method to reduce computational complexity. Performance of the SCMA codebook based on interleaving is evaluated by comparing it with selected codebooks for SCMA multiplexing. The metrics used for performance evaluation purposes include bit error rate (BER), peak to average power ratio (PAPR), and minimum Euclidean distance (MED), as well as complexity. The results demonstrate that the suggested codebook with chaotic interleaving offers performance that is equivalent to that of the conventional codebook based on interleaving. It is characterized by lower MED and higher BER compared to computer-generated and 16-star QAM codebook design approaches, but its complexity is lower than that of the conventional codebook based on interleaving.**

***Keywords — chaotic interleaving, codebook design, dimension rotation, Euclidean distance, sparse code multiple access***

## 1. Introduction

Non-orthogonal multiple access (NOMA) is a technology that may potentially help achieve high spectral efficiency and massive connectivity in 5G systems [1]. In a NOMA system, two or more users are placed on top of a single physical resource (such as power, frequency, time, or code) to offer an overloading ratio that is greater than one [2]. Unlike conventional orthogonal multiple access technologies, NOMA is capable of supporting numerous users. Power-domain multiplexing and code-domain multiplexing are the two categories into existing, dominant NOMA schemes may be divided. The equivalent schemes are power-domain NOMA and multiple access with low-density spreading [1]. NOMA is based on its ability to support many users with a limited quantity of resources. Frequency-division multiple access (FDMA), time-division multiple access (TDMA) for 2G, code-division multiple access (CDMA) for 3G, and orthogonal frequency-division multiple access (OFDMA) for 4G are all orthogonal multiple access (OMA) techniques that are most commonly used in contemporary wireless communication systems [3]. NOMA supports high connectivity rates, as it allows more people to communicate simultaneously than other OMA methods. Multiple users can share the same frequency resources by utilizing the near-far effect and by relying on cutting-edge uplink NOMA techniques [2], [3].

A non-orthogonal technique known as sparse code multiple access (SCMA) is principally based on the multi-dimensional (MD) codebook. Different bitstreams are directly mapped by SCMA to various sparse codewords, as illustrated in Fig. 1, where each user has a predefined codebook (there are 6 users). The placements of zeros in various codebooks are distinct to help users avoid colliding, and all codewords in the same codebook have zeros in the same two dimensions. Two bits are assigned to a complicated codeword for each user. Passwords of all users are multiplexed over four communal orthogonal resources [4].

## 2. Related Work

An effective codebook for the SCMA system has been widely researched in recent years. Cai *et al.* [5] developed a codebook for the optimization of the MD constellation using rotation and interleaving. A lattice that was originally intended for creating multi-dimensions for codebooks is rotated in the process. Interleaving is additionally used to improve communication over a fading channel. Performance showed that such a method offers only minor improvements compared to low-density signature (LDS). It achieves $10^{-4}$ BER at 15 dB SNR for a fading channel with reduced PAPR, while for uplink resource allocation it achieves $10^{-4}$ BER at 21 dB. In paper [8], Bonilla uses a similar approach as [6], [7], but in this case it relies on a minimum Euclidean distance (MED)-based detector at the receiver, with interleaving and phase rotation. It achieves $10^{-4}$ BER for codebooks with 4 arrays, 8 arrays, and 16 arrays with $E_b/N_o$ 5 dB, 6 dB, and 6.5 dB, respectively. In [9], Liu recommended segmentation and combining for MED enhancement based on the factor graph. He showed that irregular LDS are quite similar to the recommended approach at higher SNR, emphasizing the significance of an irregular factor graph of LDS. It achieves $10^{-4}$ BER at 10 dB SNR over a fading channel with LDPC coding. Yu [10] suggested a star-QAM-based codebook design to improve MED. Four constellation point vectors that represent the elements of

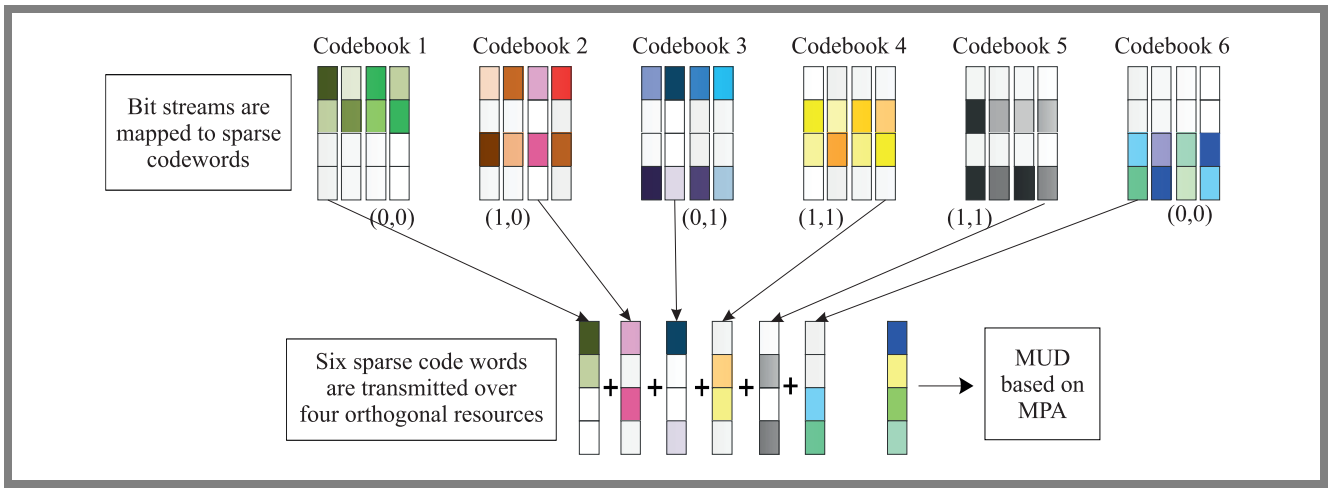



**Fig. 1.** Encoding and multiplexing in SCMA.

an $n$-dimensional star constellation is created as part of the process.

Yasmine *et al.* [11] utilized trellis coded modulation (TCM) to improve MED values and star-QAM constellations. Using this technique, they were able to obtain $10^{-4}$ BER at 8–10 dB SNR across an AWGN channel. In article [12], codebooks are designed using golden angle modulation (GAM) constellations which provide outstanding error rate performance in both uplink and downlink Rayleigh fading channels. Paper [13] illustrates development of a low complexity codebooks algorithm for almost optimum codebooks. In [14], researchers suggested to use partial Gaussian approximation, threshold-based edge selection and Gaussian approximation (ESGA) to significantly reduce receiver complexity. In paper [15], the rotation angle research zone is reduced as much as possible using a posteriori analysis and a second-order rational polynomial is then used to find a nearly optimal angle using the Levenberg-Marquardt algorithm. Such an approach is useful when fitting non-linear least squares curves. Compared to an exhaustive search, this technique is substantially simpler.

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

- In contrast to existing interleaving techniques, this work uses chaotic interleaving relying on Arnold's cat map;
- A detailed SCMA codebook design with 8 users using the proposed interleaving method is presented;
- An interleaving formula alternative to the one presented in [7] is derived;
- A formula for complexity reduction using the proposed method is given.

The structure of this paper is as follows. Section 2 describes the SCMA codebook design steps using the interleaving approach. Section 3 explains the proposed method and presents an illustrative example codebook design. Section 4 covers the simulation results, while Section 5 presents conclusions and future outlooks.

## 3. SCMA Codebook Design Using Interleaving

The SCMA system's codebook is a crucial factor for allowing the users to be multiplexed over the same resource. Therefore, it is essential in determining how users encode a bitstream using the codebook structure. The following summarizes the encoding procedure which results in the creation of the codebook. Mother constellations (MC) are produced using gray mapping coding vectors, constellation rotation and interleaving design. Then, in order to form a multidimensional base constellation, such a structure is rotated. The base constellation is rotated once more to obtain a collection of codebooks for further users, after the entire structure has been obtained for one user. This process is performed in the following manner [7]:

**Creation of the mother constellation**

1) Let $S_1$ be a subset of lattice $Z^2$ defined as:

$$S_1 = \{Qm(1+\mathrm{i})|Qm = 2m-1-M,\ m = 1,\ldots,M\}, \quad (1)$$

where $Z$ is the set of integers and $M$ is the modulation order.

2) Set vector $S_1$ from Eq. (1) based on gray mapping and assuming $M = 4$. Then the $S_1$ vector which is the starting point for constructing the MC is represented as:

$$S_{11} = -3 - 3\mathrm{i}, \quad S_{12} = -1 - 1\mathrm{i},$$
$$S_{13} = 1 + 1\mathrm{i}, \quad S_{14} = 3 + 3\mathrm{i}.$$

3) $S_N$ is generated from the preceding step and defined as $S_N = S_1 U_N$, where $U_N$ is a multidimensional phase rotation matrix:

$$U_N = \mathrm{diag}(1\mathrm{e}^{1\theta_{l-1}}) \subset C^{MN}. \quad (2)$$

4) $\theta_{l-1}$ is the factor that rotates vector $S_1$ to achieve the multi-dimension in the MC, the base vector is $S_1$, and $N$ stands for dimensions after rotating the vector. The rotating factor is defined as:

$$\theta_{l-1} = (l-1)\frac{\pi}{MN}, \quad l = 1,\ldots,N. \quad (3)$$

Setting $N = 3$ and $M = 4$ allows the angle of rotation phases in the **MC** to be calculated as:
For $l = 1$, we have $\theta_{l-1} = \theta_0 = (1-1) \times \frac{\pi}{(4)(3)} = 0$.
For $l = 2$, we have $\theta_{l-1} = \theta_0 = (2-1) \times \frac{\pi}{(4)(3)} = \frac{\pi}{12}$.
For $l = 3$, we have $\theta_{l-1} = \theta_0 = (3-1) \times \frac{\pi}{(4)(3)} = \frac{\pi}{6}$.

For $\theta_0 = 0$:

$$S_{11} = -3 - 3\mathrm{i}, \quad S_{12} = -1 - 1\mathrm{i},$$
$$S_{13} = 1 + 1\mathrm{i}, \quad S_{14} = 3 + 3\mathrm{i}.$$

When $\theta_1 = \frac{\pi}{12}$, we have:

$$S_{21} = -2.1213 - 3.6742\mathrm{i}, \quad S_{22} = -0.7071 - 1.2247\mathrm{i},$$
$$S_{23} = 0.7071 + 1.2247\mathrm{i}, \quad S_{24} = 2.1213 + 3.6742\mathrm{i}.$$

When $\theta_2 = \frac{\pi}{6}$, we have:

$$S_{31} = -1.0981 - 4.0981\mathrm{i}, \quad S_{32} = -0.3660 - 1.3660\mathrm{i},$$
$$S_{33} = 0.3660 + 1.3660\mathrm{i}, \quad S_{34} = 1.0981 + 4.0981\mathrm{i}.$$

5) The $N$-dimensional matrix **MC** with $M$ points can be expressed by rotating the $S_N$ vector as:

$$\mathbf{MC} = (S_1, S_2, \ldots, S_N)^{\mathrm{T}} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1N} \\ \vdots & \ddots & \vdots & \vdots \\ S_{N1} & \cdots & \cdots & S_{MN} \end{bmatrix}.$$

6) To enhance the performance of the codewords in the presence of fading, the vector components must be interleaved and rearranged as follows [5], [7]:

After the $S$ vectors have been obtained, the interleaving of the even-numbered $S$ vectors is completed ($S_2$ is symbolized by $S_e$). For interleaving, only even dimensions (rows) of the **MC** are reordered. For example, after interleaving, the $S_e$ vector is $S_e^*$, where $e$ is an even number in $N$ and is:

$$S_e^* = \left\{-S_e, \frac{M}{2+1}, \ldots, -S_e, \frac{3M}{4}, S_e, \frac{3M}{4+1}, \ldots, -S_e, \right. \quad (4)$$
$$\left. M, S_e, M, \ldots, -S_e, \frac{3M}{4}, S_e, \frac{3M}{4}, \ldots, S_e, \frac{M}{2+1}\right\},$$

or in an easier form:

$$S_e^* = \left\{-S_e \mathrm{Rem}\left(\frac{\frac{M}{2}}{M+1}\right) + \cdots + S_e \mathrm{Rem}\left(\frac{M}{M+1}\right) \right. \quad (5)$$
$$\left. -S_e \mathrm{Rem}\left(\frac{4M}{3\,(M+1)}\right) + \cdots + S_e \mathrm{Rem}\left(\frac{2M}{M+1}\right)\right\}.$$

For $M = 4$, the interleaved vector becomes:

$$S_2^* = [-0.707 - 1.225\mathrm{i} \quad 2.121 + 3.674\mathrm{i}$$
$$-2.121 - 3.6742\mathrm{i} \quad 0.707 + 1.225\mathrm{i}],$$

In main constellation, we place $S_1$, $S_2^*$ and $S_3$:

$$\mathbf{MC} = \begin{bmatrix} -3.0000 - 3.0000\mathrm{i} & -1.0000 - 1.0000\mathrm{i} \\ -0.7071 - 1.2247\mathrm{i} & -1.0981 - 4.0981\mathrm{i} \\ -2.1213 - 3.6742\mathrm{i} & 0.3660 + 1.3660\mathrm{i} \\ 1.0000 + 1.0000\mathrm{i} & 3.0000 + 3.0000\mathrm{i} \\ 2.1213 + 3.6742\mathrm{i} & -0.3660 - 1.3660\mathrm{i} \\ 0.7071 + 1.2247\mathrm{i} & 1.0981 + 4.0981\mathrm{i} \end{bmatrix}.$$

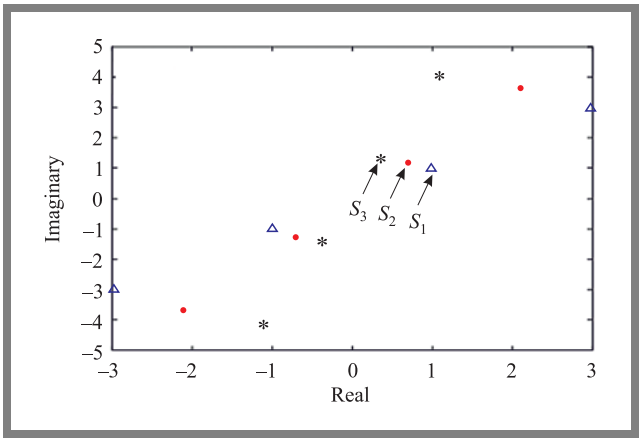



**Fig. 2.** Main constellation and rotation.

### Codebook design

After the complex MC has been obtained, its components are added to the sparse code words. The codewords for the various symbols transmitted by user $j$ are concatenated to create codebook $\mathbf{X}_j$ for that single user, where each symbol is represented by a column vector with $K$ rows. Two of the rows have complex numbers in them, and the value of the other two equals 0. Consequently, the SCMA codebook $\mathbf{X}_j$, often known as for $j$-th user, is generated as follows [5], [7]:

$$\mathbf{X}_j = V_j \mathbf{\Delta}_j \mathbf{MC} \qquad j = 1, 2, \ldots, J. \quad (6)$$

A sparsely populated $V_j$ represents the dispersion matrix for each user, where $V \in B^{K \times N}$, mapping the $K$-dimensional codeword $\mathbf{X}$ onto the $N$-dimensional complex constellation point. In this state, if $K = 6$, $N = 3$ and $J = 8$, matrices $V_j$, $j = 1, 2, \ldots, J$ are:

$$\mathbf{V}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{V}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{V}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{V}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{V}_5 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{V}_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{V}_7 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{V}_8 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Rotation of the corresponding $\mathbf{MC}$ is required by the codebook design for different users. The phase rotation angles [5], [7] are:

$$\varphi_u = (u-1)\frac{2\pi}{Mdf} + e_u \frac{2\pi}{M} \quad u = 1, \ldots, df, \tag{7}$$

where $e_u$ is an integer representation of $Z$ and the overloading factor or the number of elements present simultaneously in a given subcarrier is represented by $df$. The user's dimension on the interfering layers must remain at least as far apart from other layers' dimensions, thus $\varphi_u$ is the minimum ideal rotation phase angle.

## 4. Proposed Method

A new even dimension interleaving method using Arnold's cat map is proposed here. The Arnold's cat map is a square matrix expressed as [16]:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \bmod N. \tag{8}$$

where $N$ is the maximum index of the dimensional matrix (i.e. $N \times N$ elements matrix) to be interleaved using Arnold's cat map. The coordinates of each element which the ordered pair represents $(X, Y)$ are in the range of $[0 \ldots 1]$ when modulo 2 is applied. Such a transformation is mainly used for interleaving pixels in image processing applications. It creates a discrete-time dynamic system in which the $\mathbf{\Gamma}_{\text{cat}}$ mapping iterations control the evolution as:

$$\mathbf{\Gamma}_{\text{cat}} = \begin{bmatrix} X_{n+1} \\ Y_{n+1} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix} \bmod N, \tag{9}$$

where $N = \sqrt{M}$.

To improve performance of SCMA and reduce computational complexity, we propose using Arnold's cat map for interleaving the mother constellation using Eq. (9) instead of the interleaving process expressed by Eq. (5). To perform the interleaving of vector $S_2$ using Arnold's cat map, $\mathbf{S}_2$ should be converted into a $N \times N$ square matrix and processed. When $N$ is 2, $S_2$ would be a $2 \times 2$ matrix as:

$$\mathbf{S}_2 = \begin{bmatrix} -2.1213 - 3.6742\text{i} & 0.7071 + 1.2247\text{i} \\ -0.7071 - 1.2247\text{i} & 2.1213 + 3.6742\text{i} \end{bmatrix}.$$

Then, the locations of each element in the $\mathbf{S}_2$ matrix are changed using Arnold's cat map, as given in Eq. (9), where $\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$ is the location of elements before interleaving and $\begin{bmatrix} X_{n+1} \\ Y_{n+1} \end{bmatrix}$ is the corresponding location after interleaving. Hence, $\mathbf{S}_2$ after interleaving would be:

$$\mathbf{S}_2 = [2.1213 + 3.6742\text{i} \quad -2.1213 - 3.6742\text{i} \quad 0.7071 + 1.2247\text{i} \quad -0.7071 - 1.2247\text{i}].$$

$\mathbf{MC}$ after the proposed interleaving becomes:

$$\mathbf{MC} = \begin{bmatrix} -3.000 - 3.000\text{i} & 1.000 - 1.000\text{i} \\ 2.121 + 3.674\text{i} & -1.098 - 4.098\text{i} \\ 0.707 + 1.224\text{i} & 0.366 + 1.366\text{i} \\ 1.000 + 1.000\text{i} & 3.000 + 3.000\text{i} \\ -2.121 - 3.674\text{i} & -0.366 - 1.366\text{i} \\ -0.707 - 1.225\text{i} & 1.098 + 4.098\text{i} \end{bmatrix}.$$

The main constellation and rotation using the proposed chaotic interleaving can be explained using Algorithm 1.

**Algorithm 1.** Mother constellation creation.

**Input:** $k, j, df, \mathrm{dv}, M, N$. $j$ is the number of users, $k$ is the number of resources or subcarriers, $df$ is the number of users on one resource, dv is the number of resources for one user.
**Result:** creation of main constellation $\mathbf{MC}$
**Begin**
**1: For** $m = 1$ to $M$
**2:** $Q_m = 2m - 1 - M : S_1 = Q_m(1 + \mathrm{i})$
**3: End for** $M$
**4: For** $r = 1$ to $N$
**5:** $\theta_{r-1} = (r-1) \cdot \frac{\pi}{MN} : \mathrm{UN} = \mathrm{diag}(1\mathrm{e}^{1\theta r-1}) :$
$S_{1r} = Un(r, 1) \cdot S_1$
**6:** Select even elements of $S_{1r}$ only
**7:** Perform interleaving using Arnold's cat map
**8: End for** $N$
**9:** Construct main constellation: $\mathbf{MC} = [S_1, S_{2'}, \ldots, S_N]$
**10: End**

The values calculated from Eq. (7) if $M = 4$ and $df = 4$ are: $\varphi_1 = 0$, $\varphi_2 = \frac{\pi}{8}$ and $\varphi_4 = \frac{3\pi}{8}$. These are the ideal parameters that maintain the constant Euclidean distance between the codewords and are based on Latin squares and the structure of the $\varphi_1$ mother codebook [5], [7]. The rotation angles factor graph that follows this principle is:

$$\mathbf{F}_\varphi = \begin{bmatrix} \varphi_1 & \varphi_2 & \varphi_3 & 0 & 0 & 0 & \varphi_4 & 0 \\ \varphi_2 & 0 & 0 & \varphi_3 & \varphi_4 & 0 & \varphi_1 & 0 \\ 0 & \varphi_1 & 0 & \varphi_2 & 0 & \varphi_3 & 0 & \varphi_4 \\ \varphi_3 & 0 & \varphi_4 & 0 & \varphi_1 & 0 & 0 & \varphi_2 \\ 0 & \varphi_4 & 0 & \varphi_1 & 0 & \varphi_2 & \varphi_3 & 0 \\ 0 & 0 & \varphi_2 & 0 & \varphi_3 & \varphi_4 & 0 & \varphi_1 \end{bmatrix}. \tag{10}$$

As long as the minimal dimensional distance between the layers is maintained, there are several alternatives for how to arrange the phase rotation, depending on the codebook

structure. The angles are formulated in matrix $\mathbf{F}_{\varphi j}$ as:

$$\mathbf{F}_{\varphi j} = \begin{bmatrix} \varphi_1 & \varphi_2 & \varphi_3 & \varphi_3 & \varphi_4 & \varphi_3 & \varphi_4 & \varphi_4 \\ \varphi_2 & \varphi_1 & \varphi_4 & \varphi_2 & \varphi_1 & \varphi_2 & \varphi_1 & \varphi_2 \\ \varphi_3 & \varphi_4 & \varphi_2 & \varphi_1 & \varphi_3 & \varphi_4 & \varphi_3 & \varphi_1 \end{bmatrix}, \quad (11)$$

From the $\mathbf{F}_{\varphi j}$, the rotation operator $\mathbf{\Delta}_j$ is produced as follows:

$$\mathbf{\Delta}_j = \mathrm{diag}(F_{\varphi j}), \quad \forall j = 1, 2, \ldots, J \quad (12)$$

In this case, when $J = 8$, we have:

$$\mathbf{\Delta}_1 = \begin{bmatrix} \varphi_1 & 0 & 0 \\ 0 & \varphi_2 & 0 \\ 0 & 0 & \varphi_3 \end{bmatrix} \quad \mathbf{\Delta}_2 = \begin{bmatrix} \varphi_2 & 0 & 0 \\ 0 & \varphi_1 & 0 \\ 0 & 0 & \varphi_4 \end{bmatrix}$$

$$\mathbf{\Delta}_3 = \begin{bmatrix} \varphi_3 & 0 & 0 \\ 0 & \varphi_4 & 0 \\ 0 & 0 & \varphi_2 \end{bmatrix} \quad \mathbf{\Delta}_4 = \begin{bmatrix} \varphi_3 & 0 & 0 \\ 0 & \varphi_2 & 0 \\ 0 & 0 & \varphi_1 \end{bmatrix},$$

$$\mathbf{\Delta}_5 = \begin{bmatrix} \varphi_4 & 0 & 0 \\ 0 & \varphi_1 & 0 \\ 0 & 0 & \varphi_3 \end{bmatrix} \quad \mathbf{\Delta}_6 = \begin{bmatrix} \varphi_3 & 0 & 0 \\ 0 & \varphi_2 & 0 \\ 0 & 0 & \varphi_4 \end{bmatrix}$$

$$\mathbf{\Delta}_7 = \begin{bmatrix} \varphi_4 & 0 & 0 \\ 0 & \varphi_1 & 0 \\ 0 & 0 & \varphi_3 \end{bmatrix} \quad \mathbf{\Delta}_8 = \begin{bmatrix} \varphi_4 & 0 & 0 \\ 0 & \varphi_2 & 0 \\ 0 & 0 & \varphi_1 \end{bmatrix}.$$

SCMA codebooks $\mathbf{X}_j$ for the $j$-th user using Eq. (6) and after normalizing the values are:

$$\mathbf{X}_1 = \begin{bmatrix} 0.322 - 0.322\mathrm{i} & -0.108 - 0.108\mathrm{i} \\ 0.108 + 0.108\mathrm{i} & 0.322 + 0.322\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} -0.113 - 0.195\mathrm{i} & 0.338 + 0.585\mathrm{i} \\ -0.338 - 0.585\mathrm{i} & 0.113 + 0.195\mathrm{i} \\ -0.259 - 0.966\mathrm{i} & -0.086 - 0.322\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.086 + 0.322\mathrm{i} & 0.259 + 0.966\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix},$$

$$\mathbf{X}_2 = \begin{bmatrix} -0.322 - 0.322\mathrm{i} & -0.108 - 0.108\mathrm{i} \\ 0.106 + 0.108\mathrm{i} & 0.322 + 0.322\mathrm{i} \\ -0.051 - 0.089\mathrm{i} & 0.154 + 0.267\mathrm{i} \\ -0.259 - 0.966\mathrm{i} & -0.086 - 0.322\mathrm{i} \\ -0.154 - 0.267\mathrm{i} & 0.051 + 0.089\mathrm{i} \\ 0.086 + 0.322\mathrm{i} & 0.259 + 0.966\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix},$$

$$\mathbf{X}_3 = \begin{bmatrix} -0.478 - 0.478\mathrm{i} & -0.159 - 0.159\mathrm{i} \\ 0.159 + 0.159\mathrm{i} & 0.478 + 0.478\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ -0.167 - 0.289\mathrm{i} & 0.500 + 0.866\mathrm{i} \\ -0.054 - 0.201\mathrm{i} & -0.018 - 0.067\mathrm{i} \\ -0.500 - 0.866\mathrm{i} & 0.167 + 0.289\mathrm{i} \\ 0.018 + 0.067\mathrm{i} & 0.054 + 0.201\mathrm{i} \end{bmatrix},$$

$$\mathbf{X}_4 = \begin{bmatrix} 0.000 + 0.000\mathrm{i} & -0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ -0.113 - 0.195\mathrm{i} & 0.338 + 0.585\mathrm{i} \\ -0.054 - 0.201\mathrm{i} & -0.018 - 0.067\mathrm{i} \\ -0.338 - 0.585\mathrm{i} & 0.113 + 0.195\mathrm{i} \\ 0.018 + 0.067\mathrm{i} & 0.054 + 0.201\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} -0.707 - 0.707\mathrm{i} & -0.236 - 0.236\mathrm{i} \\ 0.236 + 0.236\mathrm{i} & 0.707 + 0.707\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix},$$

$$\mathbf{X}_5 = \begin{bmatrix} 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \\ 0.000 + 0.000\mathrm{i} & 0.000 + 0.000\mathrm{i} \end{bmatrix}$$

$$\begin{bmatrix} -0.707 - 0.707\mathrm{i} & -0.236 - 0.236\mathrm{i} \\ 0.236 + 0.236\mathrm{i} & 0.707 + 0.707\mathrm{i} \\ -0.051 - 0.089\mathrm{i} & 0.154 + 0.267\mathrm{i} \\ -0.118 - 0.440\mathrm{i} & -0.039 - 0.147\mathrm{i} \\ -0.154 - 0.267\mathrm{i} & 0.051 + 0.089\mathrm{i} \\ 0.039 + 0.147\mathrm{i} & 0.118 + 0.440\mathrm{i} \end{bmatrix},$$

$$\mathbf{X}_6 = \begin{bmatrix} 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ -0.478-0.479\text{i} & -0.159-0.159\text{i} \\ 0.104+0.180\text{i} & -0.104-0.180\text{i} \\ 0.159+0.159\text{i} & 0.476+0.478\text{i} \\ 0.035+0.060\text{i} & -0.035-0.060\text{i} \end{bmatrix}$$

$$\begin{bmatrix} 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ -0.258-0.966\text{i} & -0.086-0.322\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.086+0.322\text{i} & 0.259+0.966\text{i} \end{bmatrix},$$

$$\mathbf{X}_7 = \begin{bmatrix} -0.147-0.147\text{i} & -0.049-0.059\text{i} \\ 0.049+0.049\text{i} & 0.147+0.157\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \end{bmatrix}$$

$$\begin{bmatrix} 0.047+0.082\text{i} & -0.047-0.082\text{i} \\ 0.016+0.027\text{i} & -0.016-0.027\text{i} \\ -0.259-0.966\text{i} & -0.086-0.322\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.086+0.322\text{i} & 0.259+0.966\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \end{bmatrix},$$

$$\mathbf{X}_8 = \begin{bmatrix} 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ -0.707-0.707\text{i} & -0.236-0.236\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.236+0.236\text{i} & 0.707+0.707\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \end{bmatrix}$$

$$\begin{bmatrix} 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.000+0.000\text{i} & 0.000+0.000\text{i} \\ 0.104+0.180\text{i} & -0.104-0.180\text{i} \\ 0.025-0.092\text{i} & -0.008-0.031\text{i} \\ 0.035+0.060\text{i} & -0.035-0.060\text{i} \\ 0.008+0.031\text{i} & 0.025+0.092\text{i} \end{bmatrix},$$

The codebook for each user from the main constellation can be calculated using Algorithm 2, and the codebook of the first user is shown in Fig. 3.

**Algorithm 2.** Codebook design from main constellation

**Input:** Set $k, j, df, \mathrm{dv}, M, N$

**Result:** Creation of codebook for each user from main constellation $\mathbf{MC}$

**Begin**
**1: For** $j = 1$ to $J$
**2:** Compute the dispersion matrix for each user $V_j$
**3: For** $u = 1$ to $df$
**4:** Compute the phase rotation angle $\varphi_u = (u-1)\frac{2\pi}{M\,df} + eu\frac{2\pi}{M}$.
**5:** Using factor graph of rotation angles matrix $F_\varphi$ in Eq. (10), obtain rotation operator $\mathbf{\Delta}_j = \mathrm{diag}(F_{\varphi j})$
**6:** Construct the codebook of each user $\mathbf{X}_j = V_j \mathbf{\Delta}_j \mathbf{MC}$
**7: End for** $u$
**8: End for** $j$
**9: End**

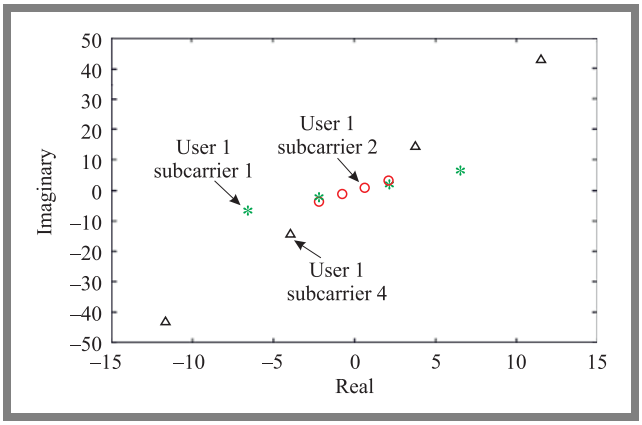



**Fig. 3.** Codebook of the first user.

## 5. Simulation Results

Matlab was used to evaluate the effectiveness of the designed chaotic interleaving-based SCMA codebooks. Performance of the designed SCMA system was measured in terms of MED, PAPR, computational complexity, and BER over a noisy channel. The minimum square Euclidean distance of the codeword from the mother codebook was computed using the following formula [4]:

$$d^2_{\min} = \begin{cases} 8[N_0^{\sim} + 4N_e^{\sim}] & M = 4 \\ 8N & M > 4 \end{cases}, \tag{13}$$

where $N_0^{\sim}$ and $N_e^{\sim}$ are the numbers of odd and even dimensions. respectively. On the other hand, PAPR was computed using [8]:

$$\mathrm{PAPR}(\mathbf{MC}) = 10\log\left(\frac{3[No(M-1)^2 + Ne]}{N(M^2-1)}\right). \tag{14}$$

Comparisons with various methods from the literature were conducted to assess the effectiveness of the proposed chaotic interleaving-based SCMA codebook. [5], [7], [9], and [11] have been selected for comparison – see Tab. 1. Fig. 4 shows a comparison of BER for different codebook design methods, while Fig. 5 illustrates the comparison of BER for different interleaving methods. It can be seen from these figures that MED is maximized when $M = 4$ and the proposed method achieves almost the same performance as in [5], [7] in terms of MED, PAPR, and obtainable SNR at $\mathrm{BER} = 10^{-4}$. The

**Tab. 1.** Comparison of different SCMA methods.

| Method | $J$ | $K$ | $M$ | $N$ | $df$ | SNR at BER $= 10^{-4}$ | PAPR | MED |
|---|---|---|---|---|---|---|---|---|
| MDconst [5], [7] | 6 | 4 | 4 | 2 | 3 | 21 dB | 0 | 2 |
| Proposed | 6 | 4 | 4 | 2 | 3 | 22 dB | 0 | 2 |
| MDconst [5], [7] | 8 | 6 | 4 | 3 | 4 | 23 dB | 1.96 | 1.64 |
| Proposed | 8 | 6 | 4 | 3 | 4 | 24 dB | 1.96 | 1.64 |
| MDconst [5], [7] | 6 | 4 | 16 | 2 | 3 | 27 dB | 0 | 0.88 |
| Proposed | 6 | 4 | 16 | 2 | 3 | 28 dB | 0 | 0.88 |
| 16-star QAM [9] | 6 | 4 | 4 | 2 | 3 | 10 dB | 0 | 2.11 |
| CG SCMA [11] | 6 | 4 | 4 | 2 | 3 | 8-9 dB | 0 | 2.16 |

results also show that the 16-star QAM constellation, as well as the segmentation and CG-SCMA methods achieve better BER than other methods due to the fact that coding associated with modulation is used.

Computational complexity is an important performance indicator that should be taken into account while designing codebooks. Low memory usage, as well as high speed of encoding and decoding activities guarantee good performance metrics. The described interleaving method is simpler in terms of computational cost, as it needs fewer mathematical operations.

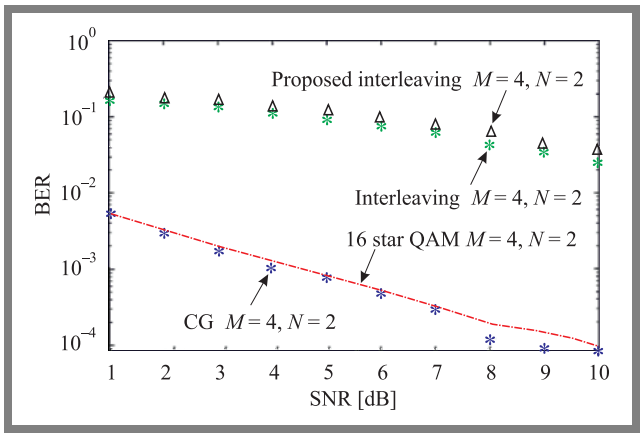



**Fig. 4.** Comparison of BER of different SCMA methods.

From Eq. (5), one may notice that 11 operations, such as division and multiplication, are required to perform interleaving for $M = 4$, while 23 operations are required for $M = 8$. Therefore, the number of operations required in the interleaving method presented in [5], [7] is $3 \times M - 1$. On the other hand, in the proposed interleaving method, according to Eq. (9), the required number of operations is 8 for $M = 4$, and 16 for $M = 8$ ($2 \times M$). The percentage reduction can be expressed as:

$$\text{Reduction} = \frac{M-1}{3M-1} \cdot 100[\%]. \tag{15}$$

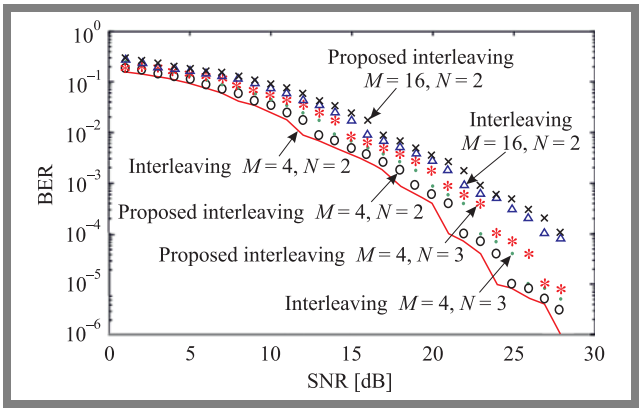



**Fig. 5.** BER of interleaving methods.

**Tab. 2.** Comparison of different SCMA interleaving methods in terms of complexity.

| Number of codewords $M$ | Number of multiplications and divisions | | Percentage reduction |
|---|---|---|---|
| | Interleaving method [5], [7] | Proposed chaotic interleaving | |
| 4 | 11 | 8 | 27% |
| 8 | 23 | 16 | 30% |
| 16 | 47 | 32 | 32% |

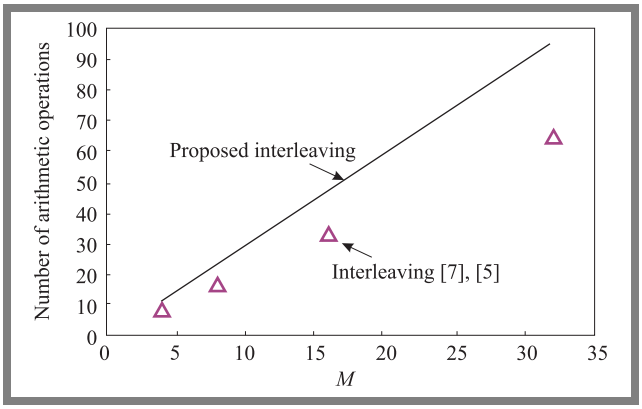



**Fig. 6.** Complexity of the proposed interleaving scheme compared with papers [5] and [7].

## 6. Conclusions

The paper presents an efficient SCMA codebook design based on interleaving. It covers several codebook design techniques that have been documented in the literature. The proposed method reduces computational complexity and offers, by relying on the Arnold's cat map for interleaving, better performance than other existing interleaving approaches. Future research should concentrate on improving MD constellations by including sophisticated models or machine learning techniques, such as neural networks.

## References


[1] Z. Ding, Z. Yang, P. Fan, and H.V. Poor, "On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users", *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, 2014 (DOI: 10.1109/LSP.2014.2343971).

[2] Z. Liu and L-L. Yang, "Sparse or dense: a comparative study of code domain NOMA systems", *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4768–4780, 2021 (DOI: 10.1109/TWC.2021.3062235).

[3] E. Catak, F. Tekce, O. Dizdar, and L. Durak-Ata, "Multi-user shared access in massive machine-type communication systems via superimposed waveforms", *Physical Communication* 37 (2019), 100896, 2019 (DOI: 10.1016/j.phycom.2019.100896).

[4] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends", *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015, (DOI: 10.1109/MCOM.2015.7263349).

[5] D. Cai, P. Fan, X. Lei, Y. Liu, and D. Chen, "Multi-dimensional SCMA codebook design based on constellation rotation and interleaving", *83-rd IEEE Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2016 (DOI: 10.1109/VTCSpring.2016.7504356).

[6] S.A. Hussain, *et al.*, "A review of codebook design methods for sparse code multiple access", *Indonesian Journal of Electrical Engineering and Computer Science*, pp. 927–935, 2021 (DOI: 10.11591/ijeecs.v22.i2.pp927-935).

[7] Y.M.J. Licea, "Resource allocation for uplink code-domain non-orthogonal multiple access", *Ph.D. thesis, Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, Manchester University*, 2021 (URL: https://www.research.manchester.ac.uk/portal/files/205626604/FULL_TEXT.PDF).

[8] J.L.L. Bonilla, S.V. Beltrán, I.S. Rivera, and F.M. Pinón, "Construction of SCMA CodeBooks using the phase rotation method", *IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, pp. 1–8, 2018 (DOI: 10.1109/ROPEC.2018.8661414).

[9] 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 (DOI: 10.1109/ACCESS.2018.2876030).

[10] L. Yu, P. Fan, D. Cai, and Z. Ma, "Design and analysis of SCMA codebook based on star-QAM signaling constellations", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10543–10553, 2018 (DOI: 10.1109/TVT.2018.2865920).

[11] Y.M. Tabra and B.M. Sabbar, "New Computer Generated-SCMA Codebook With Maximised Euclidian Distance for 5G", *Iraqi Journal of Information & Communications Technology*, vol. 2, no. 2, pp. 9–24, 2019 (DOI: 10.31987/ijict.2.2.64).

[12] Z. Mheich, L. Wen, P. Xiao, and A. Maaref, "Design of SCMA codebooks based on golden angle modulation", *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1501–1509, 2019 (DOI: 10.1109/TVT.2018.2886953).

[13] 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 (DOI: DOI: 10.1109/TCOMM.2020.2974213).

[14] F. d. Silva, D. Le Ruyet, and B.F. Uchoa-Filho, "Threshold-Based Edge Selection MPA for SCMA", *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2957–2966, 2020 (DOI: 10.1109/TVT.2020.2966333).

[15] M. Vameghestahbanati, "Designing Multidimensional Constellations and Efficient Detection Schemes for Sparse Code Multiple Access (SCMA) Systems", *Ph.D. dissertation*, Carleton University, (URL: https://curve.carleton.ca/f481d0f2-ec66-44b4-a783-1c602d2c210a).

[16] G. Peterson, "Arnold's Cat Map", *Math 45 – Linear Algebra Fall*, 1997 (DOI: 10.3840/002296).



**Sura S. Mohamed** obtained her B.Sc. in Information and Communication Engineering Department in College of Information Engineering in 2005 at AL-Nahrain University, Iraq. She worked as Information and Communication Engineer in private communication company. Since 2019 she is a M.Sc. student at Information and Communication Engineering Department in College of Information Engineering at AL-Nahrain University, Iraq. She is interested in the wireless communication systems and information theory and coding.
E-mail: suhaibsura@gmail.com
Al-Nahrain University, College of Information Engineering



**Hikmat N. Abdullah** obtained his B.Sc. in Electrical Engineering in 1995, M.Sc. in Communication Engineering in 1998 at University of Al Mustansiryah, Iraq and Ph.D. in Communication Engineering in 2004 at University of Technology, Iraq. From 1998 to 2015 he worked as a lecturer in the Electrical Engineering Department, at Al-Mustansiryah University, Iraq. Since 2015 he works as full professor in College of Information Engineering at Al-Nahrain University, Iraq. He is a senior member of IEEE since 2014. He is interested in subject wireless communication systems and cognitive radio networks.
E-mail: hikmat.abdullah@coie-nahrain.edu.iq
Al-Nahrain University, College of Information Engineering

Paper

# Deep Learning-based SNR Estimation for Multistage Spectrum Sensing in Cognitive Radio Networks

Sanjeevkumar Jeevangi, Shivkumar Jawaligi, and Vilaskumar Patil

*Department of Electronics and Communication Engineering, Faculty of Engineering and Technology (Co-Ed), Sharnbasva University, India*

https://doi.org/10.26636/jtit.2022.164922

**Abstract — Vacant frequency bands are used in cognitive radio (CR) by incorporating the spectrum sensing (SS) technique. Spectrum sharing plays a central role in ensuring the effectiveness of CR applications. Therefore, a new multi-stage detector for robust signal and spectrum sensing applications is introduced here. Initially, the sampled signal is subjected to SNR estimation by using a convolutional neural network (CNN). Next, the detection strategy is selected in accordance with the predicted SNR levels of the received signal. Energy detector (ED) and singular value-based detector (SVD) are the solutions utilized in the event of high SNR, whilst refined non-negative matrix factorization (MNMF) is employed in the case of low SNR. CNN weights are chosen via the Levy updated sea lion optimization (LU-SLNO) algorithm inspired by the traditional sea lion optimization (SLNO) approach. Finally, the outcomes of the selected detectors are added, offering a precise decision on spectrum tenancy and existence of the signal.**

***Keywords — cognitive radio, improved NMF, LU-SLNO system, optimized CNN, spectrum sensing***

**Tab. 1.** List of acronyms.

| Abbreviation | Description |
|---|---|
| ACD | Auto-correlation based detector |
| AWGN | Additive white Gaussian noise |
| AOA | Arithmetic optimization algorithm |
| BES | Bald eagle search |
| CDF | Cumulative distribution function |
| CFD | Cyclostationary feature-based detector |
| CR | Cognitive radio |
| CNN | Convolutional neural network |
| CRN | Cognitive radio networks |
| DWT | Discrete wavelet transform |
| ED | Energy detector |
| FAR | False alarm probability |
| IoT | Internet of Things |
| HGS | Hunger games search |
| LA | Lion algorithm |
| LU-SLNO | Levy updated SLNO |
| MME | Maximum-minimum eigen value detectors |
| MAF | Moving average filtering |
| MSWF | Multistage Wiener filter |
| MNMF | Modified non-negative matrix factorization |
| MSS | Multi-band spectrum-sensing |
| NBSS | Numerous narrow-band spectrum sensing |
| PRO | Poor rich optimization |
| PUs | Primary user |
| PSD | Power spectral density |
| ROC | Receiver operating curve |
| RF | Radio frequency |
| SUs | Secondary user |
| SS | Spectrum sensing |
| SVD | Singular value based detector |
| SNR | Signal to noise ratio |
| SLNO | Sea lion optimization |
| TSA | Taylor series approximation |
| WF | Wiener filter |
| WBSS | Wide-band spectrum sensing |

## 1. Introduction

The scarcity of radio spectrum in bands below 6 GHz is turning into a severe issue in current wireless communication technologies. Legal restrictions and static allotment of channels are the major issues resulting in the lack of viable radio spectrum [1], [2]. According to [3]–[5], under-utilization of the available spectrum is the root cause of spectrum scarcity [3]–[5]. By enabling unlicensed (secondary) users to access legacy networks, the cognitive radio (CR) technology promises to solve this spectrum shortage problem in situations in which the spectrum is underutilized by its licensed (primary) users [6]–[8].

The most important task in the CR phase is spectrum sensing (SS), or the precise locating of spectral holes (white spaces).

NBSS techniques, such as cyclo-stationary feature identification, energy identification, and matched filtering [9]–[12] have been the focus of recent research. Unfortunately, NBSS algorithms concentrate on taking advantage of spectral opportunities over a narrow frequency range. Cognitive radio networks (CRNs) need to take advantage of spectral opportunities over a wide frequency range, from hundreds of MHz to tens of GHz, in order to achieve the expected high throughput [8], [13], [14].

Another technique, known as WBSS, has received a lot of interest recently as well, and much research is currently conducted in this specific area [15], [16]. In general, it involves determining the frequency location of each sub-band of the sensed RF spectrum, regardless of the form of PSD, starting with wavelet decomposition for detecting matching sub-bands accessible for the SU [8], [17]. This paper aims to contribute in the following aspects:

– It presents the MSS-CRN system paradigm, in which CNN is used to measure SNR;
– After estimating SNR values, the detection scheme is chosen based on SNR at the receiver;
– SNR is categorized based on two threshold values. For high SNR, ED and SVD techniques are used, whereas for low SNR, MNMF is used;
– CNN weights are chosen in an optimum manner by using the LU-SLNO algorithm.

Table 1 summarizes the acronyms and abbreviations used in this paper.

## 2. Related Works

A multi-stage robust detector suitable for CR was reported by Kaliappan *et al.* [18]. For bands with lower and higher SNR, the selected model used two parallel connected detectors. ED and SVD were used for high SNR bands, while CFD and ACD were used for low SNR bands. According to empirical results, the suggested method beats previous models in terms of complexity.

Moayad *et al.* in [19] defined the spectrum accessing process as a multi-stage rate optimization/channel assignment. A novel resource-based channel assignment technique was developed allowing for the appropriate utilization of the available temporal frequency units. Additionally, this approach relied upon packet segmentation capabilities to significantly increase network throughput.

Jay *et al.* [20] introduced a three-part approach for analyzing localization (ROC)-based sensing performance of MSS: the best spectrum-sensing strategies maximized the area under ROC curves and could be used as a performance scale for another MSS method and to assess the energy costs of MSS. Haobo *et al.* in [21] introduced MSWF to improve effectiveness of the detection process. Additionally, based on total energy obtained from all routes, the specific locations were obtained. As such an approach did not call for the evaluation of noise variances, the selected model was resistant to noise uncertainty. The selected system has increased detection capabilities, making it suitable for CR appliances. Kim *et al.* [22] created a unique spectrum sensing and CR-sharing technique for IoT devices. A unique theory known as reciprocal fairness was developed to ensure spectrum sensing and CR sharing. The accepted model relied also on the game theory and turned out to be more effective than other examined models.

Edge identification and noise removal operations were carried out concurrently using DWT-oriented models developed by Abhishek *et al.* [8]. A MAF strategy was presented at various levels of DWT-oriented models in order to provide higher detection performance in low SNR scenarios. Low-scale DWT values were also used to obtain enhanced performance and shorter computational time. Compared to similar models, the suggested multi-task approach was simpler and more effective.

Rajput *et al.* [23] evaluated an adaptive covariance cut-off approach, comparing it against current techniques, such as energy detection and MME. Such parameters as accuracy and complexity of sensing ROC curves were taken into account there. Additionally, for every detector, the effects of the signal bandwidth in relation to the observation bandwidth were investigated. Vijayakumar and Malarvizhi [24] researched an SDR implementation of CR using an energy detector and a Wiener filter (WF) to eliminate self-interference. WF predicted the self-interfering signal which was then removed from the received signal before being moved to the energy detector. The outcomes show improvements in the probability of detection and improved efficiency.

Eappen and Geoffrey [25] introduced a hybrid PSO-GSA model to detect spectrum holes with improved energy utilization. The inclusion of mutation and crossover components in PSO-GSA enabled the recommended method to recognize spectral gaps efficiently, while maintaining optimum transmission capacity, sensor bandwidth, and spectral power density. Saber *et al.* [26] suggested a streamlined implementation of spectrum detection based on real signals produced by an Arduino Uno board and a 433 MHz ISM wireless transmitter. The received signal was detected using artificial neural networks, support vector machines, decision trees, and k-nearest neighbors, using five support vectors: linear, quadratic, cubic, medium Gaussian, and coarse Gaussian. The results showed that the use of machine learning, in particularly of SVM and ANN techniques, ensures better spectrum sensing capabilities.

Table 2 provides a summary of MSS for CRN evaluations, obtained from the literature. SVD with a high SNR level and a short sensing time suffers from synchronization problems [18]. The MAC protocol used in [19] improves throughput and spectrum efficiency, but only allows a limited basic rate on each channel. The binary search approach employed in [20], offers great detection performance, while consuming less energy. However, the setting threshold level needs optimization. The MSWF model researched in [21] achieves minimal complexity and good detection performance. How-

ever, perfect energy detection was not possible. The concept of reciprocal fairness from [22], which improved flexibility and lowered the throughput loss ratio, lacks a researched behavior model. Additionally, DWT used in [8] offers high SNR and increased FAR levels. However, shadowing effects remain unimproved. CDF, known from [23], offers high detection probability and high SNR, but has to be optimized in terms of its noise performance. In [24], a WF-based model was presented which offers improved accuracy and minimized overhead. However, certain functions need to be executed for a higher number of cycles. The hybrid PSO-GSA model presented in [25] to identify spectrum holes with better energy efficiency has good convergence and does not adhere to local optima, but its computational complexity grows with the population size. In [26], a streamlined implementation of spectrum detection based on authentic signals is proposed. The results demonstrate that using machine learning, particularly SVM and ANN techniques, will result in better spectrum sensing efficiency. Nevertheless, targeting categories overlap each other.

A new multi-stage detector is proposed for reliable signal and spectrum detecting devices. When deployed in an MSS-CRN, the process is designed as follows (Fig. 1):

– first, the sampled signal is subjected to SNR estimation,
– SNR is estimated using CNN,
– after evaluating SNR, SNR estimate of the message data is used to determine the tracking system,
– two levels of SNR are determined, MNMF is utilized with low SNR settings, while ED and SVD are used in high SNR scenarios,
– the LU-SLNO technique is used to select CNN weights in the best possible manner.

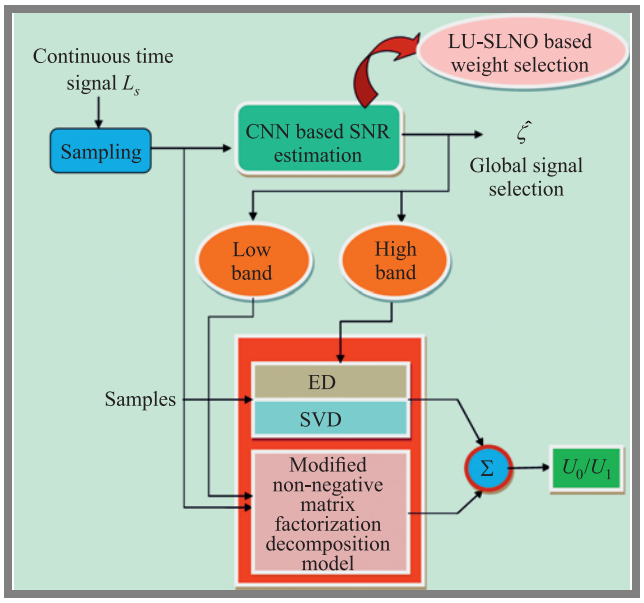



**Fig. 1.** MSS CRN system concept with LU-SLNO.

**Tab. 2.** Review of MSS in CRNs, based on existing papers.

| Paper | Schemes developed | Features | Issues |
|---|---|---|---|
| [18] | SVD | • High SNR<br>• Minimal sensing time | • Synchronization issues may occur |
| [19] | MAC protocol | • High spectrum efficiency<br>• Enhanced throughput | • Each channel supports a predetermined basic rate only |
| [20] | Binary search method | • High detection performance<br>• Minimized energy consumption | • Needs more consideration in terms of threshold selection |
| [21] | MSWF model | • Minimal complexity<br>• High detection performance | • Ideal detection of energy could not be achieved |
| [22] | Reciprocal fairness concept | • Better flexibility<br>• Minimized throughput loss ratio | • Need consideration in terms of behavior modeling |
| [8] | DWT | • Increased FAR<br>• High SNR | • Shadowing effects are not considered |
| [23] | CDF | • High detection probability<br>• Higher SNR | • Noise issues have to be improved |
| [24] | WF | • Avoids overhead<br>• High accuracy | • Certain functionalities require running more cycles |
| [25] | PSO-GSA model | • Does not stick to local optima<br>• Better convergence rate | • Increase in computational complexity when population size increases |
| [26] | ANN+SVM | • Better accuracy | • Overlap between target classes |

## 3. CNN-based SNR Estimation

A common representation of the signal recognition formula may by determined as:

$$z(t) = \begin{cases} m(t), & U_0 \\ s(t) + m(t) & U_1 \end{cases}. \tag{1}$$

where, $z(t)$ refers to the received signal, signal to identify is noted by $s(t)$, $m(t)$ refers to AWGN, while $U_0$ and $U_1$ refer to null and positive hypotheses, respectively.

Based on the determined SNR of the incoming signal, the multi-stage detector chooses a suitable detection method. AWGN with zero mean $\delta$ and unity variance $\sigma^2$ degrades the arriving time domain signal. The estimating technique is based on the signal's sampled noise component variance extraction [27], [28].

SNR estimation is performed as follows:

– Calculate the square of count of samples

$$x(m) = (z(m))^2 \quad m = 1, 2. \ldots, L_s, \tag{2}$$

– Address $x(m)$ samples as a lower-order polynomial for TSA.

– Estimate noise difference $\sigma_m^2$ from the foremost coefficients of the appropriate polynomial.

– Calculate signal power $\hat{Q}_s$ as:

$$\hat{Q}_s = \frac{1}{L_s} \sum_{m=1}^{L_s} (z(m))^2 - \sigma_m^2. \tag{3}$$

– SNR $\zeta$ is approximated as:

$$\hat{\zeta} = 20 \log \left( \frac{\hat{Q}_s}{\sigma_m^2} \right). \tag{4}$$

### *3.1. CNN Based Training*

The taxonomy of CNN accuracy is renowned to be excellent, particularly for visual sources. By convoluting the input image's features, a CNN can extract more relevant information and more detailed features, which should further increase the accuracy of SNR calculations. A CNN [29] comprises 3 layers: convolution, fully connected, and pooling layers with their features formulated as:

$$Z'_{r,t,w} = \mathrm{wet}_w^{l\ \mathrm{T}} PI_{r,t}^l = D_w^l, \tag{5}$$

where, $\mathrm{wet}_w^l$ is the weight, $D_w^l$ is the bias of $w$-th the filter linked to the $l$-th layer. At the central location $(r, t)$ of the $l$-th layer, patch input is indicated as $PI_{r,t}^l$. The activation value $\mathrm{act}_{r,t,w}^l$ related with convolutional features $z_{rtw}^l$ and is calculated as:

$$\mathrm{act}_{r,t,w}^l = \mathrm{act}\left(Z_{r,t,w}^l\right). \tag{6}$$

In the pooling layer, function $\mathrm{pool}(\bullet)$ is linked with $\mathrm{act}_{m,h,w}^l$, and $c_{r,t,w}^l$ is calculated as:

$$C_{r,t,w}^l = \mathrm{pool}\left(\mathrm{act}_{m,h,w}^l\right), \quad \forall (m.h) \in \mathrm{NN}_{r,t} \tag{7}$$

in which $\mathrm{NN}_{r,t}$ is neighbor $s(r, t)$. The CNN loss $CL$ is:

$$CL = \frac{1}{wn} \sum_{h=1}^{wm} l\left(\theta, C^{(h)}, F^{(h)}\right). \tag{8}$$

where the general constraint related with $\mathrm{wet}_w^l$ and $D_w^l$ is denoted by $\theta$. The output, the labels and the $h$-th input feature are denoted as $F^{(h)}$, $C^{(h)}$, and $PI^{(h)}$, respectively.

## 4. Proposed LU-SLNO Algorithm

The SLNO optimization method well-known from [30] illustrates sea lions' propensity to hunt. The algorithm is enhanced here to increase its effectiveness and convergence. Traditional optimization techniques have shown that self-improvement is beneficial [31]–[38]. Sea lions' delicate whiskers are a distinguishing feature that aids in pinpointing the precise location of their prey. Four stages of the process are defined: tracking, social hierarchy, attacking, and encircling prey to determine the properties of LU-SLNO.

The tracking method is defined as:

$$\mathrm{DIS} = \left| 2\vec{A}\vec{\Re}(t) - \vec{X}(t) \right|. \tag{9}$$

in which DIS denotes the distance between the sea lion and its prey, the vector position of the sea lion and the targeted prey is given by $\vec{X}(t)$ and $\vec{\Re}(t)$, while $t$ denotes the present iteration. Conventionally, identifies an arbitrary vector. In LU-SLNO, is evaluated based on a sinusoidal chaotic map.

In each cycle, the sea lion pays more attention to the victim. Equation (10) specifies the numerical model of this process:

$$\vec{X}(t+1) = \vec{\Re}(t) = \mathrm{DIS}\,\aleph. \tag{10}$$

where $(t+1)$ denotes the future iteration and $\aleph$ is continuously minimized across the iterations from 0 to 2.

The vocalization phase is illustrated in Eqs. (11)–(13) which describe finding the food and inviting other sea lions to participate in a joint attack. The speed of the leader's sound is denoted as $\overrightarrow{X_{\text{leader}}}$ and the speed of sound in the air and in water is designated by $\vec{P}_2$ and $\vec{P}_1$, respectively.

$$\overrightarrow{X_{\text{leader}}} = \left| \frac{\vec{P}_1(1 + \vec{P}_2)}{\vec{P}_2} \right|, \tag{11}$$

$$\vec{P}_1 = \sin\theta, \tag{12}$$

$$\vec{P}_2 = \sin\phi. \tag{13}$$

In the exploration phase, the SLnO model performs a universal search when $\aleph > 1$, defined as:

$$\mathrm{DIS} = \left| 2\vec{B}\vec{X}_{\mathrm{rnd}}(t) - \vec{X}(t) \right|. \tag{14}$$

$$\vec{X}(t+1) = \overrightarrow{X_{\mathrm{rnd}}}(t) - \mathrm{DIS}\,\aleph. \tag{15}$$

If $(\overrightarrow{\mathrm{SP}_{\text{leader}}} > 0.25)$, the attacking phase is conducted.

The attacking phase is employed to describe how a sea lion assaults the prey:

$$\vec{X}(t+1) = \left| \vec{\Re}(t) - \vec{X}(t).\cos(2\pi l) \right| + \vec{\Re}(t), \tag{16}$$

where the distance between the sea lion and its prey is identified by $\vec{\Re}(t) - \vec{X}(t)$, the absolute value is denoted by $||$ and the random number is referred by $l$.

In LU-SLNO, the sea lion's attack is based on the Levy method and is given as:

$$\vec{X}(t+1) = \left| \vec{\Re}(t) - \vec{X}(t).\cos(2\pi l) \right| + \vec{\Re}(t) \oplus \mathrm{levy}(\alpha) \tag{17}$$

where $t_{\max}$ implies the maximal iteration, while $ra_1$ and $ra_2$ imply arbitrary integers. In addition, Cauchy's mutation is performed for better convergence.

$$\mathrm{levy}(\alpha) = 2\left(1 - \frac{t}{t_{\max}}\right) \frac{ra_1 \cdot \beta}{|ra_2|^{\frac{1}{2}}}. \tag{18}$$

## 5. Multi-stage Detector

For applications requiring reliable signal and spectrum detection, i.e. for tactical adaptive communication networks, a multi-stage integrated detector is proposed. In this method,

two detectors (SVD and ED) are connected in parallel for servicing higher SNR regions, while one detector (MNMF) is used for lower SNR regions. The high SNR band covers the $0\ldots-10$ dB range, while the low band deals with SNR of $-10\ldots-20$ dB.

For the higher SNR band, probability $\mathrm{pr}=1$ when the approximated SNR $\zeta$ exceeds $-9$ dB. Total probability of detection $\mathrm{pd}$ and average sensing time $\mathrm{st}$ are expressed by:

$$\mathrm{pd} = \max(p^{\mathrm{ed}}, p^{\mathrm{svd}}) \tag{19}$$

$$\mathrm{st} = (s^{\mathrm{ed}}, \mathrm{st}^{\mathrm{svd}}) \quad \text{for} \quad \mathrm{pd} \geqslant 0.9. \tag{20}$$

For the lower SNR band, i.e. when $(G_2 > \zeta > G_3$ if $\zeta$ is below 9 dB then $\mathrm{pr}=0$ and $(1-\mathrm{pr})=1$. Total probability of detection $\mathrm{pd}$ and average sensing time $\mathrm{st}$ are:

$$\mathrm{pd} = \max(p^{\mathrm{MNMF}}) \tag{21}$$

$$\mathrm{st} = (\mathrm{st}^{\mathrm{MNMF}}) \quad \text{for} \quad \mathrm{pd} \geqslant 0.9. \tag{22}$$

For energy detection estimation, signal samples are doubled and then concatenated to obtain the message signal. Next, using the determined detection threshold $\mu_1$, noise energy is estimated as:

$$\mu_1 = \sqrt{2L_s}\sigma_m^2 P^{-1}(Q_{fa}) + L_s\sigma_m^2, \tag{23}$$

where $p^{-1}(\cdot)$ implies a converse $P$ function and $Q_{fa}$ implies FAR.

Finally, the decision is made on the basis of:

$$Z \rightarrow \begin{cases} \text{if} & \sum_{j=1}^{L_s} |z_i|^2 < \lambda_1, \quad U_0 \\ \text{else} & \qquad\qquad\quad U_1 \end{cases}, \tag{24}$$

where $z_j$ implies a $j$-th sample of $z(t)$.

### 5.1. SVD-Based Detection

The linear system signal processing approach and statistics greatly benefit from the SVD technique. It serves as an additional method for figuring out the matrix's eigenvalues [18], [39]–[42].

For SVD detection purposes, we start by choosing the columns count of matrix $V$, so that $k < V < L_s - k$, where $k$ implies the number of dominating singular values [41]. In most cases, when there are several samples $L_s$, $V \leqslant 20$. In the second step, it is combined into a Hankel matrix $HM$, as:

$$HM = q(x+1-1) \;\; n = 1, 2, \ldots, L \;\text{ and }\; l = 1, 2, \ldots, N. \tag{25}$$

After factorizing the matrix, we obtain singular values, i.e. $\lambda_{\min}$ and $\lambda_{\max}$ and then calculate the threshold value $\lambda_2$ as:

$$\lambda_2 = \frac{\left(\sqrt{L_s}+\sqrt{V}\right)^2}{L_s}\left(1+\frac{O\left(\sqrt{L_s}+\sqrt{V}\right)^{-\frac{2}{3}}}{(L_s\cdot V)^{\frac{1}{6}}}\right). \tag{26}$$

where $O$ implies the Tracy Widom's operation of $Q_{fa}$ [41], [42]:

In the next step, the proportion with threshold $\lambda_2$ is evaluated. If $\lambda_{\max} = \lambda_{\min} > \lambda_2$ , the signal is available and that proves the hypothesis. Otherwise, the signal is not present, pointing at $U_0$.

### 5.2. Modified NMF

NMF estimates the fundamental matrix $H(n \times l)$ and the coefficient matrix $W(k \times m)$ from the original matrix [43] $Y \approx HW$. A multi-variant linear equation serves as the foundation of NMF. Each column of $Y$ is roughly portrayed as a linear fusion of the vector form of $H$:

$$Y_i = \sum_{j=1}^{k} H_j W_{ji}. \tag{27}$$

where, $i = 1, 2, \ldots$ column count in $Y$.

As per the modified NMF, $W^*$ is computed by limiting the summary of values in each column of $W$ to one, as:

$$w_{ij}^* = \frac{W_{ij}}{\sum_{l=1}^{k} W_{ij}} \;\text{ or }\; \frac{\sum_{l=1}^{m} Wt_i}{\sum_{i=1}^{k} \frac{wt_i}{W_{ij}}}, \tag{28}$$

where $wt$ refers to weight.

### 5.3. Sensing Time

The effectiveness of the suggested detection method is evaluated using $\mathrm{st}$ and $\mathrm{pd}$ metrics. The proposed $\mathrm{st}$ detector is formulated as:

$$\mathrm{st} = \mathrm{pr}\left(\mathrm{st}^{\mathrm{ed}}, \mathrm{st}^{\mathrm{svd}}\right) + (1-\mathrm{pr})\cdot \mathrm{st}^{\mathrm{MNMF}}. \tag{29}$$

where $\mathrm{st}^{\mathrm{ed}}$, and $\mathrm{st}^{\mathrm{svd}}$ refer to the mean sensing time of ED, SVD, and MNMF detectors, respectively.

## 6. Simulation Results

In order to simulate the effectiveness of CNN + LU-SLNO, Matlab was used together with the DeepSig dataset and radio ML 2016.10A from [44]. The simulation parameters and specifications are presents in Tab. 3, while representations of the sample signal used are shown in Fig. 2.

**Tab. 3.** Simulation parameters.

| Parameter | Specification |
|---|---|
| Input signals | QPSK, BPSK, QAM64, PAM4 |
| Channel | AWGN |
| SNR range | 0 to $-20$ [dB] |
| High SNR region | 0 [dB] >= $x$ >= $-9$ [dB] |
| Low SNR region | $-9$ [dB] > $x$ >= $-20$ [dB] |
| Detector used | ED, SVD (low SNR band) NMF (high SNR band) |
| Probability of false alarm $P_{fa}$ | 0.1 |

The proposed CNN + LU-SLNO model was validated over ANN+SVM [26], CNN+PSO-GSA [25], CNN + HGS, CNN + AOA, CNN + PRO, CNN + BES, CNN + LA, CNN + SLNO, AWGN (CFD) [18] and AWGN (ACD) [18], using a synthetic dataset created with GNU Radio and consisting of 11 modulations at different signal-to-noise ratios (8 digital and 3 analogue).

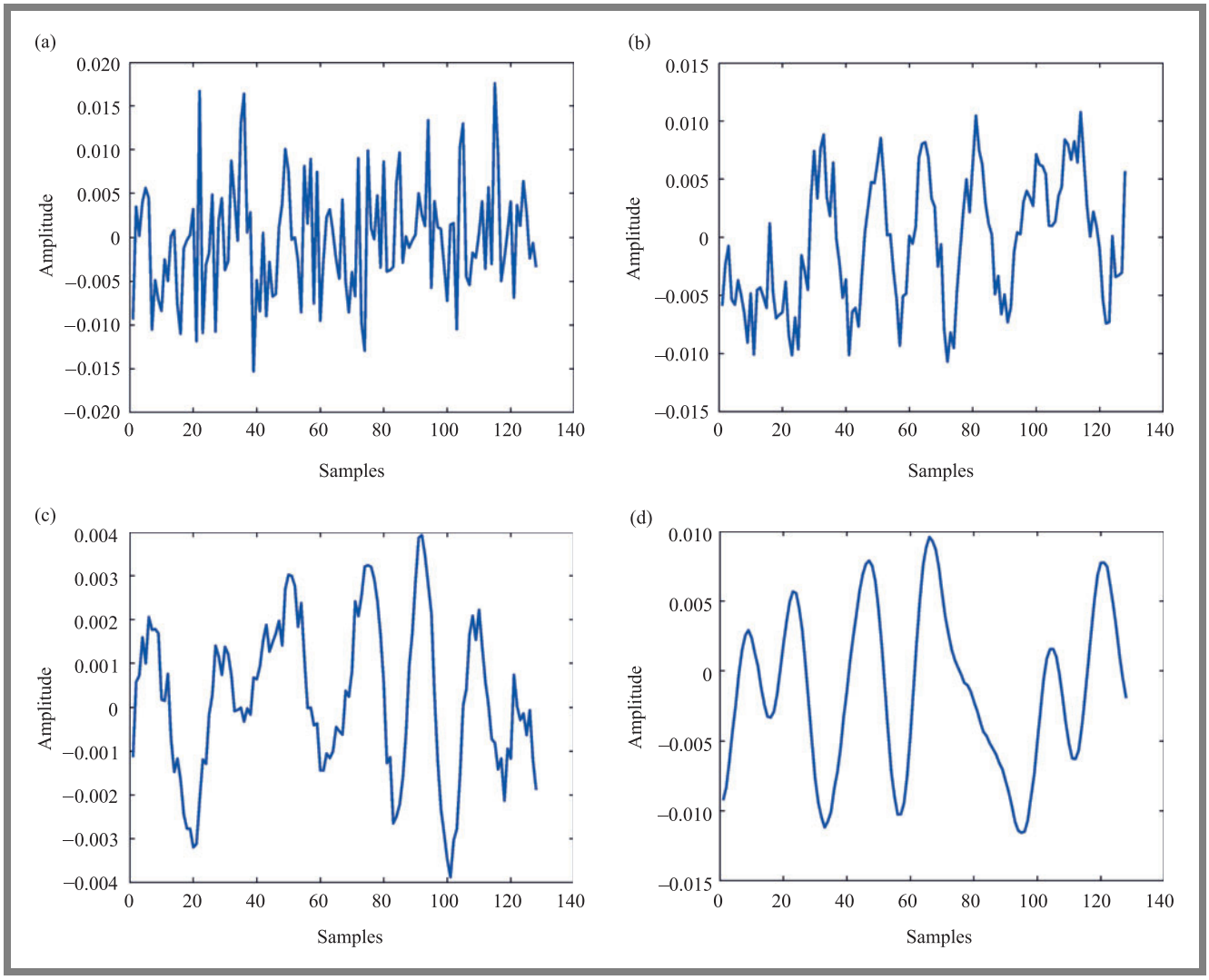
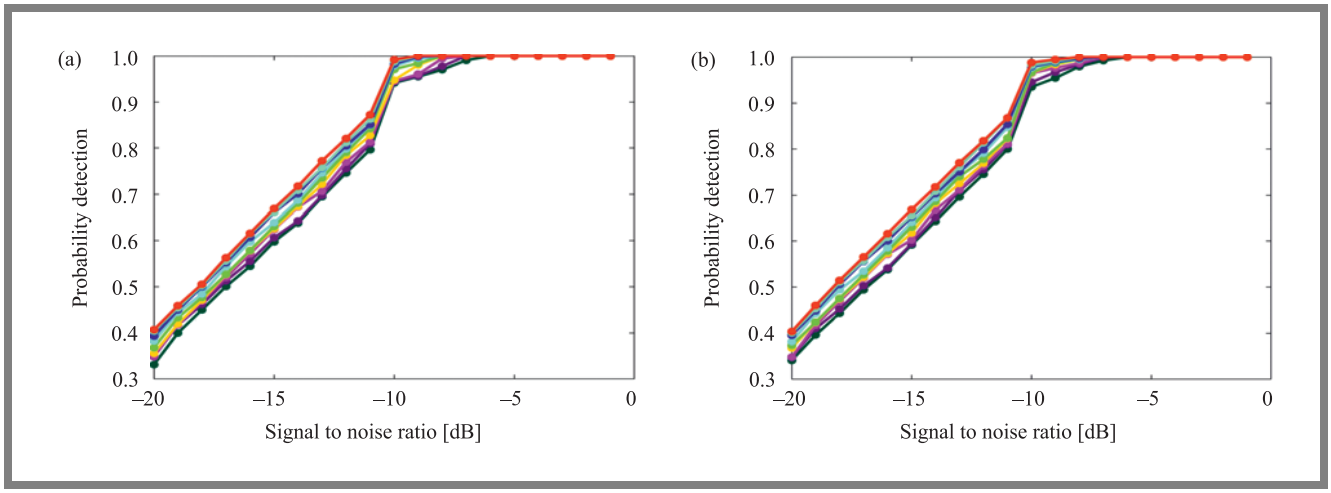



**Fig. 2.** Sample signal: a) BPSK, b) QPSK, c) PAM 4, and d) QAM 64.



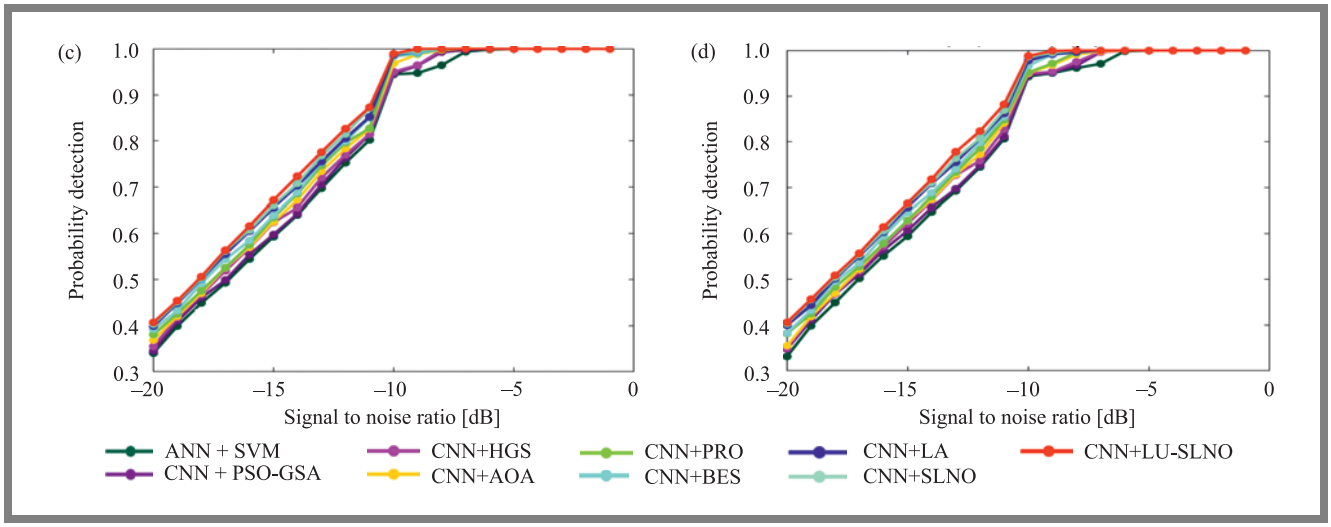



**Fig. 3.** Analysis of probability detection via CNN + LU-SLNO, compared with other schemes, for ED and: a) BPSK, b) PAM 4, c) QAM 64, d) QPSK.

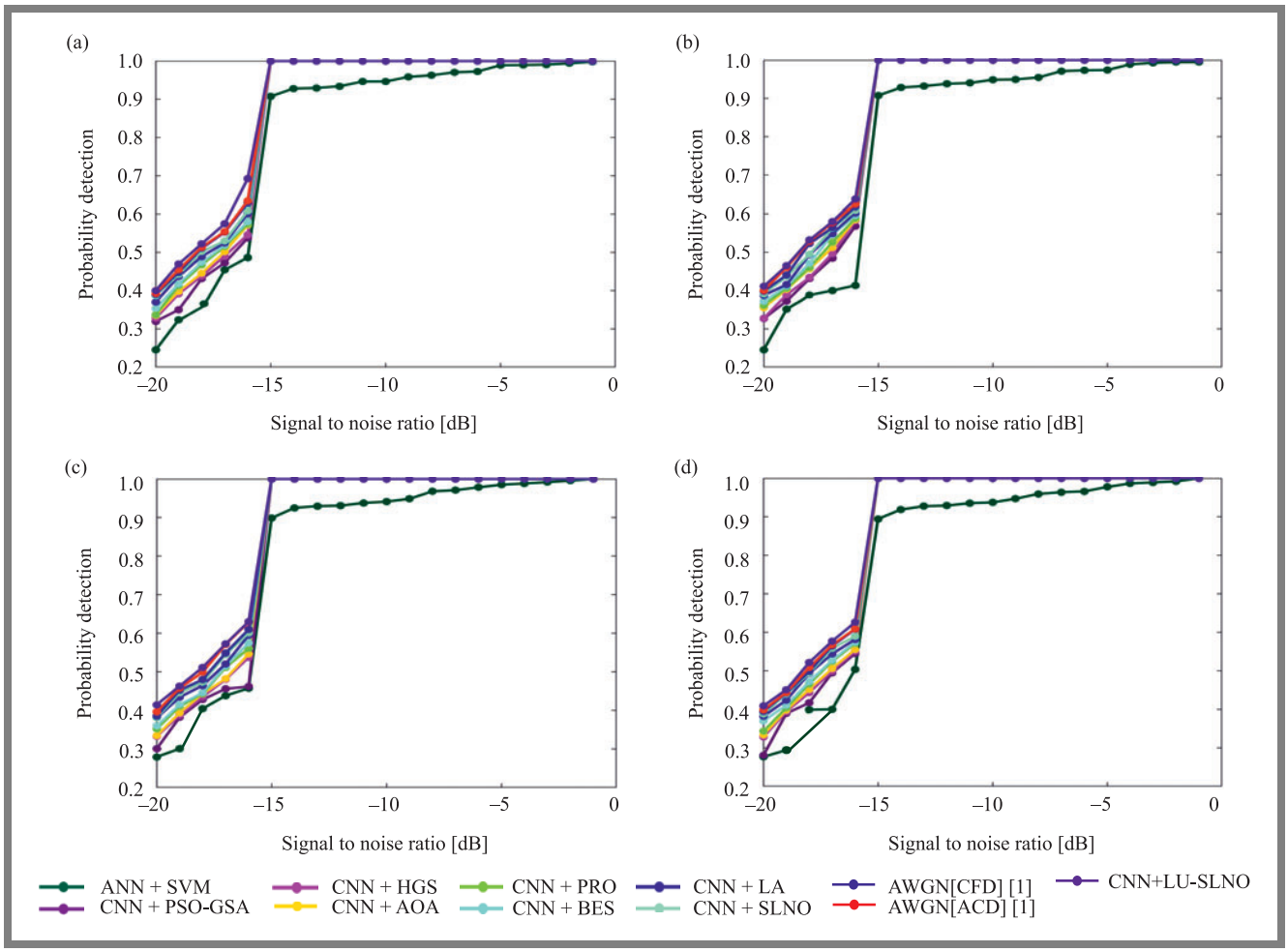



**Fig. 4.** Analysis of probability detection via CNN + LU-SLNO, compared with other schemes for MNHF and: a) BPSK, b) PAM 4, c) QAM 64, (d) QPSK.

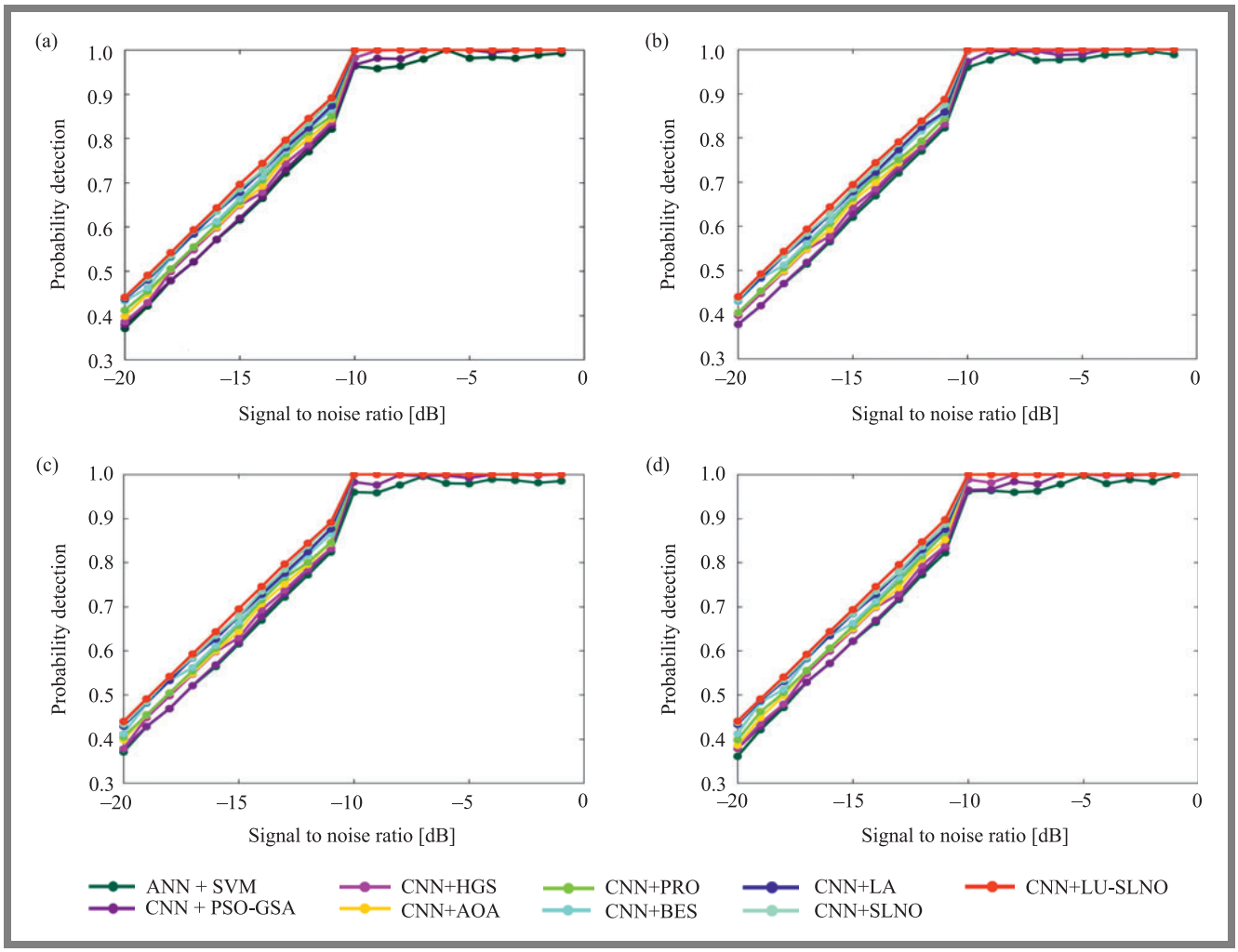



**Fig. 5.** Probability detection curves for CNN + LU-SLNO and SVD using: a) BPSK, (b) PAM 4, c) QAM 64, d) QPSK.

### *6.1. Analysis on Probability Detection*

Probability detection is calculated in CNN + LU-SLNO using traditional systems with BPSK, QPSK, PAM 4, and QAM 64 modulation, as well as using conventional schemes pictured in Figs. 3–5.

Figure 3 analyzes the results obtained using ED. Fig. 4 presents the MNMF simulation, and Fig. 5 shows the analysis of SVD. Here, CNN + LU-SLNO shows outcomes that are superior to those achieved for other schemes. Better probability detection is clearly visible, and for all SNRs, probability detection using CNN + LU-SLNO equals nearly 1.

**Tab. 4.** Results of error rate statistical analysis.

| Metrics | Std dev | Best | Variance | Mean | Worst |
|---|---|---|---|---|---|
| ANN+SVM | 1.330 | 1.770 | 1.908 | 3.231 | 0.686 |
| CNN+PSO-GSA | 1.317 | 1.735 | 1.890 | 3.200 | 0.680 |
| HGS | 1.237 | 3.146 | 1.530 | 1.887 | 0.669 |
| AOA | 1.183 | 3.280 | 1.400 | 2.068 | 0.667 |
| PRO | 1.144 | 3.190 | 1.308 | 2.099 | 0.890 |
| BES | 1.004 | 3.155 | 1.008 | 2.125 | 1.103 |
| LA | 1.100 | 3.183 | 1.209 | 2.036 | 0.848 |
| SLNO | 1.196 | 3.137 | 1.431 | 1.936 | 0.818 |
| LU-SLNO | 1.263 | 3.110 | 1.595 | 1.815 | 0.666 |

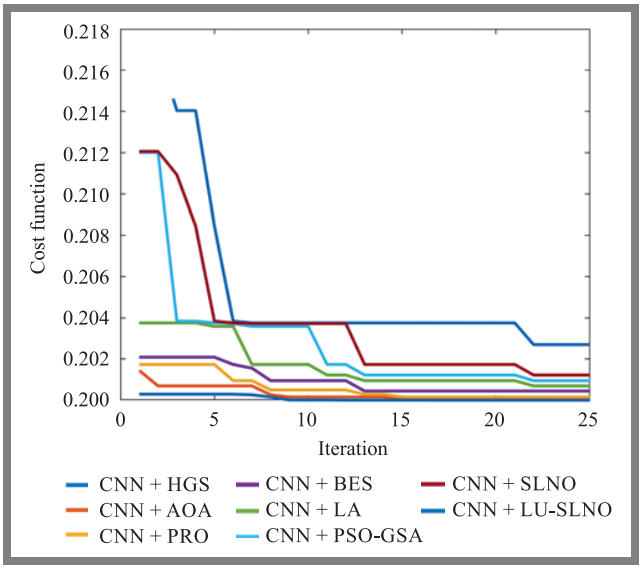



**Fig. 6.** Convergence analysis of LU-SLNO scheme and others method.

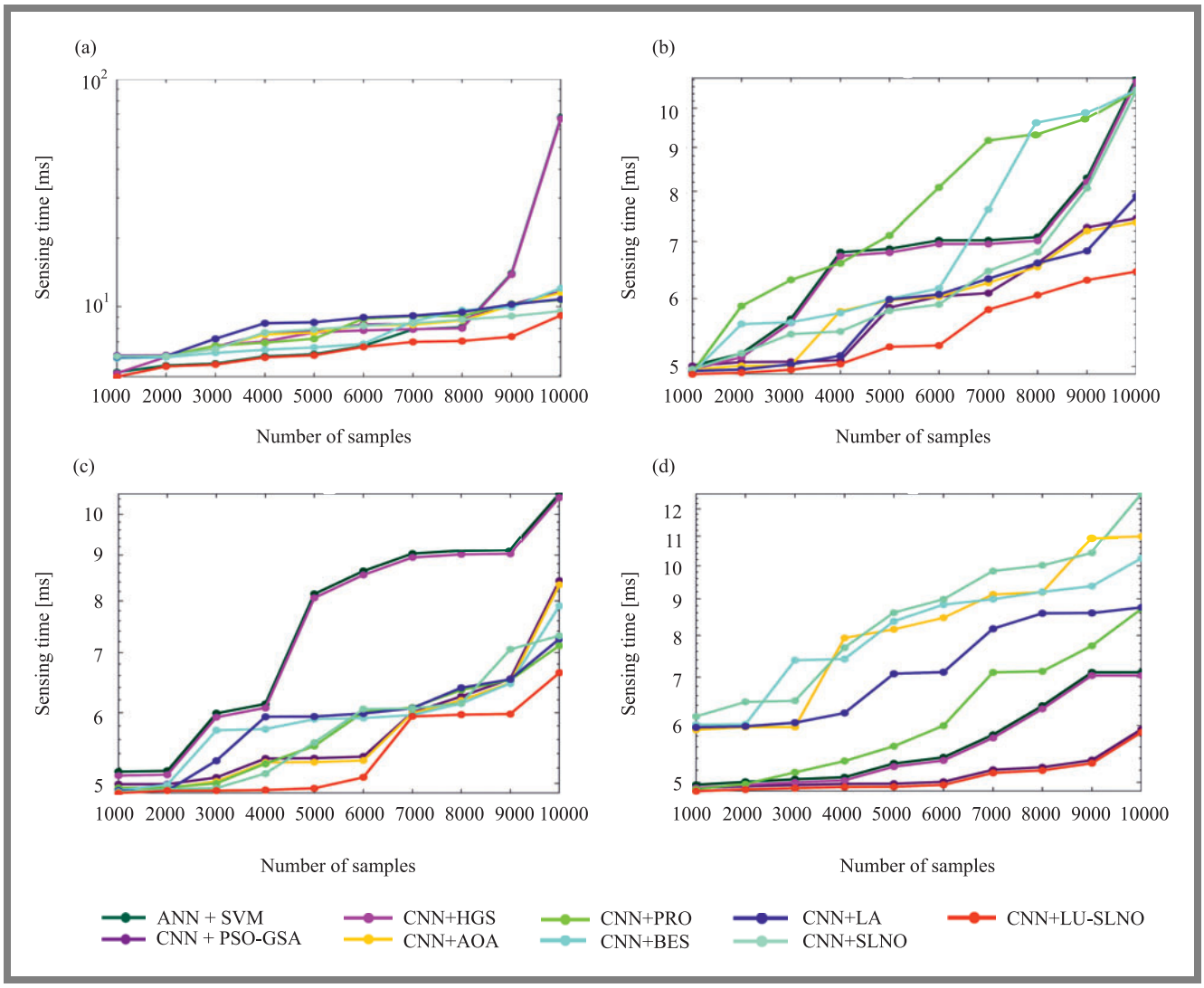
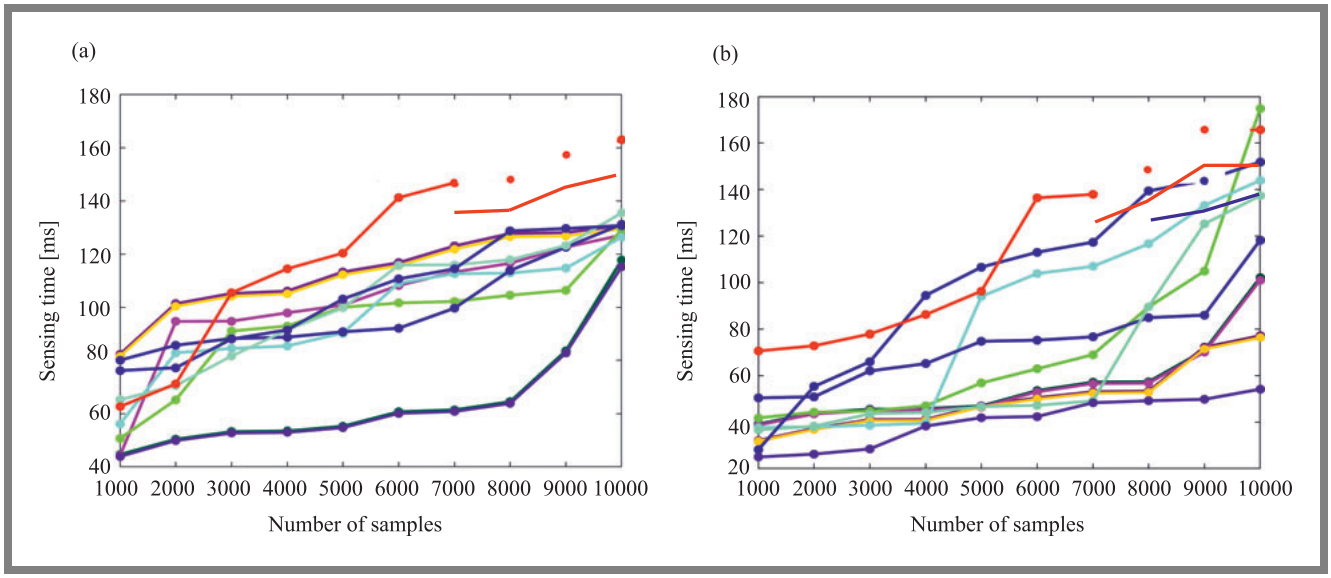



**Fig. 7.** Analysis of ED sensing time for CNN + LU-SLNO compared with other schemes for: a) BPSK, b) PAM 4, c) QAM 64, d) QPSK.



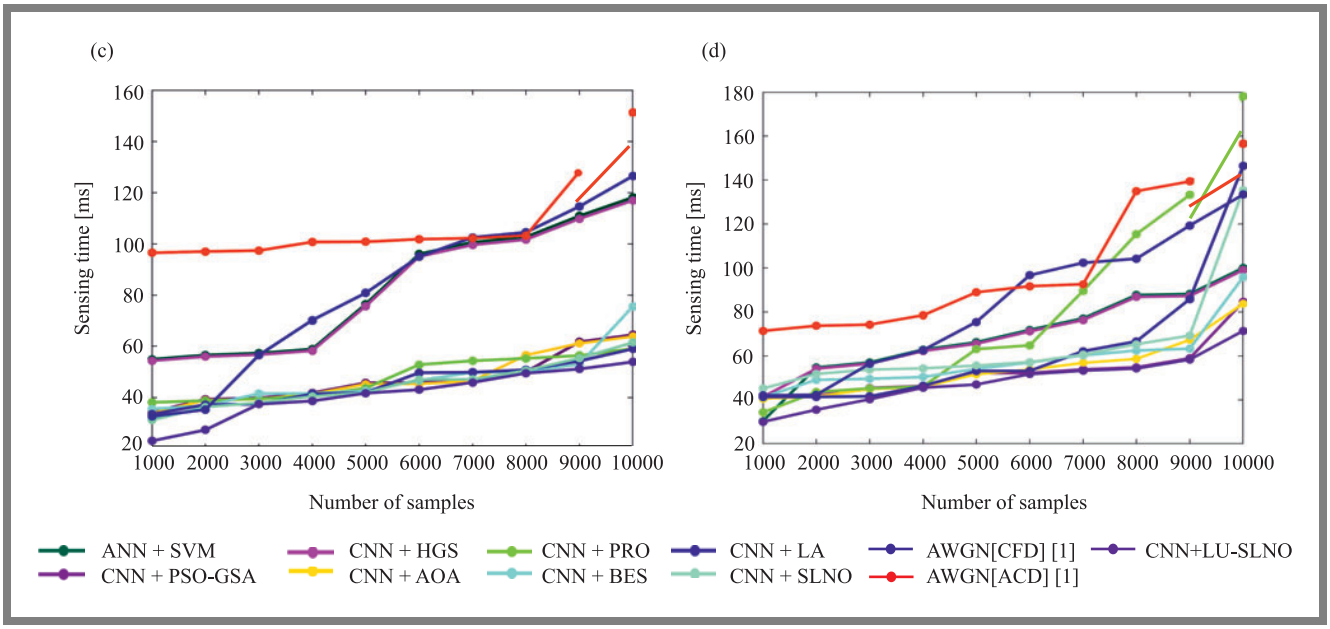



**Fig. 8.** Analysis of sensing time for CNN + LU-SLNO and other schemes (MNMF) for: a) BPSK, b) PAM 4, c) QAM 64, d) QPSK.

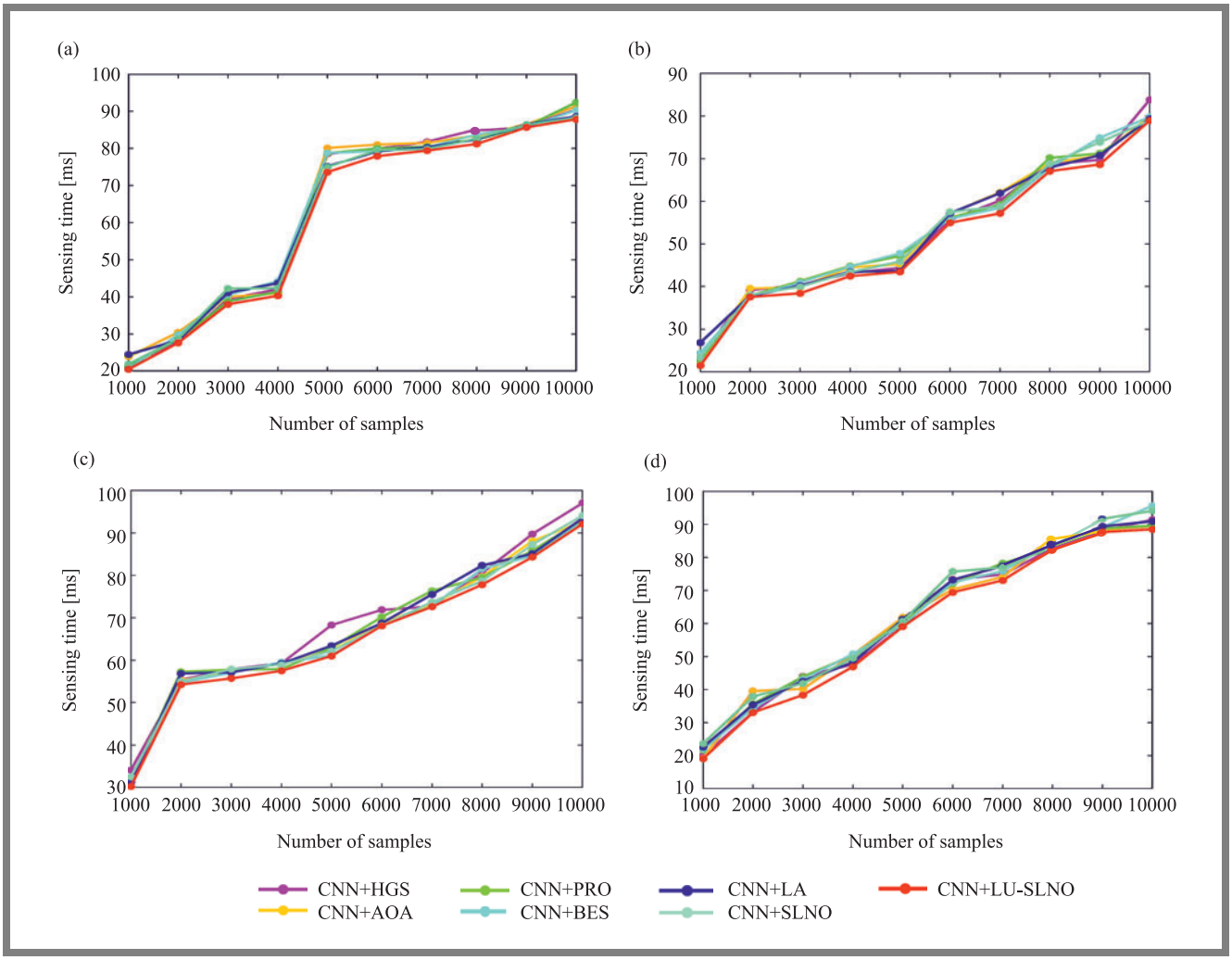



**Fig. 9.** Analysis of SVD sensing time for CNN + LU-SLNO compared with other schemes for: a) BPSK, b) PAM 4, c) QAM 64, d) QPSK.

### *6.2. Statistical Analysis*

Table 4 summarizes the statistical analysis of CNN + LU-SLNO versus other schemes. CNN + LU-SLNO shows almost no gain in terms of its output error rate over its competitors.

### *6.3. Convergence Analysis*

Figure 6 compares the cost of the LU-SLNO technique, for several iterations, with HGS, AOA, PRO, BES, LA, PSO-GSA, and SLNO. The cost of the LU-SLNO approach is lower for generations 0 to 25. In this case, the deployment of LU-SLNO results in lower convergence.

## 7. Sensing Time Analysis

The analysis of sensing time performed using the CNN + LU-SLNO method is presented in Figs. 7–9. The sensing time should be shorter to ensure improved system performance. The sensing time increases with the growing sample count. However, the sensing time using CNN + LU-SLNO is shorter when compared with others methods.

## 8. Conclusion

This article proposes a new MSS-CRN framework in which SNR estimation was relied upon in the channel selection process in cognitive radio. The detection strategy was selected in accordance with the estimated SNR of the received signal, using low and high values. For high SNR scenarios, ED and SVD were used, whereas for low SNR scenarios, MNMF was used. Together with the selected CNN weights and detectors, all SNRs and probability detection relying on CNN + LU-SLNO was superior than in other methods.

## References


[1] S. Thompson, A-T. Fadi, K.J. Suresh, N. Balamurugan, and S. Sweta, "Artificial intelligence inspired energy and spectrum aware cluster based routing protocol for cognitive radio sensor networks", *Journal of Parallel and Distributed Computing*, vol. 142, pp. 90–105, 2020 (DOI: 10.1016/j.jpdc.2020.04.007).

[2] L.A. Vasquez-Toledo, B. Borja-Benítez, J. Alfredo, and Tirado-Mendez, "Mathematical analysis of highly scalable cognitive radio systems using hybrid game and queuing theory", *AEU – International Journal of Electronics and Communications*, vol. 127, Article 153406, 2020 (DOI: 10.1016/j.aeue.2020.153406).

[3] E.G. Carla, R.C. Mario, and K. Insoo, "Relay selection and power allocation for secrecy sum rate maximization in underlying cognitive radio with cooperative relaying NOMA", *Neurocomputing*, 2021 (DOI: 10.1016/j.neucom.2020.08.082).

[4] G. Kamlesh, S.N. Merchant, and U.B. Desai, "A novel multistage decision fusion for cognitive sensor networks using AND and OR rules", *Digital Signal Processing*, vol. 42, pp. 27–34, 2015 (DOI: 10.1016/j.dsp.2015.04.).

[5] J. Qadir, A. Baig, A. Ali, and Q. Shafi, "Multicasting in cognitive radio networks: Algorithms, techniques and protocols", *Journal of Network and Computer Applications*, vol. 45, pp. 44–61, 2014 (DOI: 10.1016/j.jnca.2014.07.024).

[6] S.M. Budaraju and M.A. Bhagyaveni, "A novel energy detection scheme based on channel state estimation for cooperative spectrum sensing", *Computers and Electrical Engineering*, vol. 57, pp. 176–185, 2017 (DOI: 10.1016/j.compeleceng.2016.08.017).

[7] R. Sekaran, S.N. Goddumarri, and D. Gupta, "5G Integrated Spectrum Selection and Spectrum Access using AI-based Frame work for IoT based Sensor Networks", *Computer Networks*, vol. 186, 2020 (DOI: 10.1016/j.comnet.2020.107649).

[8] A. Kumar, S. Saha, and R. Bhattacharya, "Wavelet transform based novel edge detection algorithms for wideband spectrum sensing in CRNs", *AEU – International Journal of Electronics and Communications*, vol. 84, pp. 100–110, 2017 (DOI: 10.1016/j.aeue.2017.11.024).

[9] S.N. Kumar and K. Bikshalu, "An extensive survey on cognitive radio based spectrum sensing using different algorithms", *Materials Today: Proceedings*, 2020 (DOI: 10.1016/j.matpr.2020.11.482).

[10] P. Feng, Y. Bai, and C. Liu, "A rapid coarse-grained blind wideband spectrum sensing method for cognitive radio networks", *Computer Communications*, vol. 166, pp. 234–243, 2020 (DOI: 10.1016/j.comcom.2020.11.015).

[11] R. Ahmed, Y. Chen, and L. Du, "CR-IoTNet: Machine learning based joint spectrum sensing and allocation for cognitive radio enabled IoT cellular networks", *Ad Hoc Networks*, vol. 112, Article 102390, 2020 (DOI: 10.1016/j.adhoc.2020.102390).

[12] G.P. Aswathy, K. Gopakumar, and T.P. Imthias Ahamed, "Joint sub-Nyquist wideband spectrum sensing and reliable data transmission for cognitive radio networks over white space", *Digital Signal Processing*, vol. 101, Article 102713, 2020 (DOI: 10.1016/j.dsp.2020.102713).

[13] E. Geoffrey and T. Shankar, "Hybrid PSO-GSA for energy efficient spectrum sensing in cognitive radio network", *Physical Communication*, vol. 40, Article 101091, 2020 (DOI: 10.1016/j.phycom.2020.101091).

[14] M.A. Osameh, M.A. Hisham, and A-M. Haithem, "Efficient on-demand spectrum sensing in sensor-aided cognitive radio networks", *Computer Communications*, vol. 156, pp. 11–24, 2020 (DOI: 10.1016/j.comcom.2020.03.032).

[15] K. Alok, T. Prabhat, and G. Singh, "Threshold selection and cooperation in fading environment of cognitive radio network: Consequences on spectrum sensing and throughput", *AEU – International Journal of Electronics and Communications*, vol. 117, Article 153101, 2020 (DOI: 10.1016/j.aeue.2020.153101).

[16] Y. Gao, C. Wang, and X. Bai, "GLRT-based spectrum sensing by exploiting multitaper spectral estimation for cognitive radio network", *Ad Hoc Networks*, vol. 109, Article 102289, 2020 (DOI: 10.1016/j.adhoc.2020.102289).

[17] K.G. Walid and S. Suzan, "Performance evaluation of cooperative eigenvalue spectrum sensing GLRT under different impulsive noise environments in cognitive radio", *Computer Communications*, vol. 160, pp. 567–576, 2020 (DOI: 10.1016/j.comcom.2020.06.033).

[18] K. Mourougayane, B. Amgothu, and S. Srikanth, "A robust multistage spectrum sensing model for cognitive radio applications", *AEU – International Journal of Electronics and Communications*, vol. 110, Article 152876, 2019 (DOI: 10.1016/j.aeue.2019.152876).

[19] M. Aloqaily, H. Bany, S. Jalel, and B. Othman, "A multi-stage resource-constrained spectrum access mechanism for cognitive radio IoT networks: Time-spectrum block utilization", *Future Generation Computer Systems*, vol. 110, pp. 254–266, 2020 (DOI: 10.1016/j.future.2020.04.022).

[20] J.B. Patel, S. Collins, and B. Sirkeci-Mergen, "A framework to analyze decision strategies for multi-band spectrum sensing in cognitive radios", *Physical Communication*, vol. 42, Article 101139, 2020 (DOI: 10.1016/j.phycom.2020.101139).

[21] H. Qing, H. Li, and Y. Liu, "Multistage Wiener filter aided MDL approach for wideband spectrum sensing in cognitive radio networks", *AEU – International Journal of Electronics and Communications*, vol. 73, pp. 165–172, 2017 (DOI: 10.1016/j.aeue.2017.01.012).

[22] S. Kim, "Inspection game based cooperative spectrum sensing and sharing scheme for cognitive radio IoT system", *Computer Communications*, vol. 105, pp. 116–123, 2017 (DOI: 10.1016/j.comcom.2017.01.015).

[23] R.S. Rajput, R. Gupta, and A. Trivedi, "An Adaptive Covariance Matrix Based on Combined Fully Blind Self Adapted Method for Cognitive Radio Spectrum Sensing", *Wireless Pers Commun*. vol. 114, pp. 93–111, 2020 (DOI: 10.1007/s11277-020-07352-9).

[24] P. Vijayakumar, S.W. Malarvizhi, "Wide band Full Duplex Spectrum Sensing with Self-Interference Cancellation-An Efficient SDR Implementation", *Mobile Netw Appl 22*, pp. 702–711, 2017 (DOI: 10.1007/s11036-017-0844-7).

[25] E. Geoffrey and T. Shankar. "Hybrid PSO-GSA for energy efficient spectrum sensing in cognitive radio network", *Physical Communication*, vol. 40 101091, 2020 (DOI: 10.1016/j.phycom.2020.101091).

[26] M. Saber, A. El Rharras, R. Saadane, A. Chehri, N. Hakem, and H.A. Kharraz, "Spectrum sensing for smart embedded devices in cognitive networks using machine learning algorithms", *Procedia Computer Science*, vol. 176, pp. 2404–2413, 2020 (DOI: 10.1016/j.procs.2020.09.311).

[27] M. Taleb, A. Amei, M. Jian, and J. Yingtao, "On a new approach to SNR estimation of BPSK signals", *Int J Electron Telecommun*, vol. 58, no. 3, 273–278, 2012 (DOI: 10.2478/v10177-012-0038-).

[28] A. Ijaz, A.B. Awoseyila, and B.G. Evans, "Improved SNR estimation for BPSK and QPSK signals", *IEEE Electron Lett*, vol. 45, 16 (DOI: 10.1049/el.2009.1759).

[29] J. Gu, *et. al.*, "Recent advances in convolutional neural networks", *Pattern Recognition*, vol. 77, pp. 354–377, 2018 (DOI: 10.1016/j.patcog.2017.10.013).

[30] R.M.T. Masadeh, B.A. Mahafzah, and A.-A. Sharieh, "Sea Lion Optimization Algorithm", *International Journal of Advanced Computer Science and Applications*, vol. no. 10, pp. 388–395, 2019 (DOI: 10.14569/IJACSA.2019.0100548).

[31] B.R. Rajakumar, "Impact of Static and Adaptive Mutation Techniques on Genetic Algorithm", *International Journal of Hybrid Intelligent Systems*, vol. 10, no. 1, pp. 11–22, 2013 (DOI: 10.3233/HIS-120161).

[32] B.R. Rajakumar, "Static and Adaptive Mutation Techniques for Genetic algorithm: A Systematic Comparative Analysis", *International Journal of Computational Science and Engineering*, vol. 8, no. 2, pp. 180–193, 2013 (DOI: 10.1504/IJCSE.2013.053087).

[33] S.M. Swamy, B.R. Rajakumar, and I.R. Valarmathi, "Design of Hybrid Wind and Photovoltaic Power System using Opposition-based Genetic Algorithm with Cauchy Mutation", *IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013)*, 2013, (DOI: 10.1049/ic.2013.0361).

[34] G. Aloysius and B.R. Rajakumar, "APOGA: An Adaptive Population Pool Size based Genetic Algorithm", *AASRI Procedia – 2013 AASRI Conference on Intelligent Systems and Control (ISC 2013)*, vol. 4, pp. 288–296, 2013, (DOI: 10.1016/j.aasri.2013.10.043).

[35] B.R. Rajakumar and G. Aloysius, "A New Adaptive Mutation Technique for Genetic Algorithm", *In proceedings of IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–7, 2012 (DOI: 10.1109/ICCIC.2012.6510293).

[36] B.W. Mukund and N. Gomathi, "Improved GWO-CS Algorithm-Based Optimal Routing Strategy in VANET", *Journal of Networking and Communication Systems*, vol. 2, no. 1, pp. 34–42, 2019 (DOI: 10.46253/jnacs.v2i1.a4).

[37] H.B. Sadashiv, S.F. Kodad, S.K. Ambekar, and D. Manjunath, "Enhanced Invasive Weed Optimization Algorithm with Chaos Theory for Weightage based Combined Economic Emission Dispatch", *Journal of Computational Mechanics, Power System and Control*, vol. 2, no. 3, pp. 19–27, 2019 (DOI: 10.46253/jcmps.v2i3.a3).

[38] N.J. Amolkumar and N. Gomathi, "DIGWO: Hybridization of Dragonfly Algorithm with Improved Grey Wolf Optimization Algorithm for Data Clustering", *Multimedia Research*, vol. 2, no. 3, pp. 1–11, 2019 (DOI: 10.46253/j.mr.v2i3.a1).

[39] M.H. Omar, H. Suhaidi, and A.N. Shahrudin, "Eigenvalue-based signal detectors performance comparison", *The 17th Asia Pacific conference on communications*, 2011 (DOI: 10.1109/APCC.2011.6477853).

[40] Z. Yonghong and C.L. Ying, "Eigenvalue-based spectrum sensing algorithms for cognitive radio", *IEEE Trans Commun 2009*, 57(6), pp. 1784–1793 (DOI: 10.1109/TCOMM.2009.06.070402).

[41] M.H. Omar, H. Suhaidi, A. Angela, and A.N. Shahrudin, "SVD based signal detector for cognitive radio networks", *13th International conference on computer modeling and simulation*, 2011 (DOI: 10.1109/UKSIM.2011.104).

[42] Z. Yonghong and C.L. Ying, "Maximum-minimum eigenvalue detection for cognitive radio", *IEEE eighteenth international symposium on personal, indoor and mobile radio communications*, pp. 1–5 2007 (DOI: 10.1109/PIMRC.2007.4394211)

[43] Y. Liu, Y. Liang, Q. Kuang, F. Xie, Y. Hao, Z. Wen, and M. Li, "Post-modified non-negative matrix factorization for deconvoluting the gene expression profiles of specific cell types from heterogeneous clinical samples based on RNA-sequencing data", *Journal of Chemometrics*, e2929, 2017 (DOI: 10.1002/cem.2929).

[44] –, https://www.kaggle.com/datasets/nolasthitnotomorrow/radioml2016-deepsigcom?resource=download.

---

**Sanjeevkumar Jeevangi,** Research Scholar

E-mail: sjeevangi8@gmail.com

Department of Electronics and Communication Engineering, Faculty of Engineering and Technology (Co-Ed), Sharnbasva University, India

**Shivkumar Jawaligi,** Professor and Dean

Department of Electronics and Communication Engineering, Faculty of Engineering and Technology (Co-Ed), Sharnbasva University, India

**Vilaskumar Patil,** Associate Professor

Department of Electronics and Communication Engineering, Faculty of Engineering and Technology (Co-Ed), Sharnbasva University, India

Paper

# Chaotic-based Orthogonal Frequency Division Multiplexing with Index Modulation

Sura H. Hussien and Hamsa A. Abdullah

*Department of Information and Communication Engineering, College of Information Engineering, Al-Nahrain University, Iraq*

https://doi.org/10.26636/jtit.2022.165422

**Abstract — Orthogonal frequency division multiplexing with index modulation (OFDM-IM), stands out among conventional communication technologies, as it uses the indices of the available transmit entities. Thanks to such an approach, it offers a novel method for the transmission of extra data bits. Recent years have seen a great interest in chaos-based communications. The spectrum-spreading signals used in chaotic signal modulation technologies are orthogonal to the existing mixed signals. This paper presents how well a non-coherent differential chaos shift keying communication system performs across an AWGN. Different types of detection methods are simulated, bit error rate and power spectral density are calculated and then compared with standard OFDM with index modulation. The results of simulations concerning the performance of a DCSK system, adding to the security of the proposed solution and offering a comparable bit error rate performance, are presented in the paper as well.**

***Keywords — AWGN, chaos communications, DCSK, index modulation.***

## 1. Introduction

In the literature, a wideband chaotic signal is used in chaotic communication channel as the carrier [1] in order to ensure an efficient form of communication compared with channels relying on multipath fading. As a result, the popularity of chaotic communication networks is similar to that of ultra-wideband communication systems [6]–[4] and spectrum communications [5]. The keying technique in differential chaos (DCSK) offers great performance while using a straightforward auto-correlation receiver (ACR) on time-varying or multipath fading channels, as shown in [7]. DCSK has been also thoroughly studied in a variety of communication scenarios, including wireless personal area networks (WPANs) in paper [8] and MIMO systems in articles [9], [10].

The development and analysis of an orthogonal multilevel code-shifted differential chaos shift keying built on OFDM (OFDM-OMCS-DCSK) proved that a chaotic signal that is capable of achieving high data rates with the use of the same time slot is created using the hybrid Gram-Schmidt algorithm and the Walsh coding function, thus providing a higher bandwidth efficiency compared with conventional solutions. Furthermore, BER performance of AWGN and multipath Rayleigh fading channels shows that the OFDM-OMCM-DCSK system achieves the best BER result, with a high data rate characteristic compared with OFDM-DCSK [11].

Paper [12] proposes a permutation index modulation model and a position index modulation OFDM-based chaotic multi-carrier communication system. It combines position and permutation index modulation, adding a new dimension for data transmission and enhancing spectrum efficiency. Thanks to multi-carrier modulation, the system also offers improved multipath immunity. The system's BER performance is compared with that of PI-DCSK, CCPK-OFDM, and CCPK. The outcomes of the simulations prove that it is an effective, non-coherent modulation technique [12].

In [13], the chaotic sequence was generated by modifying a tent map, one of the well-known chaos generators, to boost security. Performance tests and security analyses were conducted using the BER, mean squared error (MSE), and correlation metrics. The outcomes showed that the autocorrelation of the suggested tent map acts as a delta function, making it challenging to separate the signal's content from the noise. It also showed how much more sensitive the suggested tent map was to variations in the initial state and control parameters. Using the proposed method, the multi-carrier CSK system outperforms the CSK in terms of data transmission rate.

Paper [14] presents SR-ODBR-DCSK – a modulation system with an orthogonal double bitrate based on the DCSK technique. To transmit two signals simultaneously, the Walsh function is used and the continuous two-bit data information is modulated utilizing orthogonality of the chaotic signals, with the modulation technology generating two orthogonal chaotic signals. This method is compared with SR- DCSK and the results show good effectiveness of the SR-ODBR-DCSK system [14].

Although many systems with a chaotic map have been proposed in the literature to ensure secure transmissions using synchronization, these systems do not consider the use an OFDM-enhanced index-modulation system-based chaotic map. Therefore, the main objective of this paper is to design a chaotic-based OFDM-IM system that ensures a significant improvement in the performance in the noise channel and enhances the performance of the fading channel, as well as ensures a secure transmission via chaotic signals and achieves

better spectral efficiency. The proposed technique, transmitting data based on the indices of the active subcarriers of orthogonal frequency division multiplexing (OFDM), is known as index modulation-aided OFDM (OFDM-IM) and allows to create a multicarrier transmission technique with a high spectral and energy efficiency in 5G networks.

Index modulation (IM) is a new concept. It is a modulation technique that uses the index of several mediums to modulate the carrier, with the incoming data bit to be transmitted being divided into two parts. One part is used to modulate the phase and amplitude of the carrier and the other part is used to select the index (I) of the activated antenna that transmits the matching modulated signal. OFDM-IM may be employed in frequency, permutation, coding [15], and time domains [16].

The remainder of this article is organized as follows. Section 2 describes the proposed OFDM-IM based chaotic communication scheme. Section 3 covers the simulation results and presents a discussion, while Section 4 evaluates the performance. Finally, Section 5 presents the conclusion.

## 2. OFDM-IM with DCSK

Due to their sensitivity to initial values of signals produced by deterministic nonlinear dynamical systems [17], also such advantages of spread spectrum signals' as the reduction of multipath fading, anti-jamming, suppression of inter-user interference and secure communication, are used to define chaotic signals here. Coherent and non-coherent detection are the two basic categories in chaotic communication systems. A precise reproduction of the chaotic carrier that was used to modulate the information must be made at the receiver in coherent systems. Since chaotic synchronization at the receiver and accurate replication of the chaotic signals are not necessary for coherent systems, more chaotic digital modulation techniques have been proposed in recent years.

In DCSK, two chaotic sample functions represent each bit of information that will be transferred. While the first sample function acts as a reference, the second sample function conveys the data. By delivering a reference signal produced by a chaos generator in a row, bit 1 is transmitted, whereas bit $-1$ is sent by using a reference chaotic signal followed by an inverted copy of the same signal [18]. DCSK seems to be less sensitive to channel distortion, as distortion affects both the reference and data samples [19].

### *2.1. Chaotic Generator*

Chaotic systems can be divided into two categories: those defined by flows or differential equations and those characterized by chaotic maps. The definition of a chaotic concept is formulated as:

$$c(t) = \begin{cases} x(t), & \text{for } (l-1)T \leqslant t < \left(1-\frac{1}{2}\right)T \\ x\left(t-\frac{T}{2}\right), & \text{for } \left(l-\frac{1}{2}\right)T \leqslant t < lT \end{cases}. \quad (1)$$

If the $l$ symbol is $+1$, Eq. (1) takes the form of:

$$c(t) = \begin{cases} x, & \text{for } (l-1)T \leqslant t < \left(1-\frac{1}{2}\right)T \\ -x\left(t-\frac{T}{2}\right), & \text{for } \left(l-\frac{1}{2}\right)T \leqslant t < lT \end{cases}. \quad (2)$$

When the symbol is $-1$, $T$ and $c(t)$ stand for the bit duration and chaotic reference signal, respectively.

In this paper, three different kinds of chaotic maps are employed, namely those of the Lorenz, Henon, and quadratic type. The Lorenz three-dimensional chaotic map is generated using [20]:

$$\begin{aligned} X_{n+1} &= 1 = X_n Y_n - Z_n + bY_n \\ Y_{n+1} &= X_n \\ Z_{n+1} &= Y_n \end{aligned}, \quad (3)$$

where: $-1 < X < 1$, $-1.5 < Y < 1.5$, and $-1.5 < Z < 1.5$.

The Henon map system is generated as [21]:

$$\begin{aligned} X_{n+1} &= 1 = aX_n^2 + bY_n \\ X_{n+1} &= X_n, \end{aligned} \quad (4)$$

where $a = 1.4$ and $b = 0.3$ are system parameters.

The quadratic chaotic map is not linear but still is deterministic because the behavior of the system is defined by an equation which determines how the behavior of the map varies when the initial value of $X_0$ changes [22]:

$$X_{n+1} = a - X_n^2, \quad (5)$$

where $n$ is an iteration number and $\alpha$ is a chaotic parameter. For $\alpha = [1.5, 2]$, the system is considered as chaotic.

Without adverse effects for generalization, we assume that the subcarrier is used to send the reference chaotic signal. It is defined as:

$$c = \{x(1), x(2), \ldots, x(\beta)\}, \quad (6)$$

where $x(\beta)$ is generated by a chaotic map and $\beta$ is the spreading factor in DCSK.

The transmitter structure of an OFDM-IM-based chaotic system using $N$ subcarriers is shown in Fig. 1, while Tab. 1 presents the parameters of the system.

### *2.2. Transmitter Architecture*

Using the OFDM scheme and applying index modulation with DSCK to its signal, a total of $m$ data bits are split into $G$ subblocks within $n = N/G$ subcarriers, where $N = 2, 4, 8 \ldots$ and $1 < k < N$. The subcarrier indices can carry $P_1$ bits of index data as:

$$p_1 = \lfloor \log_2 C(n,k) \rfloor, \quad (7)$$

where $C(\cdot,\cdot)$ denotes the binomial coefficient. Each $p = m/G$ bit subblock is then used for index modulation. Due to the same and independent features of all activities within the subblocks, we can use the example of the $\beta$-th subblock, where $\beta \in \{1, \ldots, G\}$ separates the incoming $p$ bits into two parts. The first component made up of $p_1$ index bits

**Tab. 1.** OFDM-IM parameters and symbols used.

| Parameters | Description |
|---|---|
| $M$ | Number of information bits in each OFDM block |
| $p_1$ | The total number of bits that each subblock's active indices are mapped to |
| $p_2$ | Each subblock's total amount of modulation bits expressed as $p_2 = k \log_2 m$ |
| $m$ | m-ary modulation size |
| $G$ | Number of groups, with $p$ bits in each group |
| $p$ | The number of bits per group ($p = m/G$) |
| $N$ | OFDM subcarriers count |
| $K$ | Each subblock's active subcarrier count |
| $I(\beta)$ | Chosen indices for the subblock |
| $s(\beta)$ | Modulated subblock symbol |
| $c(t)$ | Chaotic map |

determines the subcarrier activation pattern (SAP) for the $\beta$-th subblock or the indices of the $k$ active subcarriers.

$$\mathbf{M} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{G,1} & b_{G,2} & \cdots & b_{G,N} \end{bmatrix}. \tag{8}$$

Assuming the subcarriers are ordered in an ascending order, the following information can be obtained:

$$I^{(\beta)} = \left\{ i_1^{(\beta)}, \ldots, i_k^{(\beta)} \right\}. \tag{9}$$

In the proposed system:

$$\mathbf{I} = \begin{bmatrix} I_{1,1} & I_{1,K} \\ I_{2,1} & I_{2,K} \\ \vdots & \vdots \\ I_{G,1} & I_{G,K} \end{bmatrix}, \tag{10}$$

where $K = 1, \ldots, k$.

The data symbol vector is produced by the second part which consists of $p_2 = k \log_2 M$ symbol bits:

$$S^{(\beta)} = \left[ s_1^{(\beta)}, \ldots, s_k^{(\beta)} \right]^{\mathrm{T}},$$

$$\mathbf{S} = \begin{bmatrix} S_{1,1} & S_{1,p_2} \\ S_{2,1} & S_{2,p_2} \\ \vdots & \vdots \\ S_{G,1} & S_{G,p_2} \end{bmatrix}, \tag{11}$$

where $S_k(\beta) \in \mathbf{S}$ with $K = 1, \ldots, k$ and $\mathbf{S}$ is the modulation of DCSK. Since there are $C(n, k)$ possible realizations of $\mathbf{I}(\beta)$, the mapping from $p_1$ bits to $\mathbf{I}(\beta)$ can be realized by a lookup table or the combinatorial method in OFDM-IM with PSK modulation, where each bit maps to one-bit complex number. In the proposed system each bit in the message is represented by 10 bits, which is a processing gain. Consequently, the transmitted signal, $e(t)$ is:

$$e(t) = \begin{cases} x_{i,k}, & \text{for } 1 < k \leqslant \beta \\ s_i x_{i,k-\beta}, & \text{for } 1 - k \leqslant 2\beta \end{cases}. \tag{12}$$

A graphical representation of Eq. (12) is presented in Fig. 1.

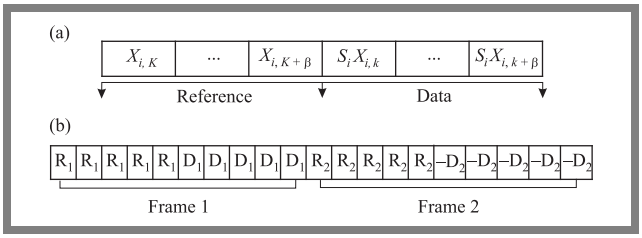



**Fig. 1.** Graphical representation of transmitted signal: a) general form and b) example for $m = 1, -1$.

Then on input the inverse FFT (IFFT) block of the same processes as those of classical OFDM are used:

$$X_T = \frac{N}{\sqrt{k}} \mathrm{IFFT}\{X_F\} = \frac{1}{\sqrt{k}} W_N^H X^F, \tag{13}$$

where $X_F$ is the cyclic prefix (CP) of length $[X(N - L + 1) \ldots X(N-1)X(N)]^{\mathrm{T}}$, $WN$: is the discrete Fourier transform (DFT) matrix with $w_N^H W_N$, $\frac{N}{\sqrt{k}}$ is the normalization term, and $L$ is the cycle prefix length.

Next, parallel-to-serial (P/S) and digital-to-analog (D/A) conversions are performed. Finally, an AWGN channel is used to transmit the generated signals (Fig. 2).

### *2.3. Receiver Architecture*

At the receiver end, after removing CP from the received signal and performing $N$-point FFT [23], the analog domain input-output relationship in the frequency domain is established by:

$$y = Hx + n, \tag{14}$$

in the formula entries of the channel matrix $\mathbf{H} = \mathrm{diag}\{h(1), \ldots, h(N)\}$ are complex Gaussian random variables with zero mean and unit variance, i.e. $h(\alpha?) \sim \mathrm{CN}(0, 1)$, and $n$ denotes the additive white Gaussian noise (AWGN) vector with $n(\alpha) \sim \mathrm{CN}(0, N_0)$. It is assumed that each non-zero symbol has an average transmit power of $\phi Es$, where $\phi = N/K$ is the power allocation coefficient and $Es$ is the average power per sub-carrier. Thus, the average signal-to-noise ratio (SNR) per active sub-carrier is given by $\gamma = \phi Es/N_0$.

Next, the output detector is compared with zero thresholds to recover the $i$-th information bit. There are three main types of detection algorithms used in the OFDM-IM system that map the incoming information bits to the subcarrier indices [24]:

– maximum likelihood (ML),

– log-likelihood ratio (LLR),

– greedy detector.

This method employs the ML detector, since the receiver has to know the set of possible indices for ML decoding. The ML detector for the OFDM-IM scheme considers all

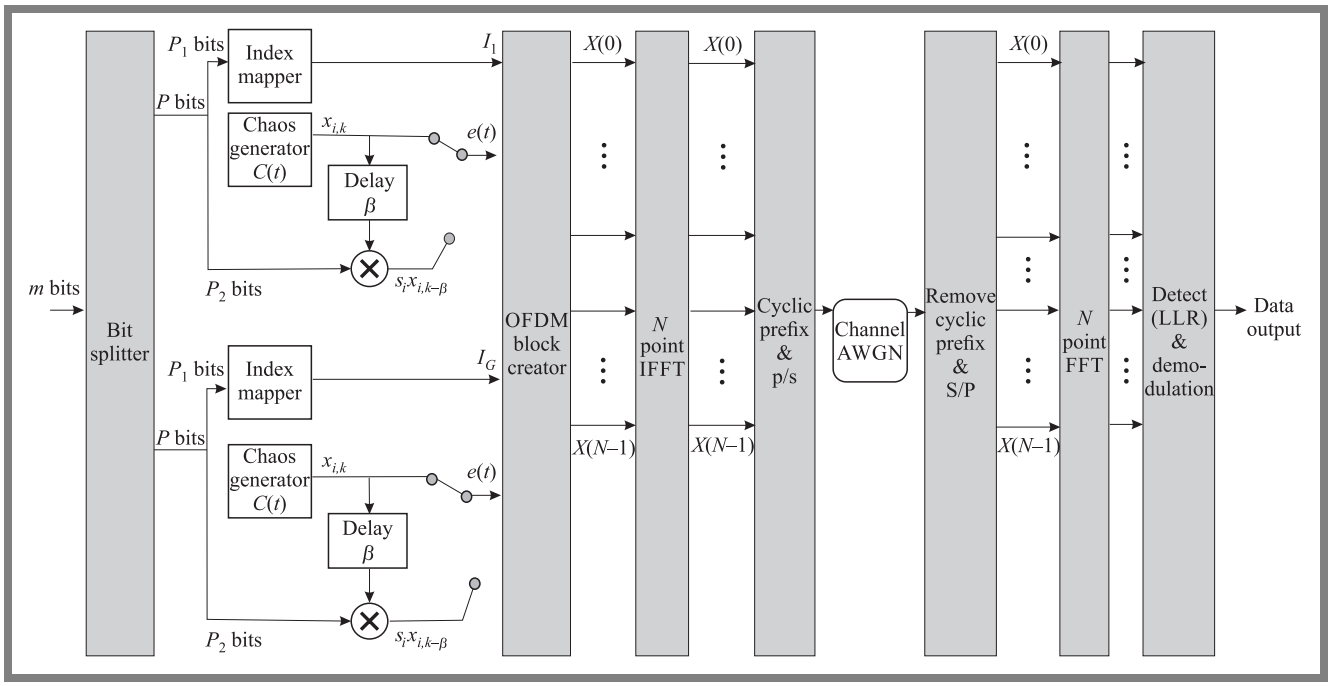



**Fig. 2.** OFDM-IM-based chaotic block diagram.

possible subblock realizations by searching for all possible subcarrier index combinations and signal constellation points, but the search complexity of the ML detector is of the order of $\sim\ O(\mathrm{CM}^k)$ per subblock, which becomes impractical for large $C$ and $k$ values.

Therefore, the receiver uses the log-likelihood ratio (LLR) detector to identify the most likely active subcarriers and the associated constellation symbols. LLR is an easy-to-use method based on combinatorial number systems. The BER simulation shown in Fig. 3 illustrates low dependability of the SAP detection method. We achieve BER about $10^{-3}$ in the greedy detector and about $10^{-4}$ with ML and LLR. Nevertheless, the greedy detector of OFDM-IM performs much worse than any other detectors due to the low reliability of the SAP detection approach.

## 3. Results and Discussion

In the proposed system, a Lorenz-type 3D map is used, as shown in Fig. 4, where the first subplot displays a chaotic map $\in \{-1, 1\}$ with a size equal to 10,000 with a spreading factor of 10. The second subplot is the message bit sequence with 100 random bits, while the third subplot is the modulated data with the same range as the chaotic map and is generated by a check bit sequence if it equals 1 or $-1$, and by copying five symbols of the chaotic map and duplicating them or inverse and send it to throw a channel. In Fig. 5, the same procedure is used but with a different initial state.

The simulation parameters of the system are summarized in Tab. 2.

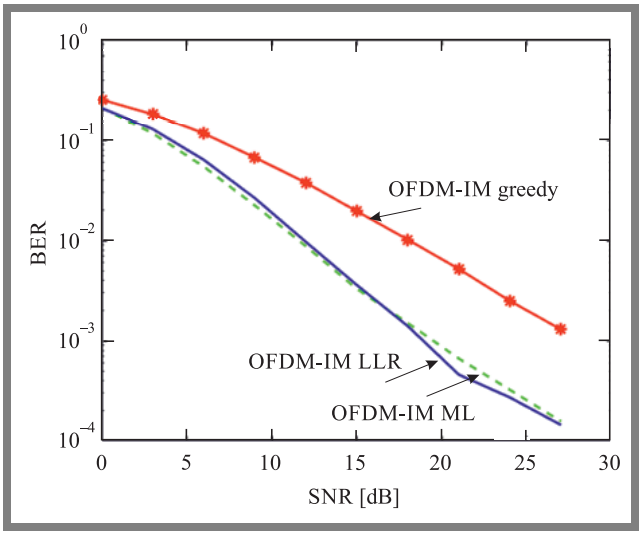



**Fig. 3.** Comparison of BER performance for different OFDM-IM detection types.

## 4. Performance Evaluation

The simulation results are presented in terms of BER probability and the Eb/No ratio, expressed in decibels, where Eb is the energy per user bit, and No is the single-sided spectral noise density. Here, we assume that channel interference is solely of the AWGN variety.

The goal of each communication system is to reduce BER. Channel disruptions, including noise and fading, are the main sources of errors. By limiting the effects of channel disruptions on the sent signal, the bit error rate can be decreased and, thus, the system optimized. Figure 6 shows BER for several chaotic map types, while Fig. 7 gives a comparison between OFDM

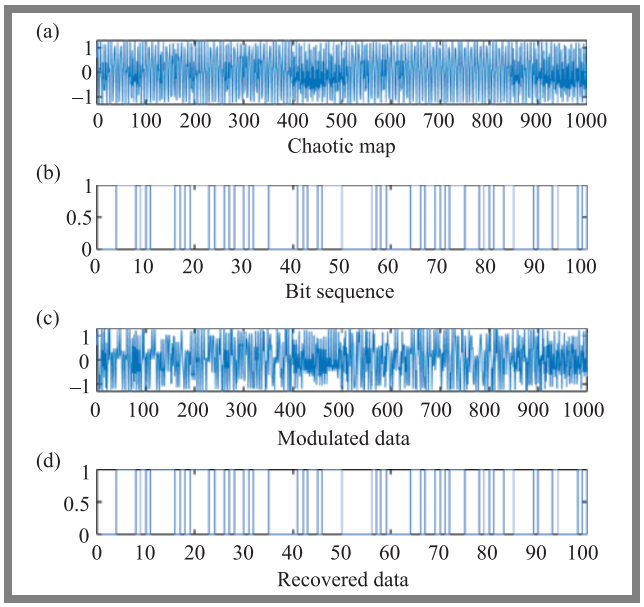



**Fig. 4.** a) Chaotic map generator, b) bit sequence, c) data modulated using DCSK, d) recovered data with the same initial value.

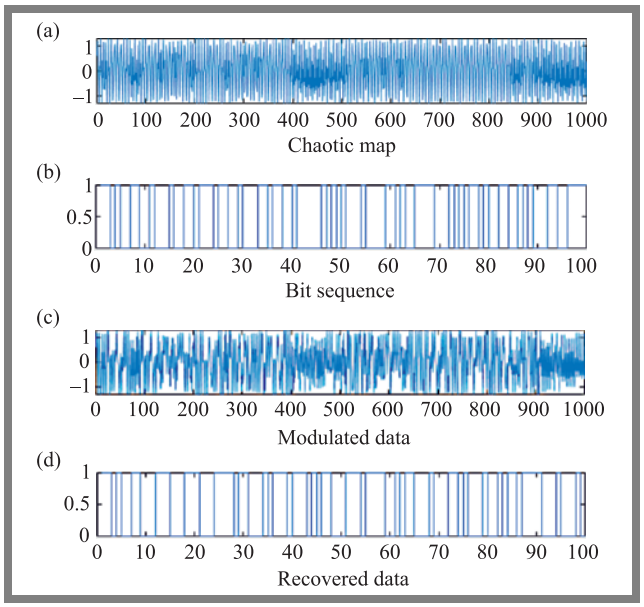



**Fig. 5.** a) chaotic map generator, b) bit sequence, c) data modulated using DCSK, d) demodulated data with a different initial state.

with index modulation and the proposed OFDM-IM-based chaotic system.

Figure 8 presents BER for an OFDM-IM-based chaotic system, for various channel state information.

Figure 9 shows a Lorenz map based on the proposed system and OFDM-IM transmitted symbols before transmission over the AWGN channel. With the Hamming window and 1024 FFT, the bandwidth of the system is 20 MHz and the PSD used of the two systems is near equal when $x = 0$, $y = -44.78$ in the chaotic system but in PSK, when $x = 0$, $y = -42.72$.

Figure 10 shows the PSD of the proposed system based on a Lorenz map and OFDM-IM transmitted symbols after transmission over the AWGN channel. The PSD of the two systems varies by approx. 20 dB, due to the effect of the

**Tab. 2.** Simulation parameters.

| Parameter values | Chaotic with OFDM-IM | OFDM-IM |
|---|---|---|
| $N$ | 4 | 4 |
| Modulation type | DCSK | PSK |
| Alphabet size, $M$ | 2 | 2 |
| $K$ | 2 | 2 |
| Number of sample per frame | 10000 | 10000 |
| Length of chaotic map | 1000 | / |
| $F_s$ | 20 MHz | 20 MHz |
| CP (cyclic prefix) | 1 | 1 |
| FFT | 64 | 64 |
| Window | Hamming | Hamming |

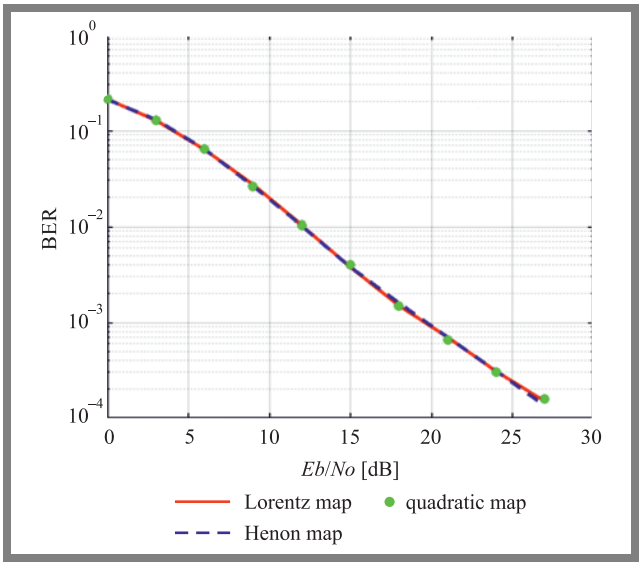



**Fig. 6.** Comparison between different types of chaotic maps.

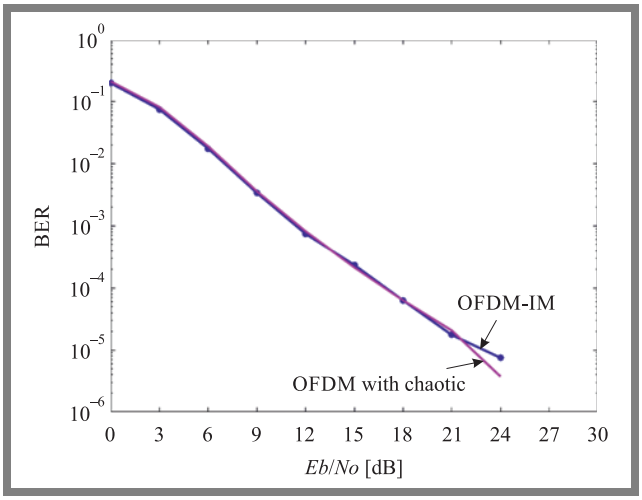



**Fig. 7.** BER Comparison between OFDM-IM and OFDM-based chaotic system.

channel on the system which transmitted the signal after the channel and the effect of AWGN when $x = 0$, $y = -62.73$ in BPSK. In the chaotic system, when $x = 0$, $y = -63.46$.

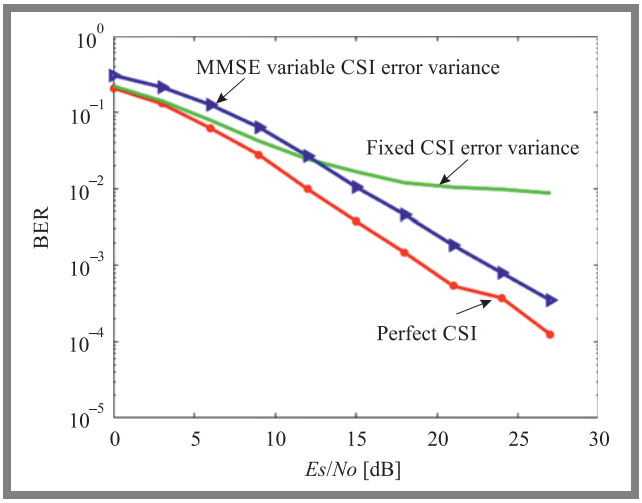



**Fig. 8.** OFDM-IM BER comparison for various channel states.

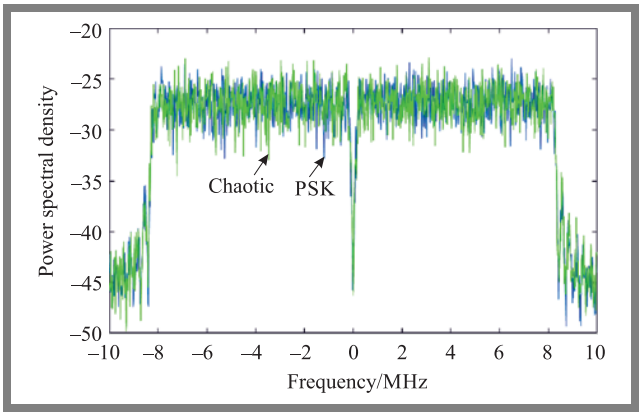



**Fig. 9.** OFDM-IM and a chaotic system before AWGN.

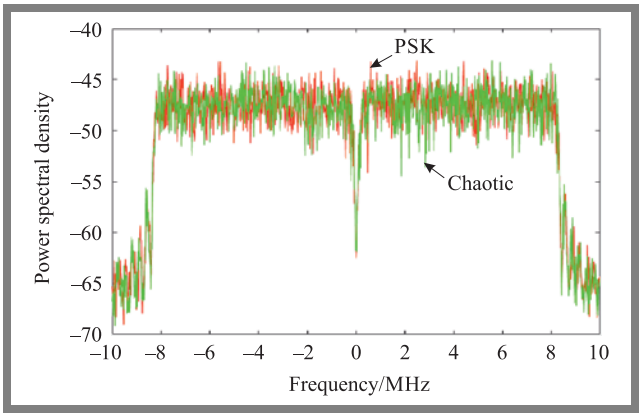



**Fig. 10.** Power spectral efficiency of OFDM-IM with chaotic after adding AWGN.

## 5. Conclusion

This paper presents a new combination of DCSK modulation with an OFDM-IM system. The goal behind using OFDM with index modulation and the DCSK system is to incorporate the benefits of both modulations, including robustness to impulse noise, high data throughput, and simplified equalization in multipath propagation. A comparison between different types of detection algorithms for OFDM-IM-based chaotic communication is presented and compared with the performance of BPSK and DCSK. The simulation results for a practical AWGN channel show that the simulated results are not worse in terms of PSD and BER than those of other traditional systems where the signal is recovered error-free at 27 dB. The proposed system provides a higher security level compared with traditional systems.

## References

[1] F.C.M. Lau and C.K. Tse, "Chaos-based digital communication systems: Operating principles, analysis methods, and performance evaluation", *Springer-Verlag*, Germany, 2003 (ISBN: 9783662051832).

[2] X. Min, W. Xu, L. Wang, and G. Chen, "Promising performance of an FM-DCSK UWB system under indoor environments", *IET Commun.*, vol. 4, no. 2, pp. 125–134, 2010 (DOI: 10.1049/iet-com.2008.0658).

[3] G. Kolumban, "UWB technology: Chaotic communications versus non-coherent impulse radio", *Proceedings of the 2005 European Conference on Circuit Theory and Design*, pp. 79–82, Ireland, 2005 (DOI: 10.1109/ECCTD.2005.1522997).

[4] S. Chen, L. Wang, and G. Chen, "Data-aided timing synchronization for FM-DCSK UWB communication systems", *IEEE Trans. Industrial Electronics*, vol. 57, no. 5, pp. 1538–1545, 2010 (DOI: 10.1109/TIE.2009.2038402).

[5] L. Wang, X. Min, and G. Chen, "Performance of SIMO FM-DCSK UWB system based on chaotic pulse cluster signals", *IEEE Trans. Circuits and Syst. I*, vol. 50, no. 9, pp. 2259–2268, 2011 (DOI: 10.1109/TCSI.2011.2112592).

[6] L. Ye, G. Chen, and L. Wang, "Essence and advantages of FMDCSK technique versus traditional spreading spectrum communication method", *J. Circuits, Systems and Signal Processing*, vol. 24, no. 5, pp. 657–673, 2005 (DOI: 10.1007/S00034-005-2413-8).

[7] M.P. Kennedy, G. Kolumban, G. Kis, and Z. Jako, "Performance evaluation of FM-DCSK modulation in multipath environments", *IEEE Trans. Circuits Syst. I*, vol. 47, no. 12, pp. 1702-1711, 2000 (DOI: 10.1109/81.899922).

[8] C.C. Chong and S.K. Yong, "UWB direct chaotic communication technology for low-rate WPAN applications", *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1527–1536, 2008 (DOI: 10.1109/TVT.2007.907089).

[9] Y. Fang, J. Xu, L. Wang, and G.R. Chen, "Performance of MIMO relay DCSK-CD systems over Nakagami fading channels", *IEEE Trans. Circuits and Syst. I*, vol. 60, no. 3, pp. 757–767, 2013 (DOI: 10.1109/TCSI.2012.2215755).

[10] G. Kaddoum and F. Gagnon, "Performance analysis of STBC-CSK communication system over slow fading channel", *Signal Processing*, vol. 93, no. 7, pp. 2055–2060, 2013 (DOI: 10.1016/j.sigpro.2012.12.020).

[11] F.S. Hasan and H.F. Fahad, "Design and analysis of an OFDM-based orthogonal multilevel code shifted differential chaos shift keying", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 3, pp. 1369–1378, 2020 (DOI: 10.11591/ijeecs.v20.i3).

[12] L. Hu, W. Xu, and L. Wang, "An OFDM-based Chaotic Chip Position Keying with Permutation Index Modulation", *14th International Symposium on Medical Information Communication Technology (ISMICT)*, Japan, 2020 (DOI: 10.1109/ISMICT48699.2020.9152730).

[13] S.S. Hasan and Z.M. Hussain, "Signal Properties under Multi-Carrier Chaos-Shift Keying", *Journal of Physics: Conference Series*, Iraq, 2021 (DOI: 10.1088/1742-6596/1804/1/012088).

[14] T. Sui, Y. Feng, Q. Jiang, F. Liu, and T. Zhang, "Design and Analysis of a Short Reference Orthogonal Double Bit Rate Differential Chaotic Shift Keying Communication", *Electronics*, vol. 11, 2022 (DOI: 10.3390/electronics11132020).

[15] E. Basar and S. Member, "Index Modulation Techniques for 5G Wireless Networks", *IEEE Communications Magazine*, vol. 54, no. 7, pp. 168–175, 2016 (DOI: 10.1109/MCOM.2016.7509396).

[16] M. Herceg, G. Kaddoum, D. Vranješ, and E. Soujeri, "Permutation index DCSK modulation technique for secure multiuser high-data-rate communication systems", *IEEE Trans. Vehi. Tech.*, vol. 67, no. 4, pp.

2997–3011, 2018 (DOI: 10.1109/TVT.2017.2774108).
[17] A. Abel and W. Schwarz, "Chaos communications-principles, schemes, and system analysis", *Proceedings of the IEEE*, vol. 90, no. 5, pp. 691–710, 2002 (DOI: 10.1109/JPROC.2002.1015002).
[18] G. Kolumban, *et al.*, "Differential chaos shift keying: A robust coding for chaos communication", *Proc. NDES96*, vol. 96, pp. 87–92, 1996 (https://www.researchgate.net/profile/Bela-Vizvari/publication/239666158_Differential_chaos_shift_keying_A_robust_coding_for_chaos_communication/links/0a85e532193a41c640000000/Differential-chaos-shift-keying-A-robust-coding-for-chaos-communication.pdf).
[19] E. Basar, U. Aygolu, E. Panayirci, and H.V. Poor, "Orthogonal frequency division multiplexing with index modulation", *IEEE Trans. Signal Process.*, vol. 61, pp. 5536–5549, 2013 (DOI: 10.1109/TSP.2013.2279771).
[20] J.C. Sprott, "Chaos and Time-Series Analysis", *Oxford University Press*, pp. 230–440, Oxford 2003 (ISBN: 9780198508403).
[21] M. Hénon, "Numerical study of quatratic area-preserving mappings", *Quart. Appl. Math.*, vol. 27, np. 3, pp. 291–312, 1969 (https://www.jstor.org/stable/43635985).
[22] N. Ramadan, H.A. Eldin, S.E. Elkhamy, and F.E.A. El-Samie, "Chaos-based image encryption using an improved quadratic chaotic map", *American Journal of Signal Processing*, vol. 6, no. 1, pp. 1–13, 2016 (DOI: 10.5923/j.ajsp.20160601.01).
[23] M. Wen, Q. Li, and X. Cheng, "Index modulation for OFDM communications systems", *Springer Singapore*, 2021 (DOI: 10.1007/978-981-15-9407-6).
[24] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain", *Proc. IEEE Int. Conf. Commun.*, USA, pp. 1009–1013. USA 1995 (DOI: 10.1109/ICC.1995.524253).

**Sura H. Hussien** received her B.Sc. in Information and Communication Engineering from the University of Baghdad in 2016, and an M.Sc. from the Information Engineering Department of Al-Nahrain University in Baghdad, Iraq. Her research interests include index modulation, wireless transmissions, physical layer security and servers.

E-mail: sura.alwash@coie-nahrain.edu.iq

Department of Information and Communication Engineering, College of Information Engineering, Al-Nahrain University, Iraq

**Hamsa A. Abdullah** obtained her B.Sc. in Information Engineering in 2005, M.Sc. in Information Engineering in 2008, and Ph.D. in Information and Communication Engineering in 2019, from Al-Nahrain University, Iraq. She has been working as a lecturer at the College of Information Engineering, at Al-Nahrain University, Iraq, since 2008. She is an Assistant Professor and an M.Sc. supervisor. She is interested in such subjects as: multimedia, security of multimedia transmissions over wireless communication systems, chaotic encryption and communication, pattern recognition, and embedded systems.

E-mail: hamsa.abdulkareem@coie-nahrain.edu.iq

Department of Information and Communication Engineering, College of Information Engineering, Al-Nahrain University, Iraq

Paper

# Multicriteria Oppositional-Learnt Dragonfly Resource-Optimized QoS Driven Channel Selection for CRNs

Ch.S.N. Sirisha Devi and Suman Maloj

*ECE Department, KL University, KLEF, India*

https://doi.org/10.26636/jtit.2022.164722

**Abstract — Cognitive radio networks (CRNs) allow their users to achieve adequate QoS while communicating. The major concern related to CRN is linked to guaranteeing free channel selection to secondary users (SUs) in order to maintain the network's throughput. Many techniques have been designed in the literature for channel selection in CRNs, but the throughput of the network has not been enhanced yet. Here, an efficient technique, known as multicriteria oppositional-learnt dragonfly resource-optimized QoS-driven channel selection (MOLDRO-QoSDCS) is proposed to select the best available channel with the expected QoS metrics. The MOLDRO-QoSDCS technique is designed to improve energy efficiency and throughput, simultaneously reducing the sensing time. By relying on oppositional-learnt multiobjective dragonfly optimization, the optimal available channel is selected depending on signal-to-noise ratio, power consumption, and spectrum utilization. In the optimization process, the population of the available channels is initialized. Then, using multiple criteria, the fitness function is determined and the available channel with the best resource availability is selected. Using the selected optimal channel, data transmission is effectively performed to increase the network's throughput and to minimize the sensing time. The simulated outputs obtained with the use of Matlab are compared with conventional algorithms in order to verify the performance of the solution. The MOLDRO-QoSDCS technique performs better than other methods in terms of throughput, sensing time, and energy efficiency.**

***Keywords — cognitive ratio network (CRN), multicriteria dragonfly optimization, oppositional learning, optimal available channel, QoS metric***

## 1. Introduction

Back in the days, efficient channel allocation was achieved in CRNs was by means of the adjustable multi-radio channel hopping (AMCH) protocol [1]. Although the protocol solves the rendezvous concern in CRN, the throughput of the network has not been improved. To handle this problem, a novel MOLDRO-QoSDCS optimization technique has been designed, relying on multiple criteria for enhancing the throughput performance of CRNs.

Machine learning was used in [2] for developing a three-phase target channel sequence (TCS) generation and allocation scheme for dynamic channel allocation in CRNs. It solves the fairness, individuality, and dynamic TCS generation issues, but the sensing time was not minimized. To solve this issue, the oppositional learning concept was designed to arrive at a global solution quicker.

Better throughput and shorter service times were obtained in CRNs by applying the extended generalized predictive channel selection algorithm (EXGPCSA), as presented in [3]. However, energy consumption of the network was not minimized. The Q learning-based dynamic optimal band and channel selection technique was developed in [4] for improving the accessible transmission time. Although the data rate was increased, QoS-aware channel selection was not performed. The cross-layered channel assignment algorithm was presented in [5] to offer a consistent route scheme in CRNs. However, the delay of the network was not reduced due to the congestion in the routes. In [6], QoS-aware routing protocols were described for selecting the channel and the route in CRNs. The research conducted has contributed to optimizing throughput and delay, but energy efficiency was not improved.

To solve these issues, this paper proposes the MOLDRO-QoSDCS technique for improving the performance of CRNs. This paper addresses the major issue of energy efficiency in the CRN channel allocation process. It also solves the issues of higher power consumption, signal-to-noise ratio sensing time, and lower spectrum utilization in QoS-aware channel selection in data communication. The proposed MOLDRO-QoSDCS technique offers the following innovative approaches:

– It uses the multicriteria dragonfly optimization algorithm to select the optimal channel available in a CRN. Such a process integrates the oppositional learning concept for finding the optimal available channel via the fitness function. On the contrary to other optimization techniques, the proposed dragonfly optimization algorithm computes multiple criteria, such as signal-to-noise ratio (SNR), power consumption and spectrum utilization in the course of the channel selection process;

– Multicriteria dragonfly optimization exploits the oppositional learning concept to initialize the opposite population with the current population for obtaining a global optimum solution in the channel selection process;
– The ranking method is employed to minimize the sensing time for obtaining the local optimal solution from the current and opposite populations. With this approach, the global optimum solution is obtained within shorter times;
– To evaluate performance of the MOLDRO-QoSDCS technique and to compare it to contemporary state-of-the-art methods, a Matlab simulation analysis is performed. It shows that the proposed method outperforms the state-of-the-art models in terms of various metrics.

The rest of this paper is organized as follows. Section 2 offers a review of the literature. Section 3 presents, in detail, the proposed MOLDRO-QoSDCS technique. Section 4 describes and discusses the simulation settings and evaluation metrics. Finally, Section 5 summarizes the paper.

## 2. Related Works

The joint FD-aware channel assignment and route selection protocol for CRN was presented, under time-changing channel conditions in [7]. However, QoS metrics are not considered in the channel selection process. An intelligent access network selection algorithm was designed in [8], depending on multiple attribute decision-making (MADM), but the throughput was not improved. A new approach was presented in [9] to assign the resources and distribution depending on the cooperative game theory in order to ensure maximized payoff. However, time and complexity have remained higher. Reinforcement learning (RL) based schemes were introduced in [10] to assign the resources in CRNs. However, no optimal channel allocation approach was proposed.

In article [11], the cross-layer algorithm, defined as multicast routing, channel selection, scheduling, and call admission control (MRCSC), was designed for the purpose of multi-radio CRNs. Utilization of the network's bandwidth was improved, but the sensing time was not measured. The cross-layer routing scheme that supports adaptive bitrate multimedia (ABM) was developed in [12] to enhance the efficiency of channel utilization, but the optimal channel selection was not researched.

A two-phase spectrum handoff methodology was presented in [13] to choose the best channel and minimize channel drop in CRNs. However, throughput was not improved. The discrete quality factors aware channel scheduling (DQFACS) method was researched in [14] to lower the recurrent transmission in a cognitive radio ad-hoc network (CRAN), but QoS-related metrics were not considered when scheduling the channels. A game theory-based algorithm was considered in [15] to assign sensors to diverse channels in order to minimize energy consumption by satisfying a few channels sensing-related constraints. Unfortunately, the throughput was lower. Another new method was designed in [16] by integrating artificial intelligence techniques and spectrum detection algorithms to allocate channels in CRNs. However, time consumption has remained higher in such an approach.

A dynamic channel selection mechanism was developed in [17]. It relies on the fuzzy inference system (FIS). It selects the channel for lowering re-transmission probability. The expected throughput, data latency, and communication reliability were achieved, but the issue of energy use was not researched. A second order Kalman filter technique was analyzed in [18] to estimate channels for sensing the spectrum in CRNs. Unfortunately, the sensing time was not reduced.

To select channels in CRNs, the multi-objective optimization based on a ratio analysis (MOORA) algorithm was developed in [19]. However, the energy efficiency of the CRN was not improved. A proactive spectrum handoff scheme was developed in [20] for choosing the target channel in CRN, but the channel collision failed to be minimized. The Brownian motion-based dragonfly algorithm (DA) was designed in [21] to address the spectrum assignment (SA) problem. However, multiple criteria were not used there. Finally, the adaptive threshold-based dragonfly optimization method was introduced in [22] for cooperative spectrum sensing in CRN in order to identify spectrum holes.

## 3. Proposed Method

The MOLDRO-QoSDCS technique is dedicated to improving cognitive radio networks by selecting the optimal channel with a better QoS. A block diagram of a QoS-aware channel selection process using the proposed MOLDRO-QoSDCS technique is shown in Fig. 1.

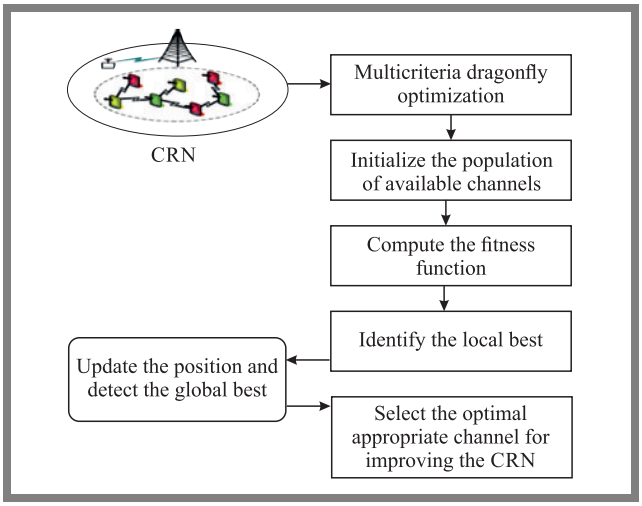



**Fig. 1.** Block diagram of the proposed MOLDRO-QoSDCS technique.

Let us consider a CRN consisting of $n$ primary users (PU) and $m$ secondary users (SU) with the available channels identified as $C = C_1, C_2, \ldots, C_n$. The proposed multicriteria oppositional-learnt dragonfly resource optimization technique is categorized as a meta-heuristic technique that is applied to determine the solution of an optimization problem. In MOLDRO-QoSDCS, the multicriteria include SNR, power

consumption, and spectrum utilization. In contrast to conventional approaches, the newly developed technique utilizes an oppositional-based learning model for acquiring the global best result within the population in a shorter period of time. By using the oppositional-based learning optimization algorithm, such parameters as convergence speed, flexibility, error tolerance, and accuracy are improved.

Roaming and searching for food sources is a natural activity of dragonflies. In the proposed technique, the dragonfly is related to the number of available channels and the food source is related to multi-criteria functions, i.e. signal-to-noise ratio, power consumption, and spectrum utilization. The population of dragonflies (i.e. available channels) is initialized and the dragonflies move within the search space.

To begin the optimization process, the channel population existing in the search space is first initialized and is given by:

$$A = \{C_1, C_2, \ldots, C_n\}, \tag{1}$$

where $A$ is the current population of channels $C_1, C_2, \ldots, C_n$. To obtain the global best solution, the proposed technique creates an opposite population from the current population through the opposition-based learning concept as:

$$A' = x_i + y_i - A, \tag{2}$$

where $A'$ is the opposite population generated based on the current population $A$, $x_i$ and $y_i$ are the lower and higher values of $A$. With this, the current $A$ and the opposite $A'$ population are created in the search space. Upon completing the population generation process, the objective function is evaluated for each dragonfly (i.e. channel) in $A$ and $A'$ according to the abovementioned multiple criteria.

First, SNR is calculated using the number of users and their transmission rates:

$$S_{nr} = \frac{T_p \cdot G}{\sigma^2 + I_p}, \tag{3}$$

where $T_p$ symbolizes the transmission power, $G$ is the gain, $\sigma^2$ represents the noise power, and $I_p$ stands for the power of interference created by other users. Subsequently, spectrum utilization is:

$$S_u = \frac{T_{\mathrm{BR}}}{\mathrm{BW}_c}, \tag{4}$$

where $S_u$ represents the spectrum utilization, $T_{\mathrm{BR}}$ refers to the transmitted bitrate that is measured in bps, and $\mathrm{BW}_c$ refers to the channel's bandwidth (in Hz). Afterwards, power consumption of the channel is determined as:

$$P_{\mathrm{con}} = P_T - P_{ut}, \tag{5}$$

where $P_{\mathrm{con}}$ indicates the channel's power consumption, $P_T$ is the total power and $P_{ut}$ denotes the amount of power used. Depending on the above-calculated metrics, fitness is calculated as follows:

$$\alpha_F = \{S_{nr}, S_u, P_{\mathrm{con}}\}, \tag{6}$$

where the maximum spectrum utilization is $S_u$ and the minimum signal-to-noise ratio is $S_{nr}$. Power consumption of the channel selected in CRN is $P_{\mathrm{con}}$. Following fitness estimation, the current and opposite populations are merged into one and the dragonflies are sorted with respect to their fitness value. The channel with higher $S_u$ and lower $S_{nr}$ and $P_{\mathrm{con}}$ values is positioned first.

$$C = \alpha_F(C_1) > \alpha_F(C_2) > \ldots > \alpha_F(C_b). \tag{7}$$

With this, $n$ best available channels are chosen from in order to identify the global best solution. Depending on fitness, the global best dragonfly in the search space is determined by performing such processes as separation, alignment, cohesion, and attraction towards the food source. To begin with the separation process, the current position of the dragonfly and its neighboring position is estimated as:

$$w = \sum_{j=1}^{n} \left[ g(t) - g_j(t) \right], \tag{8}$$

where $w$ is the separation process, $g(t)$ is the present position of an available channel at time $t$, $g_j(t)$ is the neighboring channel at time $t$ and $m$ indicates the number of neighboring dragonflies. Velocity of the dragonflies is determined and the neighborhood matching procedure is performed in the alignment process. It is computed as:

$$x = \sum_{j=1}^{m} \frac{\theta_j(t)}{m}, \tag{9}$$

where $x$ is the alignment, $\theta_j(t)$ is the velocity of neighboring dragonflies, and $m$ is the number of several neighboring dragonflies. Another process is called cohesion. Dragonflies tend to travel to the middle of their neighborhood. It is calculated as:

$$y = \sum_{j-1}^{m} \frac{g_j(t) - g(t)}{m}, \tag{10}$$

where $y$ indicates the cohesion of the dragonfly, $g_j(t)$ is the position of an adjacent dragonfly, $g(t)$ stands for the position of a recent dragonfly, $m$ is the neighborhood. Lastly, the attraction towards food sources is calculated:

$$z = g_F(t) - g(t), \tag{11}$$

where $z$ is the attraction function, $g_F(t)$ is the location of the food source, and $g(t)$ is the present dragonfly location. Finally, the position of the dragonfly gets updated to discover the global best solution as:

$$g_{t+1} = g(t) + \nabla g_{t+1}, \tag{12}$$

where $g_{t+1}$ refers to the updated dragonfly position, $g(t)$ determines the dragonfly's current position, and $\nabla g_{t+1}$ indicates the step vector for detecting the movement direction of the dragonfly as:

$$\nabla g_{t+1} = \{l_1 w + l_2 x + l_3 y + F z\} + W \cdot g(t), \tag{13}$$

where $l_1$ indicates the weight of separation ($w$), $l_2$ refers to the weight of alignment $x$, $l_3$ is the weight of cohesion $y$, $F$ is a food vector, $z$ identifies "magnetism" of the food source, $W$ indicates an inactivity weight for managing the convergence optimization, $g(t)$ represents a recent position of the dragonfly at time $t$. Next, the global best dragonfly is determined.

**Algorithm 1.** The pseudocode for of MOLDRO-QoSDCS technique.

**Input**: available channels $C = C_1, C_2, \ldots, C_n$, cognitive nodes $N = N_1, N_2, \ldots, N_n$, primary users PUs, secondary users SUs
1. **For each** $C$
2. Initialize channel populations $A = \{C_1, C_2, \ldots, C_n\}$
3. Initialize opposite population of dragonfly $A'$
4. **For each** $C$ **in** $A$ and $A'$
5. Determine objective functions
6. Measure the fitness $\varphi_F$
7. Combine $A$ and $A'$
8. Order the channel $C_1, C_2, \ldots, C_n$
9. Choose local optimum
10. Calculate $w, x, \gamma, z$
11. **If** $\alpha_{F_a} > \alpha_{F_b}$ **then**
12. Update the position of the dragonfly
13. Obtain global best dragonfly
14. **Else**
15. Go to step 6
16. **End if**
17. **End for**
18. **End for**
19. **End**

**Output**: Identify the channel with optimal QoS

Algorithm 1 shows the computation procedure of MOLDRO-QoSDCS. The current and opposite population of several channels is determined. For each population the fitness function is computed depending on the multiple criteria. Both populations are merged and the order of the dragonfly is set using the fitness score. From that, the local optimum solution is obtained. Subsequently, four different processes are performed. Then, the position of the dragonfly is updated and the optimal solution is detected for improving CRN throughput.

## 4. Performance Simulation

The performance of the proposed MOLDRO-QoSDCS technique was simulated using Matlab with 1000 cognitive radio network nodes and several users during a 10 ms period. The results are compared with a conventional adjustable multi-radio channel hopping (AMCH) protocol and an ML-enabled three-phase TCS generation and allocation scheme. Testing metrics used to evaluate the performance of both the proposed and other existing methods include sensing time, throughput, and energy efficiency.

### *4.1. Sensing Time*

Sensing time determines the channel selection period. In addition, sensing time is referred to as the number of channels sensed by SUs for data transmission. It is estimated as:

$$S_T = T[S_C] \cdot N \tag{14}$$

Simulation results of sensing time for MOLDRO-QoSDCS and methods [1] and [2] versus the number of CRN nodes are shown in Tab. 1. The MOLDRO-QoSDCS technique offers the shortest sensing time for improving the performance of

**Tab. 1.** Sensing time comparison.

| Number of CRN nodes | Sensing time [ms] | | |
|---|---|---|---|
| | ML-enabled three-phase TCS generation and allocation scheme | AMCH protocol | MOLDRO-QoSDCS technique |
| **100** | 8.2 | 5.6 | 3.4 |
| **200** | 10.6 | 8.2 | 5.6 |
| **300** | 13.5 | 11.3 | 8.2 |
| **400** | 14.5 | 12.6 | 9.1 |
| **500** | 18.6 | 15.4 | 12.3 |
| **600** | 21.4 | 18.9 | 15.4 |
| **700** | 27.6 | 24.1 | 20.6 |
| **800** | 29.5 | 26.8 | 22.5 |
| **900** | 34.2 | 31.3 | 26.8 |
| **1000** | 36.4 | 33.5 | 30.4 |

CRN and the results obtained are lower than those of state-of-the-art methods (Fig. 2).

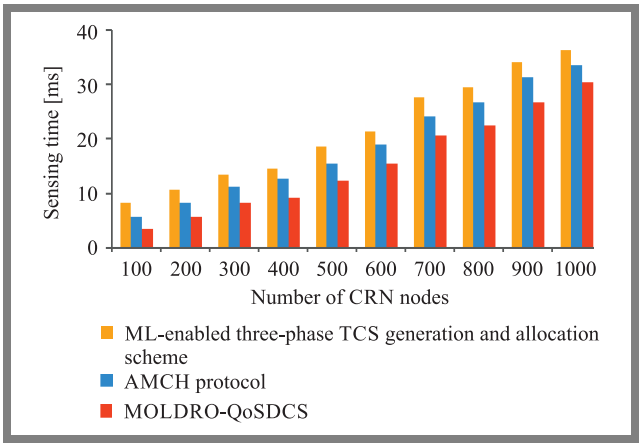



**Fig. 2.** Sensing time versus the number of nodes.

The lower sensing time achieved with the use of the MOLDRO-QoSDCS technique results from the oppositional learning concept. The ranking method detects the local optimum from the population. Contrary to existing approaches, more available optimal channels are sensed and determined in a CRN. Besides, the best channels are determined and ranked to find the local optimal one. This leads to a decrease (of up to 22%) in sensing time during the channel selection process compared to [1] and, to a decrease of 33% compared to [2].

### *4.2. Throughput*

The formula for calculating throughput is:

$$\textit{Throughput} = \frac{\mathrm{DR}}{t}, \tag{15}$$

where DR is the data rate and $t$ is the time interval (considered to be 1 s). It is determined in bits per second (bps) i.e. data transmission speed.

Table 2 shows throughput-related performance of three different methods: MOLDRO-QoSDCS, [1] and [2] for different

CRN nodes. The data obtained shows that the throughput of the proposed technique is higher than that of [1], [2] (by up to 21% compared with [1] and by 35% compared with [2]) (Fig. 3).

**Tab. 2.** Comparison of throughput for researched methods.

| Number of CRN nodes | Throughput [bps] | | |
|---|---|---|---|
| | ML-enabled three-phase TCS generation and allocation scheme | AMCH protocol | MOLDRO-QoSDCS technique |
| **100** | 100 | 120 | 165 |
| **200** | 140 | 160 | 210 |
| **300** | 190 | 220 | 280 |
| **400** | 240 | 270 | 330 |
| **500** | 330 | 360 | 420 |
| **600** | 390 | 410 | 490 |
| **700** | 450 | 480 | 560 |
| **800** | 500 | 550 | 640 |
| **900** | 610 | 650 | 720 |
| **1000** | 690 | 740 | 860 |

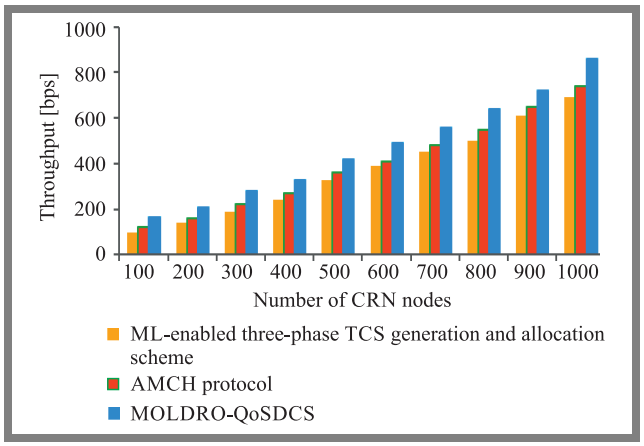



**Fig. 3.** Throughput comparison.

The newly designed optimization technique also finds the global best solution among the population through fitness function based on optimal SNR ratio, power consumption, and spectrum utilization. This additionally helps increase the throughput.

### *4.3. Energy Efficiency*

Energy efficiency is measured as the ratio between output energy and input energy used for communication in a CRN:

$$\mathrm{EE} = \frac{R_{\mathrm{out}}}{E_{\mathrm{in}}} \cdot 100\% \tag{16}$$

where EE denotes the energy efficiency, $E_{\mathrm{out}}$ is the output energy, and $E_{\in}$ is the input energy.

Table 3 shows the energy efficiency simulation results for the three methods, depending on the number of CRN nodes. From the table, it is evident that the proposed MOLDRO-QoSDC technique enhances energy efficiency to a greater extent than the two existing methods.

**Tab. 3.** Comparison of energy efficiency.

| Number of CRN nodes | Energy efficiency [%] | | |
|---|---|---|---|
| | ML-enabled three-phase TCS generation and allocation scheme | AMCH protocol | MOLDRO-QoSDCS technique |
| **100** | 73 | 78 | 85 |
| **200** | 75 | 80 | 87 |
| **300** | 76 | 81 | 88 |
| **400** | 75 | 80 | 86 |
| **500** | 76 | 81 | 87 |
| **600** | 77 | 83 | 89 |
| **700** | 76 | 82 | 88 |
| **800** | 78 | 83 | 90 |
| **900** | 75 | 81 | 89 |
| **1000** | 76 | 80 | 88 |

Figure 4 illustrates the energy efficiency performance depending on the number of CRN nodes (varying from 100 to 1000). The MOLDRO-QoSDC technique allows to achieve energy efficiency of 85% for 100-node networks, while values of 78% and 73% are obtained using [1] and [2]. The average outputs show that the energy efficiency of the MOLDRO-QoSDC technique is increased by 8% compared with [1] and by 16% compared with [2].

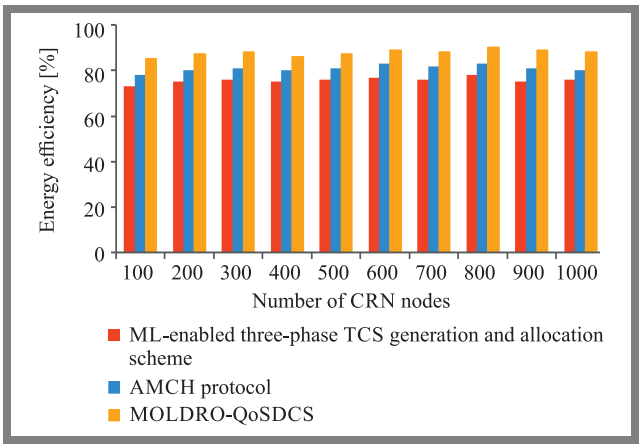



**Fig. 4.** Energy efficiency comparison graph.

## 5. Conclusion

An efficient MOLDRO-QoSDC technique is proposed to improve throughput-related performance of CRNs. MOLDRO-QoSDC increases energy efficiency and reduces the sensing time in CRNs. With the calculation of the fitness function, an optimal channel is chosen for CRN communication. Consequently, throughput is improved as well as the sensing time is shortened.

# References


[1] C. Chao, C. Chen, and H. Huang, "An adjustable channel hopping algorithm for multi-radio cognitive radio networks", *Computer Networks*, vol. 170, pp. 1–7, 2020 (DOI: https://doi.org/10.1016/j.comnet.2020.107107).

[2] T. Chakraborty and I.S. Misra, "A novel three-phase target channel allocation scheme for multi-user cognitive radio networks", *Computer Communications*, vol. 154, no. 4, pp. 19–39, 2020 (DOI: 10.1016/j.comcom.2020.02.026).

[3] J. Tlouyamma and M. Velempini, "Investigative analysis of channel selection algorithms in cooperative spectrum sensing in cognitive radio networks", *SAIEE Africa Research Journal, IEEE*, vol. 112, no. 1, pp. 4–14, 2021 (DOI: 10.23919/SAIEE.2021.9340532).

[4] S. Jang, Ch. Han, K. Lee, and S. Yoo, "Reinforcement learning-based dynamic band and channel selection in cognitive radio ad-hoc networks", *EURASIP Journal on Wireless Communications and Networking*, no. 131, pp. 1–25, 2019 (DOI: 10.1186/s13638-019-1433-1).

[5] R.N. Raj, A. Nayak, and M.S. Kumar, "Spectrum-aware cross-layered routing protocol for cognitive radio ad hoc networks", *Computer Communications*, vol. 164, pp. 249–260, 2020 (DOI: 10.1016/j.comcom.2020.10.011).

[6] R.N. Raj, A. Nayak, and M.S. Kumar, "QoS Aware Routing Protocol for Cognitive Radio Ad Hoc Networks", *Ad Hoc Networks*, vol. 113, pp. 1–22, 2021 (DOI: 10.1016/j.adhoc.2020.102386).

[7] H.B. Salameh, R. Qawasmeh, and A.F. Al-Ajlouni, "Routing with intelligent spectrum assignment in full-duplex cognitive networks under varying channel conditions", *IEEE Communications Letters*, vol. 24, no. 4, pp. 872–876, 2020 (DOI: 10.1109/LCOMM.2020.2968445).

[8] Y. Zhong, H. Wang, and H. Lv, "A cognitive wireless networks access selection algorithm based on MADM", *Ad Hoc Networks*, vol. 109, pp. 1–9, 2020 (DOI: 10.1016/j.adhoc.2020.102286).

[9] K.N.S. Gudihatti, M.S. Roopa, R. Tanuja, S.H. Manjulaa, and K.R. Venugopal, "Energy aware resource allocation and complexity reduction approach for cognitive radio networks using game theory", *Physical Communication*, vol. 42, pp. 1–41, 2020 (DOI: 10.1016/j.phycom.2020.101152).

[10] A. Kaur and K. Kumar, "Intelligent spectrum management based on reinforcement learning schemes in cooperative cognitive radio networks", *Physical Communication*, vol. 43, pp. 1–12, 2020 (DOI: 10.1016/j.phycom.2020.101226).

[11] F. Aghaei and A. Avokh, "MRCSC: A cross-layer algorithm for joint multicast routing, channel selection, scheduling, and call admission control in multi-cell multi-channel multi-radio cognitive radio wireless networks", *Pervasive and Mobile Computing*, vol. 64, pp. 1–20, 2020 (DOI: 10.1016/j.pmcj.2020.101150).

[12] A. Ali, *et al.*, "Adaptive bitrate video transmission over cognitive radio networks using cross layer routing approach", *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 935–945, 2020 (DOI: 10.1109/TCCN.2020.2990673).

[13] T. Chakraborty and J.S.S. Misra, "Design and implementation of a novel two-phase spectrum handoff scheme for QoS aware mobile users in cognitive radio networks", *Computer Networks*, vol. 195, pp. 1–27, 2021 (DOI: 10.1016/j.comnet.2021.108194).

[14] R. Dasari and N. Venkatram, "Discrete quality factors aware channel scheduling in cognitive radio ad hoc networks", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14, 2021 (DOI: 10.1007/s12652-020-02607-6).

[15] A. Bagheri, A. Ebrahimzadeh, and M. Najimi, "Energy-efficient sensor selection for multi channel cooperative spectrum sensing based on game theory", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2020 (DOI: 10.1007/s12652-020-02651-2).

[16] R. Yilmazel and N. Inanç, "A Novel approach for channel allocation In OFDM based cognitive radio technology", *Wireless Pers. Commun.*, vol. 120, pp. 307–321, 2021 (DOI: 10.1007/s11277-021-08456-6).

[17] M.W. Khan and M. Zeeshan, "QoS-based dynamic channel selection algorithm for cognitive radio based smart grid communication network", *Ad Hoc Networks*, vol. 87, pp. 61–75, 2019 (DOI: 10.1016/j.adhoc.2018.11.007).

[18] O.P. Awe, D.A. Babatunde, S. Lambotharan, and B. AsSadhan, "Second order Kalman filtering channel estimation and machine learning methods for spectrum sensing in cognitive radio networks", *Wireless Networks*, vol. 27, pp. 3273–3286, 2021 (DOI: 10.1007/s11276-021-02627-w).

[19] R.V. Awathankar, M.S.S. Rukmini, and R.D. Raut, "To mitigate with trusted channel selection using MOORA algorithm in cognitive radio network", *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, pp. 1–10, 2020 (DOI: 10.1007/s40998-020-00382-w).

[20] V. Rajpoot and V.S. Tripathi, "A novel proactive handoff scheme with CR receiver based target channel selection for cognitive radio network", *Physical Communication*, vol. 36, pp. 1–11, 2019 (DOI: 10.1016/j.phycom.2019.100810).

[21] S.N. Sanka, T.R. Yarram, K.C. Yenumala, K.K. Anumandla, and J.R.K. Kumar Dabbakuti, "Dragonfly algorithm based spectrum assignment for cognitive radio networks", *Materials Today: Proceedings*, pp. 1–4, 2021 (DOI: 10.1016/j.matpr.2020.11.301).

[22] S. Jothiraj, S. Balu, and N. Rangaraj, "An efficient adaptive threshold-based dragonfly optimization model for cooperative spectrum sensing in cognitive radio networks", *International Journal of Communication Systems*, vol. 34, no. 10, pp. 1–11, 2021 (DOI: 10.1002/dac.4829).


---


**Ch.S.N. Sirisha Devi** is pursuing a Ph.D. at KLEF (KL deemed to be University), India. She obtained her B.Tech. degree in Electronics and Communication Engineering from JNTUH in 2005 and M.Tech. in Embedded Systems from JNTUH in 2012. Her research interests are in wireless communications and signal processing.

E-mail: Siricharan163@gmail.com

ECE Department, KL University, KLEF, India



**Suman Maloj** works as Professor & Head of the Department of Electronics and Communication Engineering at KLEF (KL deemed to be University), India. He obtained his B.Tech. degrees in Electronics and Communication Engineering from JNTUA in 2001, M.Tech. in Electronics and Communication Engineering from JNTUH in 2006 and Ph.D. degree in 2014 from KL University. His research interests focus on speech coding, speech compression, as well as speech, and speaker recognition.

ECE Department, KL University, KLEF, India

Paper

# Investigation of Vehicular S-LSTM NOMA Over Time Selective Nakagami-$m$ Fading with Imperfect CSI

Ravi Shankar[1], Bhanu Pratap Chaudhary[2], and Ritesh Kumar Mishra[2]

[1]*Electronics and Communication Engineering Department, SR University, India,*
[2]*Electronics and Communication Engineering Department, National Institute of Technology Patna, India*

https://doi.org/10.26636/jtit.2022.165722

**Abstract — In this paper, the performance of a deep learning-based multiple-input multiple-output (MIMO) non-orthogonal multiple access (NOMA) system is investigated for 5G radio communication networks. We consider independent and identically distributed (i.i.d.) Nakagami-$m$ fading links to prove that when using MIMO with the NOMA system, the outage probability (OP) and end-to-end symbol error rate (SER) improve, even in the presence of imperfect channel state information (CSI) and successive interference cancellation (SIC) errors. Furthermore, the stacked long short-term memory (S-LSTM) algorithm is employed to improve the system's performance, even under time-selective channel conditions and in the presence of terminal's mobility. For vehicular NOMA networks, OP, SER, and ergodic sum rate have been formulated. Simulations show that an S-LSTM-based DL-NOMA receiver outperforms least square (LS) and minimum mean square error (MMSE) receivers. Furthermore, it has been discovered that the performance of the end-to-end system degrades with the growing amount of node mobility, or if CSI knowledge remains poor. Simulated curves are in close agreement with the analytical results.**

***Keywords — inter-symbol interference, MIMO, NOMA, orthogonal frequency division multiplexing (OFDM), S-LSTM, zero-mean circularly symmetric complex Gaussian (ZM-CSCG).***

## 1. Inrtoduction

The market expects that 5G communications will be suitable for providing services with very low latency, excellent quality of service (QoS), and increased mobile broadband [1], [2].

Several technological advances, including 5G radio access networks (5G-RAN), the Internet of Things (IoT), ultra-reliable low-latency communications (URLLC), heterogeneous networks, including small cells and machine-to-machine (M2M) communications, as well as the understanding of the mechanisms of specific methodologies, are important building blocks of the 5G air interface [3], [4]. Vehicular communication (VC) solutions, such as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) schemes, are widely used in safety-, information-, mobility- and environment-related applications. Because of its unique characteristics, VC has evoked the curiosity of academics, businesses, and governments alike. V2V and V2I have been envisioned as one of the most promising schemes for improving traffic management and road safety, because they allow each vehicle to communicate with other users placed at important intersections, such as stop signs and traffic lights. A schematic representation of a V2V network is presented in Fig. 1 [1], [2].

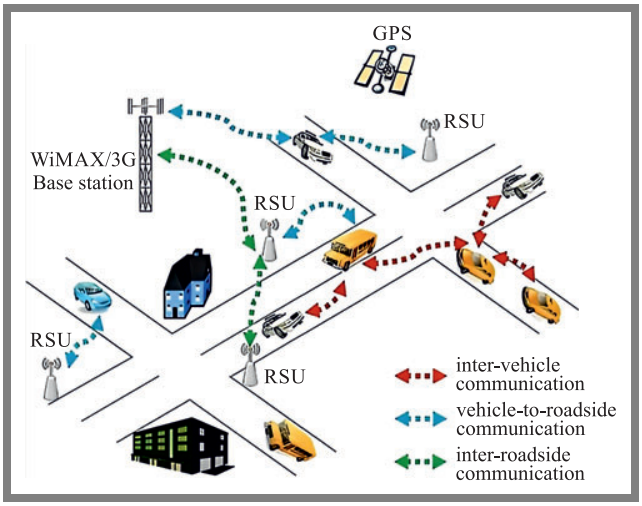



**Fig. 1.** Schematic representation of a V2V scheme.

In addition to being a new resource allocation scheme, 5G wireless networks rely also on a user-centric network concept that seeks to meet the application needs of all users participating in the connected world [5]–[7]. It is almost impossible to develop a unified technology that meets the broad variety of transmission-related needs, due to the sheer number of access schemes and use cases existing in today's interlinked, digital world. Therefore, 5G does not seek to alter the wireless 4G architecture, instead offering a unified platform that makes use of all current and envisioned technologies to serve users by providing them with access to a wide range of services. 5G aims to provide new air interfaces as well as a few new access modes by making relying on a spectrum that has recently been allocated. To be more specific, it will be built on top of contemporary wireless technologies, such as 5G URLLC, long-term evolution advanced (LTE-A), mmWave high-band 5G, lowest signal-to-interference-plus-noise ratio (SINR), and enhanced mobile broadband (eMBB).

It is crucial to outline the fundamental components of 5G in order to create a framework for the co-existence of other technologies. Secure communication in V2V and V2I NOMA networks demands minimal latency, excellent end-to-end reliability, and massive connectivity. Intelligent transportation systems (ITS) will be applied in entertainment applications, such as connected driving and smart transportation, thus increasing the requirement for spectral efficiency (SE) [8]–[11]. By enabling multiplexing in the code and the power domain, NOMA – which significantly improves energy and SE over orthogonal multiple access (OMA) – serves many users concurrently or simultaneously accessing frequency resources [12]. 5G, IoT and multimedia applications place strict demands on capacity and user access. The NOMA scheme [13], [14] offers an optimistic answer to these challenges. MIMO, cognitive cooperative relaying, full-duplex relaying, millimeter-wave, and other technologies have been employed in conjunction with NOMA to increase throughput and guarantee user fairness in a wide range of fading channel distributions [15], [16]. Figure 2 shows a MIMO-NOMA network example [10], [17].

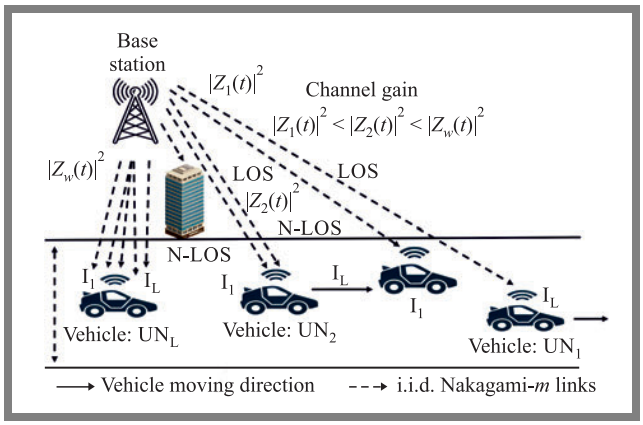



**Fig. 2.** V2V NOMA scenario over Nakagami-$m$ fading channel conditions.

## 2. Related Work

In [18]–[20], the authors examine the millimeter-wave scheme over quasi static frequency flat links. Over the frequency flat $\alpha$-$\eta$-$\kappa$-$\mu$ fading channels, paper [21] evaluates end-to-end SER performance of the relaying network in conjunction with a NOMA scheme for multiple users. In [22], cooperative NOMA's exact OP performance and the achievable data rate are examined by considering an analog relaying protocol over frequency-flat generalized Nakagami-$m$ fading links, and in [23], the authors extend that work by considering imperfect CSI conditions and additive white Gaussian (AWGN) channel noise. It is very difficult to compare channel matrices in MIMO networks. In their work [24], [25], the authors propose the generalized singular value decomposition (GSVD) approach turning MIMO channels into several single-input single-output (SISO) channels [24], [25]. The authors look at how well MIMO-NOMA networks minimize delays when limits concerning transmission time and power are imposed. This study shows that MIMO-NOMA systems offer shorter delays than their OMA-based counterparts.

Since security is one of the paramount issues in wireless networks, the authors of [25] investigate how NOMA and OMA networks compare in terms of ensuring confidentiality of information. The results show that NOMA-MIMO offers a higher secrecy rate. Additionally, the effectiveness of NOMA has been assessed for partial CSI [26]. However, there are several restrictions affecting NOMA, including the need for a perfect CSI in the transmitter and a significant amount of processing complexity in the receiver.

Utilizing deep learning (DL) methods is a great way to overcome these difficulties. The traditional SIC technique has several drawbacks. With more users, it becomes more challenging to utilize the SIC approach to accurately interpret the data. Propagation inaccuracy also has an impact on the SIC approach. Signal categorization may be used by deep neural networks (DNN) to recover a discrete sequence from a degraded signal. The SIC approach can be improved with DL as well. The authors of [27] study a beamforming-based downlink NOMA network that uses dynamic user pairing to maximize the lowest data rate of all DL users. The authors resolve the issue with non-convex optimization with mixed-integer variables, by converting a discrete domain to the analog domain using an iterative scheme depending on the internal approximation to obtain a local optimum. The suggested technique outperforms standard beam generation, NOMA with random pairing, and heuristic search schemes.

DL is a popular channel estimation approach used in 5G networks. Traditionally, channel estimation and signal detection processes are separated. Before signal detection, a pilot broadcast estimates CSI. The receiver may reconstruct the transmitted signals using the calculated CSI. End-to-end OFDM channel analysis is discussed in detail in [28], describing a DL-based technique for joint channel estimation and detection in OFDM systems. Training a DL model involves matching the received signals with the data and pilots provided. After training, the model can decode online-transmitted data without channel estimations. Furthermore, the authors analyze a DL-based single-input multiple-output (SIMO) channel estimate. By using piecewise linearity, the DL channel detector might be able to estimate a large group of functions, since a DNN with a rectified linear unit (ReLU) activation function is conceptually similar to a piecewise linear function. An alternative method for estimating the channel using the DL scheme is presented by the authors in [29]. The authors provide an alternate method for DL channel estimation. MMSE is a model-specific estimator for conditionally normal channels. Using MMSE, CNN estimators are defined.

Due to the high processing complexity and substantial variations in wireless channel conditions, current NOMA systems make it difficult to evaluate channel characteristics and establish the appropriate resource allocation strategy. In [30], the authors present a DL-aided NOMA system, where a single BS serves random NOMA users. DL learns time-selective fading NOMA systems with poor CSI. DL-LSTM networks

in NOMA systems provide automatic channel property identification. NOMA employs DL for automatic encoding, decoding, and channel identification in AWGN noise. The suggested method provides more accurate simulations. Faster than Nyquist (FTN)-NOMA is proposed in [31] to improve SE, reduce latency, and increase connectivity. The authors define sliding-window detection for DL-based FTN NOMA. This scheme's detector exceeds MMSE-decision feedback equalization (FDE).

In [32], the authors suggest DL-SIC for downlink (D/L) MIMO-NOMA. This method minimizes MMSE in MIMO-precoding NOMAs and SIC decoding. The transmitter precodes signals for multiple D/L users using superposition coding, and the receiver extracts them using DNN-built SIC decoders. In each of the SIC phases, one DNN is used to decode the signal for the current user, while another is used to reconstruct the signal for the prior user (apart from the first user). Since the user was not previously decoded, just one DNN is required to decode the signal. In [33], the authors establish a DL approach for recognizing D/L signals using the MIMO-NOMA scheme. DNN is the sole SIC receiver used to decode D/L signals broadcast in a single time slot. All aerials provide DNN with data. Most DNNs have a fully linked output layer that uses the Softmax function. Conversely, MIMO-NOMA signals, although emanating from several different transmitters, need a single decoding slot only. Therefore, the proposed output layer is constructed for data classification purposes. Each cluster has as many neurons, as there are aerials for sending signals, and each neuron in a cluster may only be capable of encoding a single state. Different power allocations and complex modulation forms degrade energy efficiency of the proposed system.

Using DL schemes such as CNNs, the authors of [34] reconstruct uplink users' MIMO channel signals. Data from several users may potentially be decoded instantly using the suggested technique which does not rely on any of the typical processes found in communication signal processing. Each step of the SIC decoding procedure in [35] employs a DNN with fully linked layers to decode input from a single user. DNN is composed of four layers: an input layer, an output layer, two hidden layers, and an intermediate hidden layer. The decoded bits that are shown to the user are generated by the DNN's output layer which is responsible for that function. After the first stage of SIC, the input layer receives the combined signal in addition to the signals that have already been decoded by the users. Signals must be transformed into bit sequences before digital neural networks (NNs) (also known as DNNs) may be trained to interpret them. The method provided in [35] to enhance DL-based SIC systems for higher order modulation will be shown in Section 3.

### *2.1. Contribution of the Paper*

DL is a popular approach relied upon for boosting error performance of MIMO-NOMA end-to-end systems. DL-based SIC uses a NN to estimate, identify, decode, and reject channels. DL-based SIC outperforms MIMO NOMA in Nakagami-$m$ fading channels. We apply the DL-based SIC from [35] to an uplink MIMO-NOMA system with BPSK modulation in a time selective Nakagami-$m$ fading channel conditions, and then:

– study the DL-based MIMO NOMA VC system over Nakagami-$m$ fading channel users and compare it with the conventional SIC-based NOMA system,
– study the DL schemes for 5G and beyond 5G communications, with a focus on the VC scenario and the imperfect CSI and conduct a comprehensive literature research on DL,
– for different values of the shape parameter, training data rates, and packet sizes, the traditional NOMA receiver's performance is compared with that of the S-LSTM-based NOMA receiver,
– to decode user signals and ensure performance with current conventional systems, it is necessary to create a DL-SIC MIMO-NOMA model for higher-order complex modulation schemes using the newly suggested architecture.

The paper is organized as follows. Assuming that CSI is imperfect and that a SIC error exists, we examine the signal and channel models in Section 3. The node mobility scenario is considered in this section, and the effects of channel estimation inaccuracy are investigated. We analyze the DL-based NOMA receiver in Section 4 and present the S-LSTM approach for enhancing both OP performance and end-to-end SER. Optimal power allocation (OPA) is determined because of a derivation for the ideal power allocation factor. Additionally, training for the S-LSTM model is provided. We discuss the simulation's outcomes in Section 5, and we conclude the study in Section 6.

## 3. System and Channel Model

### *3.1. Channel Model*

As a result of multipath propagation and node mobility, fading links transform from being frequency-flat to time-selective. ISI is caused by time-varying channels and 5G-OFDM uses a cyclic prefix (CP) to reduce it [36]–[42]. In this work, the MIMO technique is employed by the base station (BS) in order to connect with $W$ vehicular nodes (VNs) in NOMA-based V2V communication networks [10]. As the VN moves away from the BS, the fading channel profile changes from the near Rician-fading channel to a Rayleigh channel. With the help of singular value decomposition (SVD) techniques, the channels are converted to parallel fading links. To take advantage of the diversity at the VNs, the receiver uses maximum ratio combining (MRC), selection combining (SC), or SVD schemes. However, because the diversity of connections is non-identical under high fading conditions, the predicted diversity increases when the independent but non-identically distributed (i.n.i.d.) assumption is not realized [10], [30]. The i.i.d. assumption adds complexity to the study, but the conclusions are more realistic and relevant in real-time communication and V2V network system design [38],

[42]. Here, we analyze the performance of an S-LSTM-based V2V NOMA networks over i.i.d. fading connections while accounting for inaccurate CSI, and we compare findings with the existing literature. This section also examines wireless D/L V2I communications and considers the time selective Nakagami-$m$ model.

Mobile users communicating with a roadside BS move away from that BS at a comparable $v$ speed. Each VN has $T$ receive (Rx) antennas, whereas the BS has $K$ transmit (Tx) antennas. Doppler spread is experienced due to the relative motion between the VNs and the BS. i.i.d. Nakagami-$m$ fading links can be modeled by using the classical Jakes model and the first order autoregressive model. The fading channel coefficients between the BS and $w$ vehicle at $t$ time interval are expressed as [10], [17]:

$$z_w(w) = \rho_w^{t-1} z_w(1) + \sqrt{1-\rho_w^2} \sum_{l=1}^{t-1} \rho_w(t-l-1) e_w(l), \quad (1)$$

where $\rho_w = J_0 \frac{2 \cdot \pi \cdot f_c \cdot \nu}{R_s c}$ is the correlation factor for the gain of the temporally neighboring channel, the carrier frequency is denoted by $f_c$, $c$ is the light speed [4]–[8], [33], and transmission symbol rate is represented by $R_s$. The zeroth-order Bessel function of the first kind is defined as $J_0(:)$. The time selective fading component is represented as $e_w(l)$ which is modeled as the ZM-CSCG with variance $\sigma_{ew}^2$.

According to the working theory, destination tracking loops are unable to keep up with time-varying channel variances and can only estimate them during the first signaling period of each transmitted block, because node mobility causes the fading channel coefficients to change for each time instant and transform a quasi-static Nakagami-$m$ fading channel into a time selective fading channel. This is because the quantity of information that may be communicated is limited by the time selected Nakagami-$m$ fading connections, i.e. $z_w(1)$ as $\hat{z}_w(1)$. Furthermore, ISI and channel estimation error $\hat{z}_w(1)$ can be expressed as:

$$\hat{z}_w(1) = z_w(1) + \hat{z}_{\epsilon w}(1),$$

where $\hat{z}_{\epsilon w}(1)$ is the fading channel error coefficient distributed as ZM-CSCG with variance $\sigma_\epsilon^2$ modeled as, $\mathbb{CN}(0, \sigma_e^2)$. The small-scale fading channel is affected by the path loss. The estimated value of the fading channel coefficient $\hat{z}_w(1)$ is affected by the path loss exponent $\epsilon$, as a function of distance $d_w$ between BS and mobile user $w$ . At $t = 1$, the gain of the fading link is given as, $\hat{z}_w(1)|^2$ [5], [6], [17], [33]–[35]. The fading channel link gain is inversely proportional to the distance from the BS to mobile user $w$. With the increasing distance between the BS and the user, channel gain decreases. At $t = 1$, the channel gain is:

$$|\hat{z}_1(1)|^2 \leqslant |\hat{z}_2(1)|^2 \leqslant \ldots \leqslant |\hat{z}_W(1)|^2.$$

The BS uses superposition coding and power domain NOMA to correlate the signals for various vehicles in the power domain. When considering the channel order at time $t$, the power coefficients assigned to vehicles are expressed in the following order $\beta_1(t) \geqslant \beta_2(t) \geqslant \ldots \geqslant \beta_W(t)$ [3]–[7], [33].

### *3.2. System Model*

The use of DL to increase the performance of MIMO-NOMA systems has recently gained in popularity. SIC-related research based on DL uses a neural network (NN) to conduct channel estimation, detection, decoding, and discarding of the decoded signal. A DL-based SIC receiver performs better than MIMO-NOMA under Nakagami-$m$ fading channel conditions. Hence, under time-selected Nakagami-$m$ fading channel conditions, we apply a modified version of the DL-based SIC, as suggested in [35], to the uplink MIMO-NOMA system that uses BPSK modulation. After analyzing DL-based SIC in [35], the system model is formalized and DNN is applied in the SIC receiver.

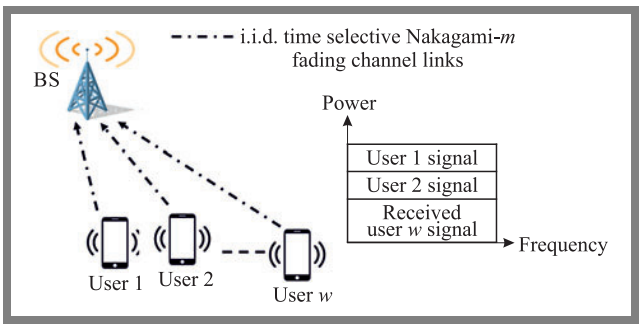



**Fig. 3.** Schematic representation of uplink in a MIMO NOMA (single cell) network.

Figure 3 illustrates the single cell uplink MIMO-NOMA scenario over time selective Nakagami-m fading channel conditions. The receiver has $T$ antennas, and $W$ represents the total number of users. The schematic diagram demonstrates that users are broadcasting their signals to the BS utilizing the same frequency resources, but at varying degrees of transmission power. According to the findings from [35]–[43], the users must transmit many frames which include both pilot and data symbols. When the frames are sent within a coherent period, the channel impulse response is assumed to be constant across the span of a single frame. Each frame has $N$ data symbols and $J$ pilot symbols. Symbols with the letter $L$ in front of them are considered pilot symbols, whereas those with the $D$ letter denote data symbols. Inspired by the analysis in [35], all users' data signals are expressed as [35], [43]–[45]:

$$\mathbf{R} = \mathbf{ZPS}^D + \mathbf{\Psi}^D, \quad (2)$$

where $S^D$ represents the data matrix as:

$$S^D = [S_1^D, S_1^D, \ldots, S_W^D]^{\mathrm{T}} \in \mathbb{C}^{W \times N},$$

where T is the transpose operator, $W$ represents the total number of users, while the total number of receive antennas is $T$. $S_w^D$ denotes the data vector of $w$-th user complex modulated data symbol and can be given as [35], [44], $S_w^D = [S_{w,1}^D, S_{w,2}^D, \ldots, S_{w,N}^D]^{\mathrm{T}} \in \mathbb{C}^{1 \times N}$ and $E(|S_{w,N}^D|^2) = 1$. The fading channel matrix $\mathbf{Z}$ is represented as [30]–[35]:

$$\mathbf{Z} = [z_1, z_2, \ldots, z_W] \in \mathbb{C}^{T \times W},$$

where $z_w = [z_{w,1}, z_{w,2}, \ldots, z_{w,T}] \in \mathbb{C}^{T \times 1}$ is the channel vector between $w$ users. Considering the time selective Nakagami-$m$ fading channel conditions, let the diag-

onal matrix $P$ be represented as [35], [43], [44], [46] $P = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_w)$. $P$ represents the total available transmit power and is equal to $\sum_{w=1}^{W} \lambda_w \leqslant P$. $\Psi^D$ is the AWGN noise at the BS modeled as $\Psi^D \sim \mathbb{CN}(0, N_0 I) \in \mathbb{C}^{T \times N}$. The same approach can be used to represent all users' pilot signals at any frame [34], [35], [43], [44], [46], [47]:

$$\mathbf{R}^L = \mathbf{ZPS}^L + \mathbf{\Psi}^L \tag{3}$$

with energy of the symbol equaling:

$$E(|S_{w,n}^L|^2) = 1.$$

$S^L$ stands for the pilot matrix and can be expressed as [44]:

$$S^L = [S_1^L, S_2^L, \ldots, S_W^L]^{\mathrm{T}} \in \mathbb{C}^{W \times J}.$$

Channel noise $\Psi^L \in \mathbb{C}^{T \times J}$ is assumed to be AWGN noise for pilot symbol transmissions at the BS. In articles [35], [43], [44] the authors have considered the equal power allocation scheme. Paper [46] analyzes mMTC with the assumption that all mobile users have access to the same bandwidth. There are several drawbacks to the equitable distribution of power. The user should be allocated less power if they are close to the BS, since the channel gain between them is higher than it would be otherwise. Due to its poor channel gain, the distant user needs a larger share of the available power. Using the dynamic power distribution technique in [35], [43], [44], [46], [47], the authors provide more power to the user with the lowest gain, while giving less power to the user with the highest gain. Availability of power at the BS is limited - let it be represented by $P$. The total available power is distributed among $W$ users. The power allocation factor $\lambda_w$ is expressed as [30]–[34]:

$$\lambda_w = \frac{P}{W} \quad \forall\, w \in \{1, 2, 3, \ldots, W\}.$$

Let user 1 be nearest the BS user $w$ is the farthest user. Channel gain decreases with distance, as $||z_1|| \geqslant ||z_1|| \geqslant \ldots \geqslant ||z_W||$. Detecting the received signals involves recovering the signal at the receiver. A standard MIMO-NOMA receiver must perform channel estimation, signal detection, and demodulation before extracting the signal. The operation of a conventional SIC receiver is shown in Fig. 4. The strength of the signals provided by users will be used to discern their meanings. To do this, we start by decoding the user who is experiencing optimal channel circumstances, and then we deduct their signal from the composite signal. Then, the user with the second-highest signal strength will have their code decrypted. The process will be repeated until the individual user whose signal strength is the weakest can be deciphered.

Operation of SIC is determined by the user count. $S_w^L$ pilot symbols allow the channel to be estimated during the pilot broadcast. Using zero forcing (ZF)-SIC and MMSE SIC, MIMO-NOMA can successfully cancel interference. The data vector (estimated value) for mobile user $w$ can be obtained by employing the ZF-SIC receiver as [21], [33]–[35], [44], [46]:

$$\hat{S}_w^D = Z_w \mathbf{R}^{\mathbf{D}}, \tag{4}$$

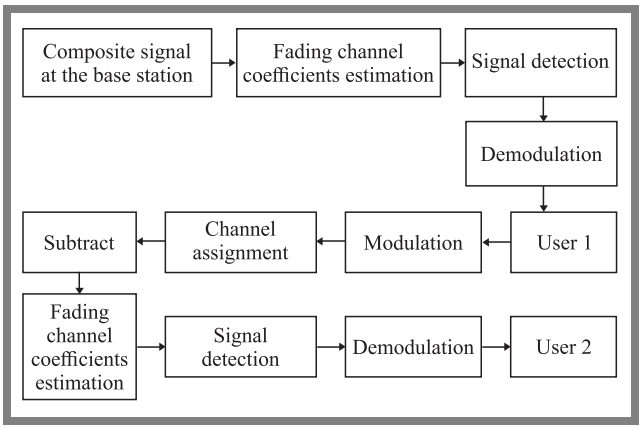



**Fig. 4.** Schematic representation of a SIC-based receiver.

where $\hat{z}_w$ is the channel estimation vector of user $w$, as:

$$\hat{z}_w = [\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_T]^{\mathrm{T}} \in \mathbb{C}^{T \times 1} \text{ and}$$
$$Z_w = \hat{z}_w^H (\hat{z}_w \hat{z}_w^H + I)^{-1}.$$

Next, the $(w+1)$-th user's signal is decoded after the decoded signal is subtracted from the received signal. Similarly, the detection for user $w$ using MMSE-SIC may be written as [32], [35], [37], [43], [44], [46]–[48]:

$$\hat{S}_w^D = W_w \mathbf{R}^{\mathbf{D}}. \tag{5}$$

where $W_w = \hat{z}_w^H (\hat{z}_w \hat{z}_w^H + \rho^{-1} I)^{-1}$ and $\rho$ is the received SNR given by [30], [33]–[35], [43], [44]:

$$\rho = \frac{P_r}{N_0}. \tag{6}$$

Also, the SINR of user $w$ is given as [30], [33]–[35], [43], [44]:

$$\text{SINR}_w = \frac{\lambda_w \rho |z_w|^2}{\sum_{j=1}^{w-1} \lambda_j \rho |z_j|^2 + 1} (w \neq 1). \tag{7}$$

Furthermore, the SINR calculation for user 1 is [30], [33]–[35], [43]–[46]:

$$\text{SINR}_1 = \lambda_1 \rho |z_1|^2. \tag{8}$$

To communicate with the SIC receiver, the DNN protocol has been established. According to the results of the preceding research, a standard SIC receiver starts by making an estimate of the CSI based on the transmitted symbols that match the pilot symbols. The estimated CSI may then be used for reconstruction of the received signal. The DNN, in contrast to the DL-based SIC technique, is trained during the pilot transmission and then utilized to recover the transmitted bits without explicitly calculating the channel state or discarding the decoded signal. This is possible because the DL-based SIC approach makes use of DL.

To decode the data from a single user, a DNN with layers that are completely linked is used at each stage of SIC. In addition to input and output layers, the DNN design includes two hidden layers. Except for the first SIC step, the DNN output layer decodes bits for each user, whereas the input layer receives the composite signal and previously decoded signals. DNNs are trained to convert the input data matrix into bit sequences for the broadcast. The user with the best

channel condition would have their data deciphered first, as in the traditional SIC system. Then, in SIC, the second-strongest user is decoded. In this manner, the procedure is repeated to decode the user at the lowest level. Modulation order C is used to determine which nodes of each DNN will be used as outputs. BPSK requires two nodes, whereas BPSK needs four. Figure 5 shows a two-user DL-based SIC architecture. After each DNN SIC, the next modulation block modifies the decoded user bits. During SIC, the user's modulated symbols are delivered to DNN.

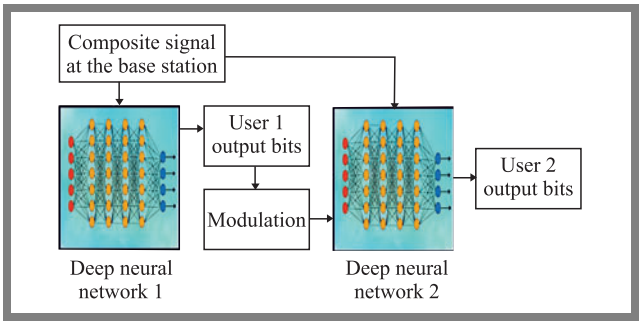



**Fig. 5.** Schematic representation of a DL-based SIC receiver.

DNN is composed of $Q$ fully linked layers, and the Softmax function is used on the output layer of each DNN. Firstly, we consider the input vector $\mathbf{s} = \{s_1, s_2, \ldots, s_c\} \in \mathbb{R}^{C\times 1}$ and we can find a way to express the output vector $\Phi = \{\Phi_1, \Phi_2, \ldots, \Phi_C\}$ using real number values in the range of 0 to 1 that add up to 1 [48]. The Adam algorithm optimizes performance by decreasing the categorical cross-entropy loss function between output and training objectives. The loss function is given as [32], [35], [43]–[48],

$$\text{loss} = \sum_{j=1}^{J}\sum_{c=1}^{C} B_{jc} \log\left(\frac{\Phi_{jc}}{2}\right), \tag{9}$$

where $B_{jc}$ is a binary ground truth indicator such that $B_{jc} = 1$ only holds if and only if the $j$-th sample belongs to the $c$-th class. The Softmax's output probability that the $j$-th input corresponds to the $c$-th class is represented by the $\Phi_{jc}$ symbol. Consider $f_w(.)$ to be the $w$-th DNN's processing component, with function $f_w(.)$ defined as $f_w(.) : \mathbb{R}^{(T+w-1)} \rightarrow \{0,1\}^{\log_2(c\times 2)}$, which maps the combined signal and the previously decoded signal of the $(w-1)$-th users to the broadcast bit of user $w$ [49]–[53]. This is possible because all users' complex signals can be split into real and imaginary parts. To estimate the function of the $w$-th DNN, the weight and bias matrix are changed after pilot symbol training. Specifically, ELU and ReLU are used as activation functions for the first and second hidden layers.

## 4. DL Based NOMA Receiver

### *4.1. S-LSTM Basics*

LSTM cells are presented in Fig. 6 [31]–[33]. Recurrent neural networks (RNNs) are known as LSTM networks, and they can learn the long-term associations that exist between sequence time steps. An LSTM network is made up of several different layers, two of which are the sequence input layer and the LSTM layer. A sequence input layer is a component of the network that is responsible for the transmission of data to the network in the form of sequences or time series. An LSTM layer will learn the long-term connections that exist between the many time steps that make up the sequence data. The LSTM architecture is shown in Fig. 7 [32], [33]. The sequence input layer is followed by the LSTM layer. Three layers are used to forecast class labels: a fully connected layer, a Softmax layer, and a classification output layer. Figure 8 [32], [33] shows a basic regression LSTM network. The first two network layers are sequence input and LSTM. The network's last layers are completely linked and regressed.

Figure 9 shows the video classification network architecture. The image sequences could be networked with the help of a sequence input layer. To independently apply convolutional processes to individual video frames, each frame needs to include a sequence folding layer, followed by convolutional layers, and finally a sequence unfolding layer. To use LSTM layers to learn from vector sequences, we must first deploy a flattened layer – Fig. 9 [32]. In its most basic form, an LSTM model comprises a single hidden LSTM layer followed by a feed-forward output layer. S-LSTM is a variant of this paradigm that includes several hidden LSTM layers, each with numerous memory cells. As more and more LSTM hidden layers are stacked, the model becomes more complex, and the technique becomes more acceptable as DL. The hierarchy provided by the DNNs' several layers is frequently credited with its effectiveness. Each layer solves a smaller subproblem before passing the solution to the next layer. The DNN may be thought of as a processing pipeline, with each layer performing a specific task and passing the data on to the next layer to be processed.

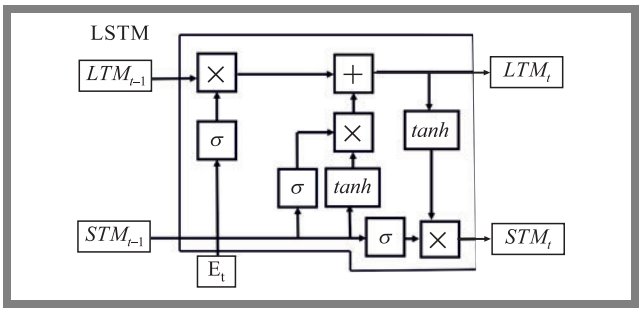



**Fig. 6.** Schematic representation of an LSTM cell.

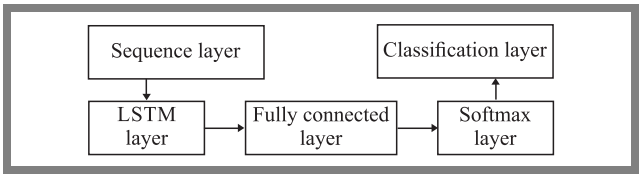



**Fig. 7.** Schematic representation of a simple LSTM network for classification.

Adding extra hidden layers to an NN multilayer perceptron makes it more complex. High degrees of abstraction can be achieved by combining the learnt representation from pre-

vious layers with the additional hidden layers. Lines, forms, and objects are just a few examples here. A single, suitably deep hidden layer. Most functions may be approximated using multilayer perceptrons. An alternative approach that uses fewer neurons and trains more quickly involves deepening the network. After all, depth optimization is a form of representational optimization.

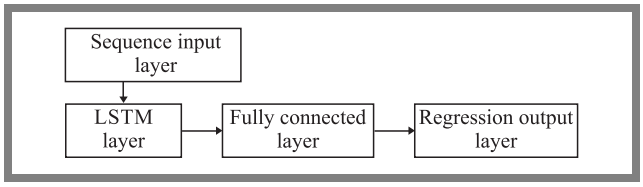



**Fig. 8.** Schematic representation of an LSTM network for regression.

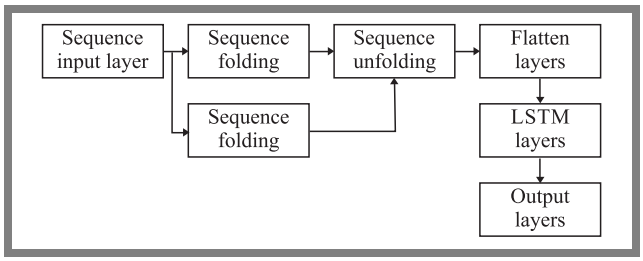



**Fig. 9.** Schematic representation of a video classification network.

In [25], the authors proposed S-LSTM or deep LSTMs for their speech recognition application. The authors achieved this by solving a difficult standard problem. RNNs are deep in time, because they use hidden states. This paper examines the question of whether RNNs may also take advantage of the benefits of depth in space, i.e. the practice of placing numerous, recurrent hidden layers one atop the other, the way that feedforward layers are built in conventional deep networks. Layered LSTMs have recently solved difficult sequence prediction tasks. S-LSTM architecture is a name given to an LSTM model that is made up of multiple, different layers of LSTMs. A higher-level LSTM transmits values to a lower-level LSTM. Each input time step should have a separate output time step (Fig. 10), instead just of one such step. It is necessary to generate one output time step for each input time step, and to generate one output for each input [29].

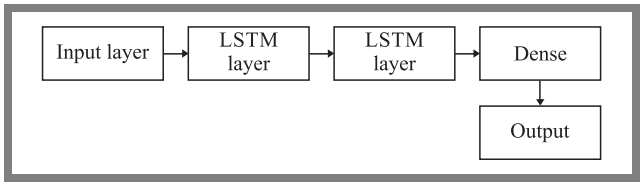



**Fig. 10.** LSTM architecture.

### *4.2. Model Training*

OFDM data packets consist of 92 carriers, and each packet consists of four OFDM symbols. Three pilots are assigned to detect fading channel coefficients. The number of bits per carrier is 4, because channel detection is a very complex task and more bits are assigned per subcarrier. After the carriers have been assigned, the next step is to create a feature vector (FV) and, despite the fact that data symbols are of a complicated nature in the training stage, this step must be completed before moving on. The symbols include both real and imaginary components in their construction. The number of subcarriers influences the size of the FV in a significant way. The FV dimension may be expressed as $92\times4\times2 = 736$. The S-LSTM NOMA channel estimator acquires the ability to comprehend the signal connected to the $k$-th subcarrier after having the key labels included in the training process. A label is a numerical representation of a signal sent by two users together. There will be 30 combinations/labels, since BPSK symbols are being transmitted by both users. DNNs are created with Python and Matlab by attaching DL layers to the DL Toolbox and GPU accelerator. Tensor flow and the Sigmoid activation function are employed in the receiver analysis. A fully linked layer comes after the S-LSTM layer. This layer has an output size of 30 bits and contains 280 hidden units. The classification layer is responsible for generating putative labels that will be used to map the signals that are being concurrently sent by both users, and the Softmax layer is the one that will apply a Softmax function to the input.

## 5. Simulation Results

In this section, performance comparisons have been presented between S-LSTM and other conventional NOMA receivers, in terms of node mobility, SIC error, and imperfect CSI. Further, the S-LSTM NOMA receiver is trained using simulation data and the end-to-end SER and OP are examined. Performance of the system is investigated for each subcarrier, over the various SNR regimes. The effect of ISI and Doppler spread is thought to be reduced in both online training and offline training by assuming that the fading links are time-selective and there is imperfect CSI. All i.i.d. fading links receive a distinct random face shift from each OFDM packet, enabling the analysis of time-selective fading channel conditions. This proposed scheme is trained using 550,000 OFDM samples over 280 epochs, while considering all real-time propagation scenarios, i.e. those involving SIC error and imperfect CSI. The accuracy of S-LSTM-based receivers is evaluated using optimal or maximum probability receivers. In simulations, the carrier frequency is 25 GHz, CP duration is 20 and 25 s, the number of subcarriers is 92, and the number of multiple paths is 60, 70, 80, and 90. BPSK constellations have been used, and the maximum delay spread is set to 25 $\mu$s. Table 1 shows the simulation parameters.

### *5.1. OP Performance in Time-selective Fading Channels with VNs and Imperfect CSI*

OP performance and average SER performance of the different modulation schemes for NOMA-based 5G V2V networks have been compared to validate the analytical findings given in the preceding sections. Consider a D/L V2V scenario in which all VNs are moving away from the BS at 130 km/h. The BS establishes connections with $VN_1$, $VN_2$, $VN_3$ and

$VN_4$ using the carrier frequency of $f_c = 25$ GHz and the transmit symbol rate of $R_s = 20$ Mbps. Assume that $VN_1$ is the farthest user, and that it has a weak fading channel gain. The channel gain also decreases significantly when the $VN_1$ moves away from the BS. For $VN_2$, the user closest to the BS, channel status fluctuates or is inversely proportional to distance and node mobility in the simulations. Channel coefficients change very fast, and gain is highly reduced as node mobility rises. For 25 GHz and 30 GHz frequency bands, the path loss exponent values are 1.64 and 3.10, respectively. The power allocation factors in the simulations for $VN_1$, $VN_2$, $VN_3$, and $VN_4$ are 0.60, 0.25, and 0.10, 0.05, respectively. The dynamic power allocation approach is employed. The target data transmission rate is fixed at 2 bps/Hz, yielding the threshold SNRs of 2, 4, 6, and 8 for $VN_1$, $VN_2$, $VN_3$, and $VN_4$, respectively. An autoregressive process with a variation of 0.01% at a certain point in time may be used to simulate a time-selective fading channel.

**Tab. 1.** Simulation parameters.

| Simulation settings | Type or value |
|---|---|
| Total number of vehicular users | 10 |
| Number of Tx/Tr | 20/10 |
| DL optimization scheme | Stochastic gradient descent |
| Training rate | 0.004 |
| Number of complex modulated symbols | 5000 |
| Number of hidden layers | 280 |
| Hidden layer activation function | Parametrized ReLU |
| Number of training epochs | 280 |
| OPA factor | 0.75, 0.25 |
| Complex modulation scheme | BPSK |
| Fading channel used in simulation | Time selective i.i.d. Nakagami-$m$ fading links |

**Algorithm 1.** MIMO-NOMA DL-based training algorithm

1. Initialization of the DL model
2. Parameter initialization
3. Data normalization
4. Generation of training data symbols and formatting of the data symbols. In simulations, the total number of time slots is $Ts$ and the data vector is: $S^L = [S_1^L, S_2^L, \ldots, S_W^L]^{\mathrm{T}} \in \mathbb{C}^{W \times J}$. The $l$-th slot time slot data is $S^{[l-\mathrm{th}]}$.
5. Set the hidden layer and output layer's important settings, including epochs, data training rate, dropout, and output functions
6. Assigning bias and weight to DNN layers
7. Calculation of the output data vector
8. Calculation of the loss function as: $\mathrm{Loss} = -\sum_{j=1}^{J} - \sum_{c=1}^{C} -B_{jc} \log(\frac{\Phi_{jc}}{2})$
9. Use the stochastic gradient descent technique to compute the correction parameter and to update the simulation settings to seek OPA factors that provide the minimum SER and OP
10. If the loss function is not satisfactory, then recalculate the loss function
11. Using the test data, evaluate the trained DNN and generate OP vs. SNR and OP vs. SNR plots

Figure 11b shows the user's OP performance for S-LSTM MIMO-NOMA and MIMO-OMA for *m*=3. When compared to $m = 2$, the performance is better, since the diversity of benefit for $m = 3$ is better (severity of fading decreases with an increase in the value of the fading severity parameter). It can be readily seen that MIMO-NOMA outperforms MIMO-OMA when $m = 3$, by a difference of 4 dB. However, the performance decline brought on by i.i.d. consideration is more pronounced when $m = 3$, as opposed to $m = 2$. The influence of i.i.d. consideration is shown to decrease under non-line of sight conditions. Because in non-line of sight communication, the channel is no longer Rayleigh faded,

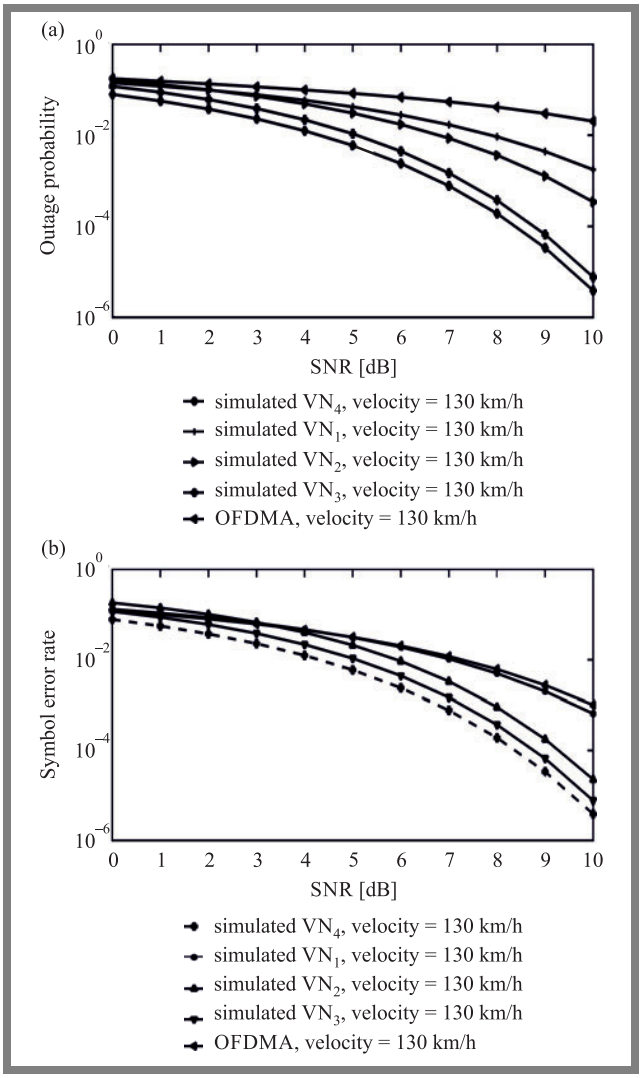



**Fig. 11.** OP vs. SNR for MIMO NOMA considering node mobility and imperfect CSI: a) $m = 2$ and b) $m = 3$.

and in this simulation, we have considered the time selective Nakagami-$m$, with more generalized fading. OP performance for $m = 4$ and 8 for the maximal ratio combining (MRC) NOMA and OMA is shown in Fig. 12a–b. Because $m = 4$ has more severe fading, $m = 8$ has better OP performance for $1 \times 4$ and $1 \times 8$ than $m = 4$. It can be readily seen that the simulation results differ from the analytical results if the likelihood of error is less than $10^{-4}$. This is because the simulation's sample size was practically limited owing to computer-related restrictions.

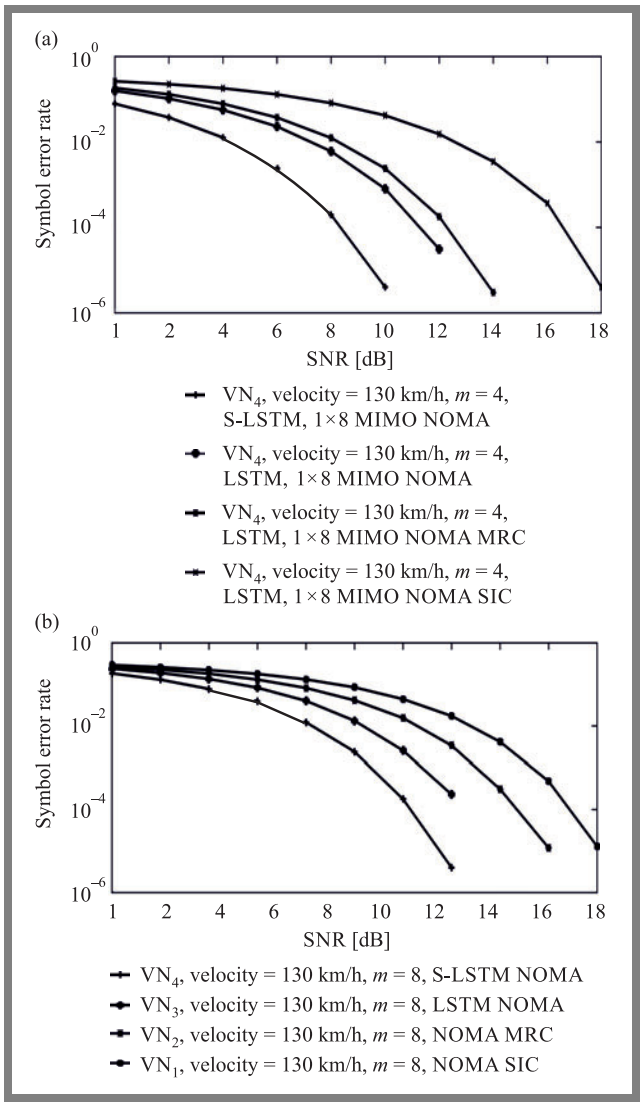



**Fig. 12.** OP vs. SNR for $1 \times 8$ and $1 \times 4$ MIMO NOMA for: a) $m = 4$ and b) $m = 8$, considering node mobility and imperfect CSI and error variance equal to $\sigma_\epsilon^2 = 0.02$.

Figure 13 shows a MIMO-NOMA with $4 \times 2$ and $8 \times 4$ antenna configurations, with imperfect CSI. The simulation results show how imperfect CSI error impacts the functionality of $UN_2$ and $UN_3$. $UN_2$ is affected by interference from $UN_1$, whereas $UN_3$ is affected by interference from both $UN_1$ and $UN_2$. As a result of an increase in SIC error, it is seen that $UN_3$ exhibits a greater rate of performance deterioration than $UN_2$. For both vehicles, $8 \times 4$ MIMO performs better than MIMO because of its higher diversity gain. Simulations of the net throughput and sum-rate between two VNs are performed for SISO-NOMA, SIMO-NOMA, and MIMO-NOMA. The power coefficients for $UN_1$ and $UN_2$ are $\beta_1 = 0.70$ and $\beta_2 = 0.30$, respectively, while the i.i.d. channel has a decaying factor of 0.40. It is assumed that the receiver has access to the ideal CSI. MMSE channel estimation also shows the sum rate performance. Periodically, the BS transmits pilot signals for the purpose of channel estimation. Even if the velocity is higher in the case of S-LSTM when compared to the S-LSTM scheme, SER performance is better compared to the LSTM NOMA scheme. This is something that can be clearly recognized.

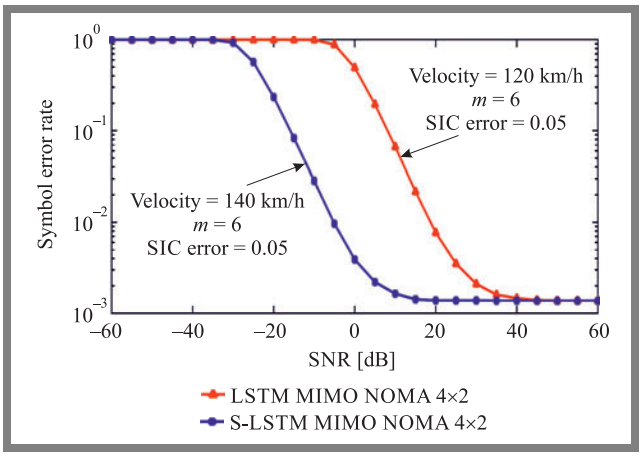



**Fig. 13.** Performance comparison between the $8 \times 4$ and $4 \times 2$ MIMO NOMA system over time selective Nakagami-$m$ fading channel considering SIC error.

### *5.2. Effect of the Number of Pilot Symbols and Clipping Noise*

As shown in Fig. 14, both MMSE and LS approaches may produce accurate estimates and detection rates when 120 pilots are used. Standard MMSE, LS, and SIC-based receivers are outperformed by the S-LSTM based NOMA detector. MMSE and SIC-based receivers' detection accuracy decreased by 15 dB SNR after limiting the number of pilots to 35, for both $VN_1$ and $VN_2$. Additional pilots make it abundantly evident that when the velocity between communication nodes rises, the fading links transform from being frequency-flat to time-selective, resulting in a decreased SER. However, the DL NOMA receiver can equal the performance of the 120-pilot example. This demonstrates that the S-LSTM-based receiver is more dependable for many pilots and can attain a higher level of performance with a lower number of pilots.

In Fig. 15, a comparison of performance is provided for various values of CP. Considering the scenario in which $length_{CP}$ is greater than $length_{impulse\ response}$, DL NOMA works considerably better as opposed to the scenario the $length_{CP}$ is less than $length_{impulse\ response}$. Because the channel is changing at every instant, it has been demonstrated in the literature [12]–[16] that neither MMSE nor LS schemes can properly detect the fading channel coefficients. Due to the ISI effects caused by temporal selectivity, the ideal ML-based NOMA receiver can no longer provide the optimal response, even in the event of perfect CSI. In a real-time communication scenario,

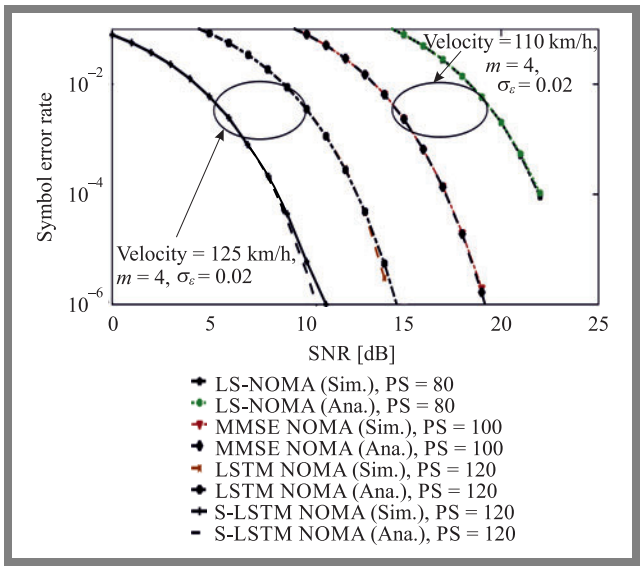



**Fig. 14.** Performance comparison between S-LSTM and conventional NOMA receivers for various pilot numbers, considering node mobility and error variance equal to $\sigma_{\epsilon}^2 = 0.02$.

frequency-selective links are needed to check the robustness of the S-LSTM NOMA receiver.

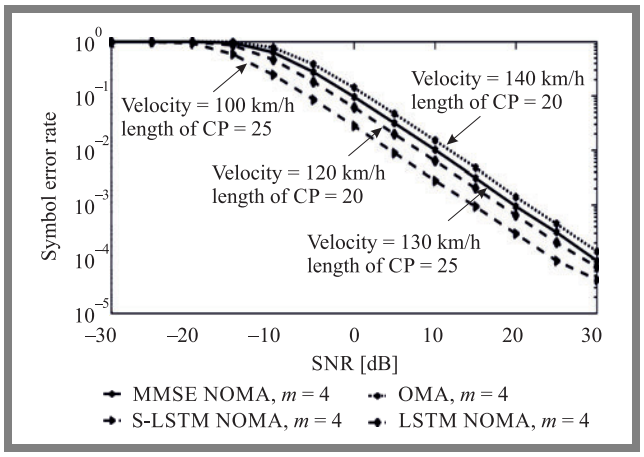



**Fig. 15.** Performance comparison between S-LSTM and conventional NOMA receivers for various CP length values, considering node mobility and channel error variance equal to 0.02.

When the influence of node mobility is disregarded, the S-LSTM NOMA receiver's end-to-end error performance under frequency flat fading channel circumstances is in near agreement with the ideal ML-based NOMA receiver. Additionally, the end-to-end error performance of the Nakagami-$m$ connections considerably improved as the shape parameter or amount of fading was increased. Additionally, the S-LSTM NOMA detector for $VN_2$ (far user) in Fig. 16 is robust in terms of signal intensity and has an impact on conventional error estimation. Performance of the DL NOMA detector is equivalent to the optimal fading channel circumstances and is more resistant to random channel fluctuations.

The simulated results have demonstrated that a reduction in node velocity improves end-to-end system performance. One of the issues that arises in connection with the use of power amplifiers, is non-linear clipping noise. To maintain

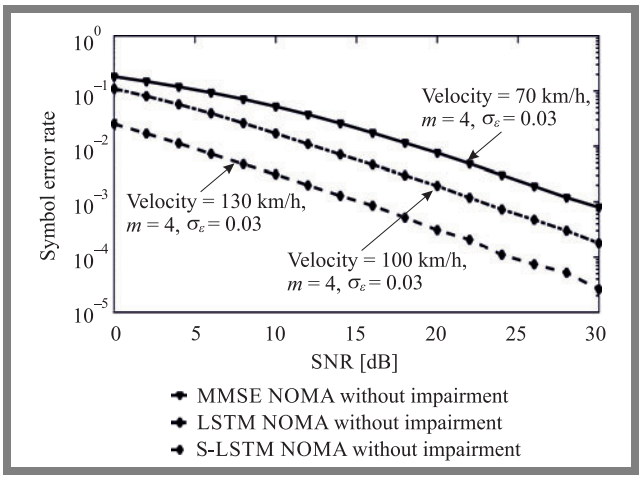



**Fig. 16.** SER vs. SNR plots considering all impairments for time-selective Nakagami-$m$ fading connections.

the linearity of the power amplifier, the envelope cancellation scheme is used in this work. In Fig. 16, end-to-end error performance comparisons have been provided between the S-LSTM, MMSE and conventional NOMA receivers, considering BPSK modulated symbols. The S-LSTM NOMA detector has been found to perform better overall than the traditional SIC-based NOMA receiver when the clipping ratio is equal to two and the SNR is more than 14 dB. As shown in Fig. 17, performance improves significantly when S-LSTM is used, but the end-to-end system performance improves with a decrease in node velocity.

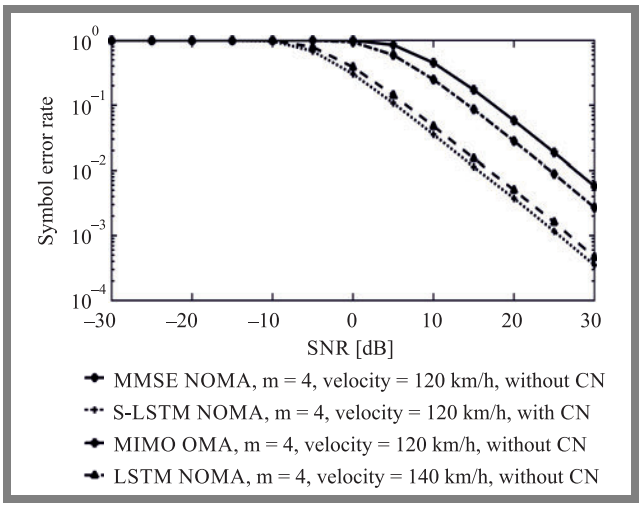



**Fig. 17.** Performance comparison between S-LSTM NOMA and other conventional NOMA schemes for various CP.

The time-varying fading channel coefficients in the online phase, considering BPSK symbols, are calculated using offline data sets. Performance difference between the offline and online states is shown in time-selective fading. Furthermore, these differences must be stable for the trained model. Figure 18 presents the results of the effect that is caused by changing the fading relationship statistics that are used during the training and testing stages. It is easy to understand that the performance of SER will improve if the number of

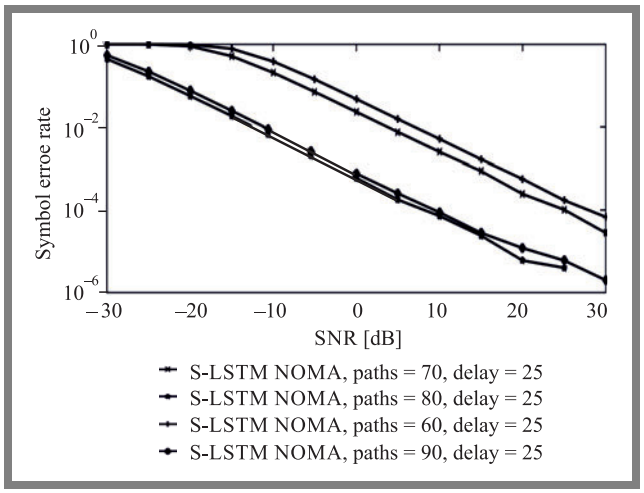



**Fig. 18.** Performance S-LSTM NOMA for various number of paths and delays with node velocity equal to 120 km/h.

possible paths is increased while maintaining the same fixed propagation delay.

### *5.3. Impact of Training Rate and Packet Length*

The achievable data rate plots for $\mathrm{VN}_S$ are demonstrated in Fig. 19, along with an investigation for the S-LSTM MIMO-NOMA receiver trained at various training rates. It is evident that a lower training rate leads to a greater data transmission rate, supporting the hypothesis that a higher training rate would result in more frequent weight changes and a bigger validation error. A lower training rate of 0.004 improves accuracy, but slows down convergence, since more updates are required.

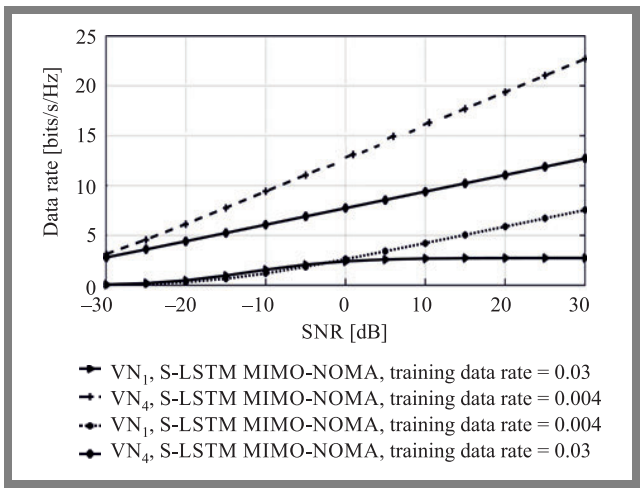



**Fig. 19.** Achievable data rates for various training data rates, considering node mobility equal to 120 km/h and error variance of $\sigma_\epsilon^2 = 0.02$.

To balance training accuracy and duration, the training rate has been adjusted to 0.03 for all other simulation scenarios. Smaller packets, however, offer reduced testing precision. Figure 20 shows the data rate that is achievable for different packet lengths when the node mobility is set to 130 km/h and the error variance is set to $\sigma_\epsilon^2 = 0.02$. Additionally, more data is used to offer a more precise estimate of the gradient for

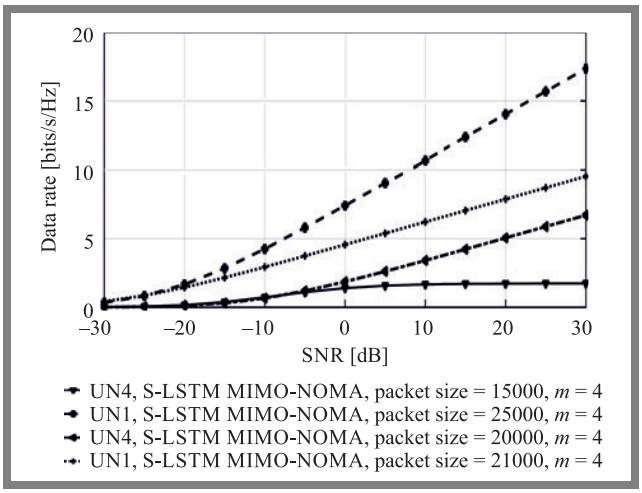



**Fig. 20.** Achievable data rate for various packet lengths, considering node mobility equal to 130 km/h.

each update, even though larger-sized packets require fewer iterations and updates of S-LSTM's parameters. Therefore, a final receiver with superior performance benefits from larger packet sizes.

## 6. Conclusion

This work presents a preliminary analysis of an S-LSTM-based MIMO-NOMA receiver and considers node mobility, SIC error and channel estimation error. Relative velocity between the BS and mobile users yields the ISI and time selectivity results in the enhanced end-to-end SER increase and OP. The simulation results show that the S-LSTM technique works better than the conventional SIC receiver and is more resistant to limited radio resources, such as SE, energy efficiency, number of pilot symbols, training rate, packet length, and CP than previous channel estimation methodologies. For more complex models, such as S-LSTM MIMO-NOMA systems, more research and testing will be performed. The entire training and testing procedure presented in this work starts with a time-selective channel profile. The impacts of node mobility and imperfect CSI will be examined for use in actual applications, in order to further evaluate the S-LSTM NOMA model's resistance to random channel profiles.

## References

[1] M.H. Eiza, *et al.*, "Investigation of routing reliability of vehicular ad hoc networks", *J. Wireless Com Network.*, vol. 179, 2013 (DOI: 10.1186/1687-1499-2013-179).

[2] E.U. Ogbodo, E. Utochukwu, A.M. Abu-Mahfouz, and A.M. Kurien, "A Survey on 5G and LPWAN-IoT for Improved Smart Cities and Remote Area Applications: From the Aspect of Architecture and Security", *Sensors*, vol. 22, pp. 6313, no. 16, 2022 (URL: https://www.mdpi.com/1424-8220/22/16/6313/htm).

[3] Q. Qiu, D. Wang, X. Du, S. Yu, S. Liu, and B. Zhao, "Security standards and measures for massive IoT in the 5G era", *Mobile Networks and Applications*, vol. 27, no. 1, pp. 392–403, 2022 (DOI: 10.1007/s11036-021-01841-2).

[4] M. Verma, J. Sheetlani, V. Mishra, and M. Mishra, "An Integrated Technique for Security of Cellular 5G-IoT Network Healthcare Ar-

chitecture", *G. Ranganathan, R. Bestak R. Palanisamy, Á. Rocha, Pervasive Computing and Social Networking, Lecture Notes in Networks and Systems*, vol. 317, pp. 549–563, Singapore, 2022 (DOI: 10.1007/978-981-16-5640-8_42).

[5] T. Xu, C. Xu, and Z. Xu, "An efficient three-factor privacy-preserving authentication and key agreement protocol for vehicular ad-hoc network", *China Communications*, vol. 18, no. 12, pp. 315–331, 2021 (DOI: 10.23919/JCC.2021.12.020).

[6] H. Badis and A. Rachedi, "Modeling tools to evaluate the performance of wireless multi-hop networks", ed.: Mohammad S. Obaidat, Petros Nicopolitidis, Faouzi Zarai, *Modeling and Simulation of Computer Networks and Systems, Morgan Kaufmann*, pp. 653–682, ISBN 9780128008874, 2015 (DOI: 10.1016/B978-0-12-800887-4.00023-7).

[7] R. Tiwari and S. Deshmukh, "Analysis and design of an efficient handoff management strategy via velocity estimation in HetNets", *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, pp. e3642, 2022 (DOI: 10.1002/ett.3642).

[8] R. Tiwari and S. Deshmukh. "Prior information-based Bayesian MMSE estimation of velocity in HetNets", *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 81–84, 2018 (DOI: 10.1109/LWC.2018.2857805).

[9] R. Tiwari and S. Deshmukh, "MVU estimate of user velocity via gamma distributed handover count in HetNets", *IEEE Communications Letters*, vol. 23, no. 3, pp. 482–485, 2019 (DOI: 10.1109/LCOMM.2019.2892962).

[10] D.K. Patel, *et al.*, "Performance Analysis of NOMA in Vehicular Communications Over i.n.i.d. Nakagami-$m$ Fading Channels", *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6254–6268, Oct. 2021 (DOI: 10.1109/TWC.2021.3073050).

[11] H. Ahmad, D.M. Ali, W.N.W. Muhamad, and M.S. Idris, "Performance analysis of NOMA in pedestrian and vehicular environments", *Journal of Physics: Conference Series, IOP Publishing*, vol. 1502, no. 1, pp. 012003, 2020 (DOI: 10.1088/1742-6596/1502/1/012003).

[12] B.P. Chaudhary, R. Shankar, and R.K. Mishra, "A tutorial on cooperative non-orthogonal multiple access networks", *The Journal of Defense Modeling and Simulation*, vol. 19, no. 4, pp. 563–573, 2022 (DOI: 10.1177/1548512920986627).

[13] M.K. Beuria, R. Shankar, and S.S. Singh, "Analysis of the energy harvesting non-orthogonal multiple access technique for defense applications over Rayleigh fading channel conditions", *The Journal of Defense Modeling and Simulation*, vol. 19, no. 4, pp. 821–828, 2022 (DOI: 10.1177/15485129211021168).

[14] R. Malladi, M.K. Beuria, R. Shankar, and S.S. Singh, "Investigation of the fifth generation non-orthogonal multiple access technique for defense applications using deep learning", *The Journal of Defense Modeling and Simulation*, 2021 (DOI: 10.1177/15485129211022857).

[15] R. Shankar, T.V. Ramana, P. Singh, S. Gupta, and H. Mehraj, "Examination of the Non-Orthogonal Multiple Access System Using Long Short Memory Based Deep Neural Network", *Journal of Mobile Multimedia*, pp. 451–474, 2021 (DOI: 10.13052/jmm1550-4646.18214).

[16] R. Shankar, *et al.*, "Examination of user pairing NOMA system considering the DQN scheme over time-varying fading channel conditions", *Journal of Information Science and Engineering*, vol. 38, no. 4, 2022 (DOI: 10.6688/JISE.202207 38(4).0010).

[17] L. Chouhan, P.K. Sharma, and N. Varshney, "Optimal Transmitted Molecules and Decision Threshold for Drift-Induced Diffusive Molecular Channel With Mobile Nanomachines", *IEEE Transactions on NanoBioscience*, vol. 18, no. 4, pp. 651–660, 2019 (DOI: 10.1109/TNB.2019.2935241).

[18] Y. Sun, D.W.K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems", *IEEE Tran. Commun.*, vol. 65, no. 3, pp. 1077–1091, 2017 (DOI: 10.1109/TCOMM.2017.2650992).

[19] M. Choi, J. Kim, and J. Moon, "Dynamic power allocation and user scheduling for power-efficient and delay-constrained multiple access networks", *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019 (DOI: 10.1109/TWC.2019.2929809).

[20] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications", *IEEE Trans. Wireless Commun*, vol. 17, no. 5, pp. 2961–2974, 2018 (DOI: 10.1109/TWC.2018.2804953).

[21] B. Soni, D.K. Patel, Y.L. Guan, S. Sun, Y. C. Chang, and J. M-Y. Lim, "Performance Analysis of NOMA aided Cooperative Relaying over $\alpha$-$\eta$-$\kappa$-$\mu$ Fading Channels", *National Conference on Communications (NCC)*, pp. 1–6, 2020 (DOI: 10.1109/NCC48643.2020.9056034).

[22] J. Men, J. Ge, and C. Zhang, "Performance analysis of non-orthogonal multiple access for relaying networks over Nakagami-$m$ fading channels", *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1200–1208, 2017 (DOI: 10.1109/TVT.2016.2555399).

[23] J. Men, J. Ge, and C. Zhang, "Performance analysis for downlink relaying aided non-orthogonal multiple access networks with imperfect CSI over Nakagami-$m$ fading", *IEEE Access*, vol. 5, pp. 998–1004, 2017 (DOI: 10.1109/ACCESS.2016.2631482).

[24] Y. Dursun, F. Fang, and Z. Ding, "Hybrid NOMA based MIMO offloading for mobile edge computing in 6G networks", *China Communications*, vol. 19, no. 10, pp. 12–20, 2022 (DOI: 10.23919/JCC.2022.00.024).

[25] Y. Dursun, K. Wang, and Z. Ding, "Secrecy sum rate maximization for a MIMO-NOMA uplink transmission in 6G networks", *Physical Communication*, vol. 53, pp. 101675, 2022 (DOI: 10.1016/j.phycom.2022.101675).

[26] D. Wan, M. Wen, F. Ji, Y. Liu, and Y. Huang, "Cooperative NOMA systems with partial channel state information over Nakagami-m fading channels", *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 947–958, 2017 (DOI: 10.1109/TCOMM.2017.2772273).

[27] K-H. Nguyen, H.V. Nguyen, V-P. Bui, and O-S. Shin, "Dynamic User Pairing for Non-Orthogonal Multiple Access in Downlink Networks", *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, pp. 111–114, 2021 (DOI: 10.1109/ICCE48956.2021.9352141).

[28] H. Ye, G.Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems", *IEEE Wireless Communications Letters*, vol. 7, 2018 (DOI: 10.1109/LWC.2017.2757490).

[29] N. David, U. Wolfgang, and W. Thomas, "Deep channel estimation", *Proc. 21st International ITG Workshop on Smart Antennas*, 2017.

[30] M. Vaezi, R. Schober, Z. Ding, and H.V. Poor, "Nonorthogonal Multiple Access: Common Myths and Critical Questions", *IEEE Wireless Communications Magazine*, vol. 26, no. 5, 2019 (DOI: 10.1109/MWC.2019.1800598).

[31] G. Gui, H. Huang, Y. Song, and H. Sar, "Deep Learning for an Effective Non-orthogonal Multiple Access Scheme", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, 2018 (DOI: 10.1109/TVT.2018.2848294).

[32] J-M. Kang, I-M. Kim, and C-J. Chun, "Deep Learning-Based MIMO-NOMA With Imperfect SIC Decoding", *IEEE Systems Journal*, vol. 14, no. 3, pp. 3414–3417, 2020 (DOI: 10.1109/JSYST.2019.2937463).

[33] J. Pan, N. Ye, A. Wang, and X. Li, "A Deep Learning-Aided Detection Method for FTN-Based NOMA", *Hindawi Wireless Communications and Mobile Computing*, vol. 2020, 2020 (DOI: 10.1155/2020/5684851).

[34] L. Chuan, C. Qing, and L. Xianxu, "Uplink NOMA signal transmission with convolutional neural networks approach", *Journal of Systems Engineering and Electronics*, vol. 31, 2020 (DOI: 10.23919/JSEE.2020.000068).

[35] M.A. Aref and S.K. Jayaweera, "Deep Learning-aided Successive Interference Cancellation for MIMO-NOMA", *IEEE Global Communications Conference*, pp. 1–5, 2020 (DOI: 10.1109/GLOBECOM42002.2020.9348107).

[36] R. He, A.F. Molisch, F. Tufvesson, Z. Zhong, B. Ai, and T. Zhang, "Vehicle-to-vehicle propagation models with large vehicle obstructions", *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2237–2248, 2014 (DOI: 10.1109/TITS.2014.2311514).

[37] M. Alouini and A.J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels", *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, 1999 (DOI: 10.1109/26.789668).

[38] D. Kalita, "An Overview on Long Short Term Memory (LSTM)", (URL: https://www.analyticsvidhya.com/blog/2022/03/an-overview-on-long-short-term-memory-lstm/).

[39] –, "Long Short-Term Memory Networks", (URL: https://in.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html).
[40] R. Shankar and R.K. Mishra, "PEP and OP examination of relaying network over time-selective fading channel", *SN Appl. Sci.*, vol. 2, pp. 1329, 2020 (DOI: 10.1007/s42452–020–3077–5).
[41] C.B. Barneto *et al.*, "Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements", *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10 pp. 4042–4054, 2019 (DOI: 10.1109/TMTT.2019.2930510).
[42] R. Shankar and R.K. Mishra, "Investigation of multiple hop cooperative communication system over time-selective Nakagami-$m$ fading channel", *Iran. J. Comput. Sci.*, vol. 3, pp. 145–168, 2020 (DOI: 10.1007/s42044-020-00054-2).
[43] C. Lin, Q. Changr, and X. Li, "A Deep Learning Approach for MIMO-NOMA Downlink Signal Detection", *Sensors*, vol. 19, no. 11, pp. 2526, 2019 (DOI: 10.3390/s19112526).
[44] K.M. Deep, "Learning SIC Approach for Uplink MIMO-NOMA System", *University of Oulu, Faculty of Information Technology and Electrical Engineering, Degree Programme in Wireless Communications Engineering, Master's thesis*, 50, p. 2022 (URL: http://jultika.oulu.fi/files/nbnfioulu-202208163315.pdf).
[45] Z. Wei, D.W.K. Ng, and J. Yuan, "Joint Pilot and Payload Power Control for Uplink MIMO-NOMA With MRC-SIC Receivers", *IEEE Communications Letters*, vol. 22, no. 4, pp. 692–695, 2018 (DOI: 10.1109/LCOMM.2018.2798658).
[46] Z. Wei, L. Yang, D. W. K. Ng, and J. Yuan, "On the Performance Gain of NOMA over OMA in Uplink Single-Cell Systems", *2018 IEEE Global Communications Conference*, pp. 1–7, 2018 (DOI: 10.1109/GLOCOM.2018.8647743).
[47] Z. Wei, D.W. Ng, and J. Yuan, "Joint Pilot and Payload Power Control for Uplink MIMO-NOMA with MRC-SIC Receivers", *IEEE Communications Letters*, vol. 22, no. 4, 2018 (DOI: 10.1109/LCOMM.2018.2798658).
[48] D. Di Caprio and F.J. Santos-Arteaga, "Enhancing the pattern recognition capacity of machine learning techniques: The importance of feature positioning", *Machine Learning with Applications*, vol. 7, pp. 100196, 2022 (DOI: 10.1016/j.mlwa.2021.100196).
[49] V. Andiappan and V. Ponnusamy, "Deep Learning Enhanced NOMA System: A Survey on Future Scope and Challenges", *Wireless Pers. Commun.*, vol. 123, pp. 839–877, 2022 (DOI: 10.1007/s11277-021-09160-1).
[50] Narengerile and J. Thompson, "Deep Learning for Signal Detection in Non-Orthogonal Multiple Access Wireless Systems", *UK/China Emerging Technologies (UCET)*, pp. 1–4, 2019 (DOI: 10.1109/UCET.2019.8881888).
[51] Narengerile, "Deep learning for signal detection in NOMA", *MATLAB Central File Exchange*, 2022 (URL: https://www.mathworks.com/matlabcentral/fileexchange/75478-deep-learning-for-signal-detection-in-noma-systems).
[52] X. Du, H. Zhang, H.V. Nguyen, and Z. Han, "Stacked LSTM Deep Learning Model for Traffic Prediction in Vehicle-to-Vehicle Communication", *IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, 2017 (DOI: 10.1109/VTCFall.2017.8288312).
[53] J. Brownlee, "Stacked Long Short-Term Memory Networks", (URL: https://machinelearningmastery.com/stacked-long-short-term-memory-networks/).

**Ravi Shankar** received his B.E. degree in Electronics and Communication Engineering from Jiwaji University, India, in 2006. He received his M. Tech. degree in Electronics and Communication Engineering from GGSIPU, India, in 2012. He received a Ph.D. in Wireless Communication from the National Institute of Technology Patna, in 2019. He was an Assistant Professor at MRCE Faridabad, from 2013 to 2014, where he was engaged in researching wireless communication networks. He is presently an Assistant Professor at SR University, Warangal, India. His current research interests cover cooperative communication, D2D communication, IoT/M2M networks and network protocols.

https://orcid.org/0000-0001-7532-3275
E-mail: ravi.mrce@gmail.com
Electronics and Communication Engineering Department, SR University, India

**Bhanu Pratap Chaudhary** is working as Research Scholar at National Institute of Technology Patna in Electronics and Communication Engineering department. He received his M. Tech. degree in Microelectronics & amp, VLSI design from National Institute of Technology, India. He received his B.Tech. degree in Electronics & Communication Engineering from CET-IILM-AHL, India. He is presently an Assistant Professor at Manav Rachna International Institute of Research and Studies, India. His current research interests cover cooperative communication, D2D communication, IoT/M2M networks and network protocols.

https://orcid.org/0000-0001-7085-4130
E-mail: bhanuc.phd18.ec@nitp.ac.in
Electronics and Communication Engineering Department, National Institute of Technology Patna, India

**Ritesh Kumar Mishra** received his B.E. degree in Electronics and Communication Engineering from Shivaji University, India, in 1998. He received his M.Tech. degree in Electronics and Communication Engineering from University of Burdwan, WB, India, in 2004. He received his Ph.D. in Wireless Communication from LNMU, Bihar, India, in 2011. He is presently working as an Assistant Professor at the National Institute of Technology Patna, Patna, India and has been engaged in researching wireless communication networks. His current research interests focus on cooperative communication, D2D communication, IoT/M2M networks.

https://orcid.org/0000-0002-7996-2231
E-mail: ritesh@nitp.ac.in
Electronics and Communication Engineering Department, National Institute of Technology Patna, India

Paper

# Reliability of Communication Systems Used in Offshore Wind Farms

Tadeus Uhl[1,2]

[1]*Maritime University of Szczecin, Szczecin, Poland,*
[2]*Flensburg University of Applied Sciences, Flensburg, Germany*

https://doi.org/10.26636/jtit.2022.166022

**Abstract — In the era of renewable energy, offshore wind farms play a very important role. The number of such installations in Europe is increasing rapidly. With the growing capacity of wind turbines installed in these farms (3, 5, 10 MW), the profitability of this type of energy systems plays an increasing role. The number of wind energy turbines installed at offshore wind farms is growing constantly as well. Once installed, the power plants must be under constant technical supervision, with reliability of electronic communication systems being a particularly important aspect in the operation of offshore wind farms. Considerations focusing on this subject form the very core of this paper. After an introduction to offshore wind farms, the following aspects will be discussed: redundant topologies, e.g. multiple HiPERRings, redundant switches and routers within the backbone networks, redundancy of the transmission media used, alternative transmission technologies, e.g. WLANs (IEEE 802.11h, IEEE 802.11g). Finally, requirements applicable to reliable electronic communication systems used in offshore wind farms will be formulated.**

***Keywords — communications technology, offshore wind farms, reliability, redundancy, telematics systems***

## 1. Introduction

The 21st century is characterized, inter alia, by major changes taking place in the field of energy sources. Renewable energy generated with the use of wind, water, sun and biological resources is relied upon ever more widely and plays an important role in the energy mix. Wind energy offers great potential here, as this source is available virtually anywhere in the world. Consequently, many companies have been stimulated [1] to design, manufacture and operate sufficiently large and profitable wind turbines. Especially in coastal areas, such a source of energy source may be very profitable. Therefore, in recent years, many offshore wind farms have been erected in numerous coastal countries of Europe (e.g. England, the Netherlands and Germany) and all over the world (e.g. USA, China, Australia) [2]. The construction of such farms is very costly and time-consuming. It also required a significant technical effort and poses a number of logistical challenges.

Safety plays an important role in the operation of offshore wind farms. Three distinct aspects may be distinguished here:
– safety of personnel performing construction, operation, and maintenance works,
– safety of navigation through wind farms,
– reliability of communications between specific wind farms systems and components.

In the first case, the focus is on identifying specific components/elements of the wind farm, as well as on describing the escape routes and alarm systems used in the event of an accident/failure. In the second case, rules need to be developed to enable shipping along very narrow (2 km wide) corridors between water farms, especially at waterway intersections. In the third case, reliable communication needs to be ensured between components/subsystems of the wind farm, regardless of the weather conditions and of the time of year. It is this area of activity that constitutes the main topic of this work. This paper systematizes and expands on the content of the internal technical report [3].

Initially, aspects related to the construction of offshore wind farms will be presented briefly. Next, a general concept of a telematics system relied upon for ensuring communication within these farms will be described. The article will focus primarily on the techniques and methods fostering reliability of this type of communication. The "redundancy" will play a special role here. At the end of the work, requirements applicable to a reliable telematics system used in offshore wind farms will be formulated. The work will conclude with a summary and outlook concerning future work.

## 2. Telematics Systems in Offshore Wind Farms

Construction of wind farms in coastal areas requires intensive cooperation and coordination between wind farm operators on the one hand and manufacturing, administration, and transport companies on the other. This is a very complicated and time-consuming process. Good management is a key factor determining whether such projects turn out to be successful from the business point of view.

Figure 1 shows a typical layout of an offshore wind farm [4]. It illustrates how large such farms may be and how close are turbines located to each other (in rows!). In order for this comprehensive system to work effectively, many specialized systems supporting its management and monitoring are necessary. Reliable communication between these systems plays a key role here.

**Fig. 1.** Offshore projects by Bernstein and Citrin in the North Sea.

In order to ensure effective and reliable operation of offshore wind farms, all components and subsystems of these farms must remain under constant supervision. Suitable telematics systems must be designed and implemented for this purpose. Figure 2 shows a block diagram of a telematics system intended for offshore wind farm projects [3].

The telematics system presented in Fig. 2 consists of numerous hardware subsystems and software components, such as:

- WET control station. It is responsible for the controlling the operation of individual wind turbines;
- Power management. It collects measurement results concerning, inter alia, voltage and current levels, turbine power, frequency of the current generates. This system also allows to control the measured parameters of the wind farm;
- Communication connection. It serves as an interface with the telecommunications network providing access to individual subsystems of the wind farms;
- Weather station. It measures the meteorological parameters, such as wind direction and force, air temperature, etc.;
- Online portal. It collects large amount of data concerning the wind farm;
- Reporting station. It offers visualizations of the current condition of and provides an overview of the historical data related to the wind farm.

Offshore wind farms are built at a considerable distance from the coast, for example in Germany: Alpha Ventus – 43 km, Borkum Riffgrund West – 50 km, Dan Tysk – 70 km, Sandbank24 – 90 km. Due to the long distances to mainland, the farms' telematic systems are connected to public networks using optical fibers. At distances of up to several kilometers, multimode 50/125 $\mu$m optical fibers with a wavelength of 1300 $\mu$m and attenuation values of 1 dB/km are used. For longer distances, which should not exceed 30 km, however, 9/125 $\mu$m single-mode optical fibers with a wavelength of 1300 $\mu$m and attenuation value of 0.4 dB/km are used. For distances of over 30 km, 9/125 $\mu$m single-mode fibers with a wavelength of 1550 $\mu$m and attenuation factor of 0.25 dB/km are relied upon.

The technical requirements that must be met by specific wind farm components are another important aspect. All parts must function reliably in harsh environmental conditions, often over periods of many years. Electronic devices must function properly over a wide temperature range ($-30^{\circ}$C to $+60^{\circ}$C) and must also prevent the ingress of moisture and contamination. IT systems used must comply with the ISO 20653 standard [5] as well.

Telematics systems for offshore wind farms need to comply with strict reliability requirements. Due to long distances between wind farms and mainland, maintenance and servicing operations may only be performed in favorable weather conditions and are governed by numerous constraints. Thus, a high level of reliability of IT systems used in wind farms must be guaranteed. Table 1 shows an overview of the specific reliability classes according to [6].

**Tab. 1.** Overview of reliability classes and unavailability time.

| Class | Reliability [%] | Down time, per year |
|---|---|---|
| Stable | 99.0 | 3.7 days |
| Available | 99.9 | 8.8 h |
| Highly available | 99.99 | 52.6 min |
| Insensitive to errors | 99.999 | 5.3 min |
| Fault tolerant | 99.9999 | 32 s |
| Free of errors | 99.9999 | 3 s |

To illustrate the importance of reliability, two failure scenarios affecting offshore wind farms will be discussed. The first one assumes that errors occur without any external interference, e.g. failure of a device supplying power to electronic components. The other one involves errors caused by deliberate human action, e.g. by switching off devices while performing maintenance work. In the first case, failures may be prevented by using such equipment as uninterruptible power supplies (UPS) which are capable of supplying electricity to electronic devices for the required period of time. In the other case, reliability depends on redundant solutions implemented within the wind farms. This topic will be covered in detail in the next section.

## 3. Redundancy in Wind Parks

When designing and building high-quality telematics systems, special attention should be paid to the redundancy of their components. Redundant elements may be connected to each other in series or in parallel. It is known from the reliability theory that the reliability of a system relying on series connections is described by:

$$P(A \cup B) = P(A) + P(B) = P(A \cap B), \tag{1}$$

For parallel connections, Eq (1) takes the following form:

$$P(A \cup B) = P(A) \cdot P(B). \tag{2}$$

For example, if $P(A) = P(B) = 0.1$, then total reliability of a system made up of components connected in series

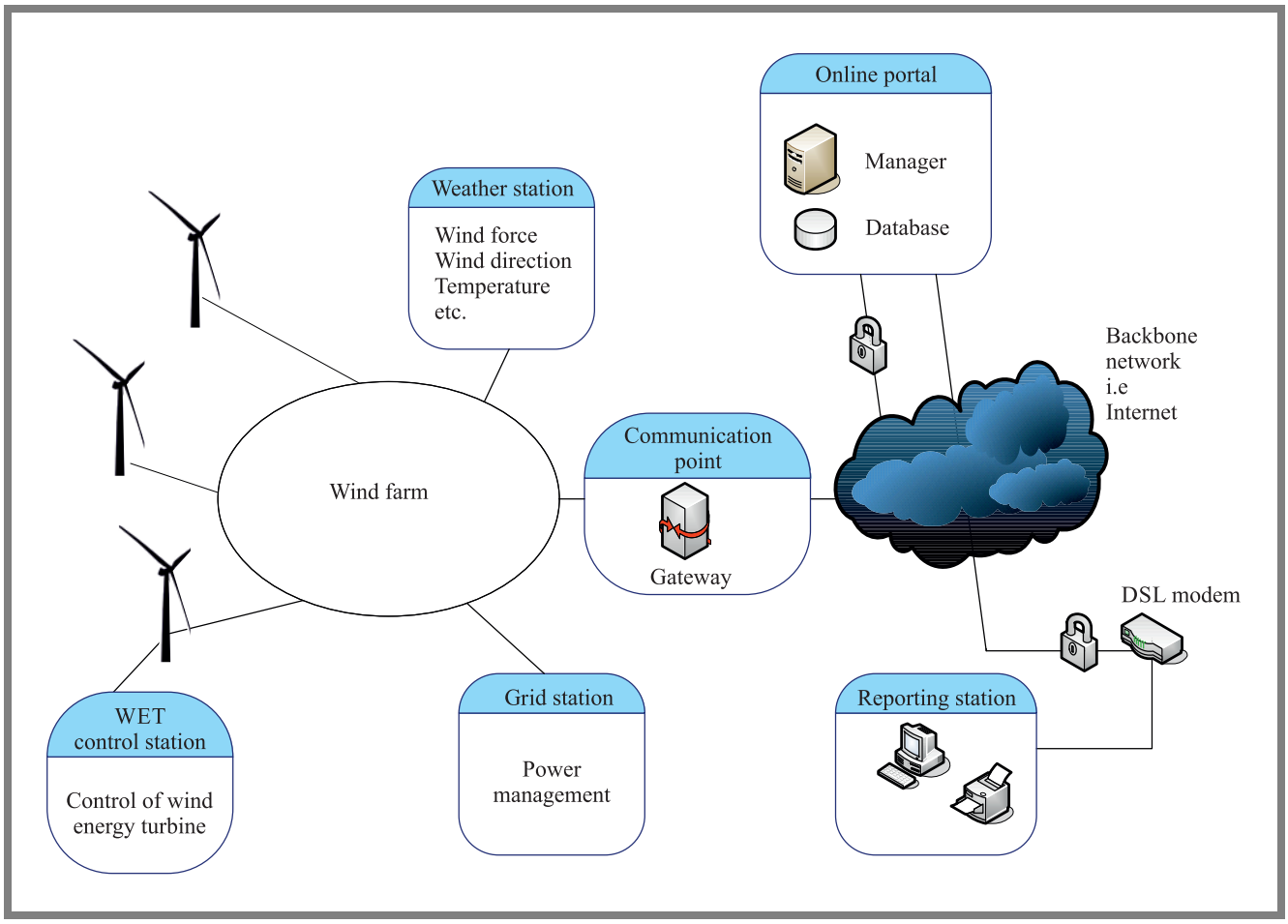



**Fig. 2.** Block diagram of a telematics system intended for offshore wind farms.

equals 0.81, and in the case of a parallel connection, it amounts to 0.99. This is a very convincing example, showing that in order to ensure high reliability of the entire system, its redundant elements should be connected in parallel. Therefore, parallel connections are used in redundant offshore wind farm systems.

### *3.1. Ring Redundancy*

The individual wind energy turbines (WET) are most often positioned linearly and at equal distances from each other (see Fig. 1). Figure 3 shows a typical ring topology connection between WETs, relying on optical fibers. Two optical fibers (transmit and receive) are sufficient for a basic connection between WETs. Four optical fibers are required for a redundant WET connection. Since Ethernet data transfer between two points can only take place using an unambiguous connection path, the primary ring structure is interrupted by a redundancy check. To overcome this problem, a redundancy manager (RM) is used. It is a functional block implemented in the switch with the smallest number and is only available once in HiPERRing [7]. If all the connections in HiPERRing are working properly, the RM transforms the ring structure into one of the linear type. Using test and control packages, the RM controls the functioning of the ring structure. If one switch fails or the optical fiber in the physical ring fails, the test packets sent via one port will not be received using the other. In this case, the RM activates the redundant connection, i.e. uses the other fiber optic path.

The reconfiguration time should be less than 500 ms. Therefore, the number of WETs in one ring is limited to approx. 10–15. If an offshore wind farm is made up of a large number of WETs, these must be linked together by multiple HiPERRings. The individual HiPERRings lead to the communications point (see Fig. 2), where they are connected to each other.

Using the multiple HiPERRing technique, it is possible to efficiently commutate redundant rings. The rings are divided into main ring and subrings. Naturally, it is possible to connect multiple subrings to one main ring. The multiple HiPERRing technique enables the start and end point of one subring to be connected to the main ring at different locations (Fig. 4).

If media redundancy protocol MRP [8] is used for commutation in the sub-rings, it must also be used in the main ring and in the LANs defined within the main ring. If a HiPERRing is used in the main rings, the MRP protocol can be used in the subrings. The formation of mesh structures between rings or their cascading are not allowed.

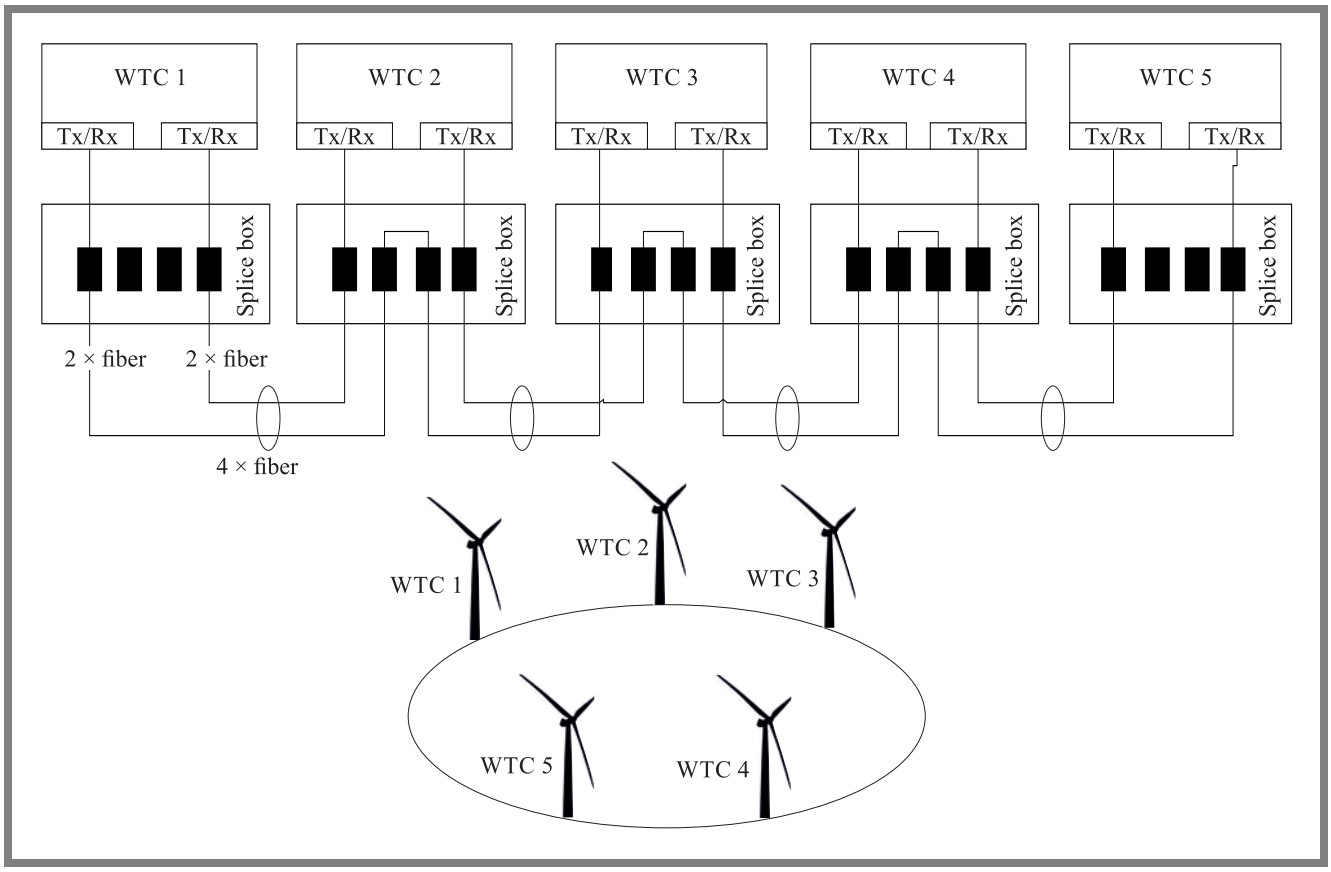



**Fig. 3.** Typical WET ring structure cabling layout with redundant optical fibers.

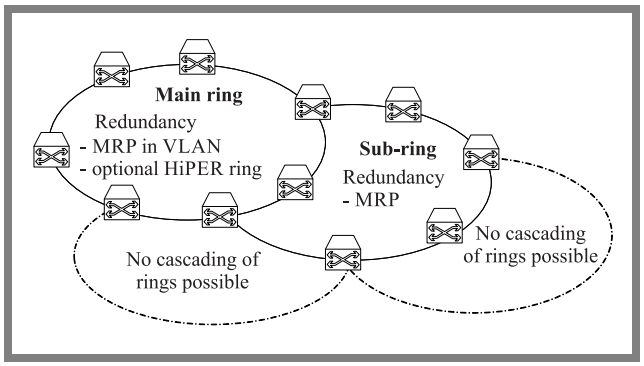



**Fig. 4.** Structure of a telematics system based on the main- and sub-rings concept.

By using additional cables, it is also possible to ensure a redundant connection between the two rings. Such an approach guarantees that communication may be maintained also if more than one ring element malfunctions. Access to other functioning components is possible by relying on the redundant cables existing within the structure. However, HiPERRing or MRP techniques cannot be used at the junction of the two rings, where the Rapid Spanning Tree Protocol [9]–[10] (RSTP) must be deployed.

### *3.2. System Redundancy*

The term "system redundancy" means that multiple IT components are used. Such redundancy is not always necessary or beneficial, and always leads to increased costs and additional implementation-related efforts.

Figure 5 shows an example solution ensuring system redundancy, in which one defective HiPERRing is connected to two independent switches in the backbone network (communication point, see Fig. 2). The failure of one switch in the backbone is immediately compensated for by a well-functioning second switch. Communication between switches in the backbone could be additionally supported by relying on the link aggregation technique [10]–[11]. The use of such a technique would be another factor boosting redundancy of the system.

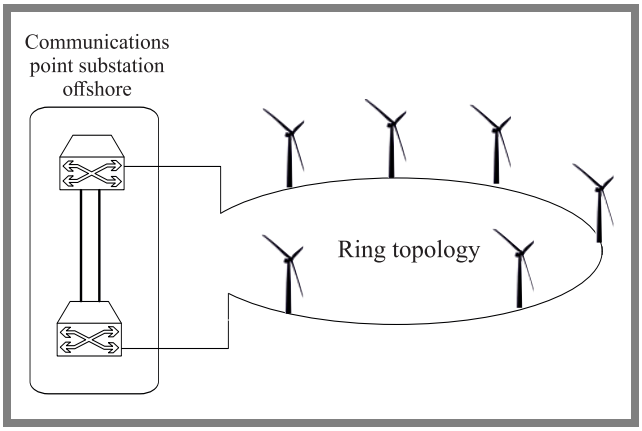



**Fig. 5.** Example of a solution offering system redundancy.

### *3.3. Media Redundancy*

Media redundancy involves the use of various forms of communication. Cables are used as the primary communication medium. In practice, as an alternative, radio channels are used as well. In order for the alternative communication medium to function effectively, the system components relying on cable and radio channels must be completely independent of each other. This means that the use of different communication media leads to the creation of separate, independently functioning communication networks.

Figure 6 shows an example of the use of various communication media. Two rings are presented, with one of them based on a cable and the other on a radio link. Both rings are connected to each other via suitable coupling elements at the communication point. Since WETs are placed in rows (see Fig. 1), the radio link must always provide a WET-WET connection. As the MRP protocol is used, it would be advisable here to ensure a low reconfiguration time. On the other hand, this type of connection requires a significant implementation effort, as each WET must be equipped with two antennas. Furthermore, even if one WET fails, the radio ring is interrupted immediately. This is a major drawback of the ring structure and, therefore, such a solution is not used in offshore wind farms.

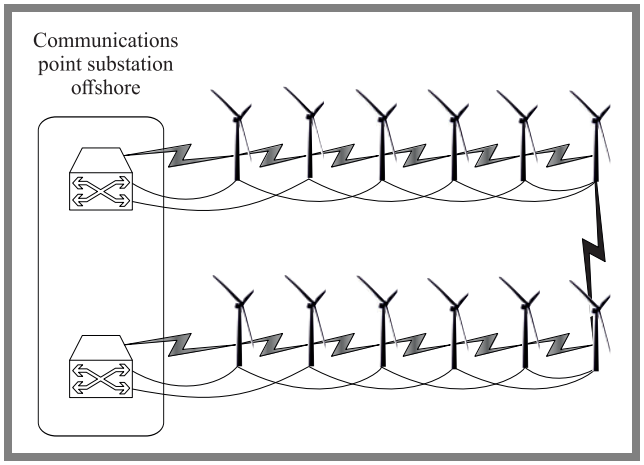



**Fig. 6.** Example of a ring-shaped radio network.

Figure 7 illustrates two cable-based ring networks and an additional star-shaped radio network. The radio network is controlled from a central point (master station). Hence, it is possible to communicate via radio channels with any WET device (slave station). This structure is dominated by point-to-multipoint connections. The insignificant implementation effort is an advantage here (only one antenna per WET is required only). Additionally, each WET can be individually pooled by the master station. Moreover, the failure of one WET does not affect the communication between other stations. These are the big advantages of the star topology. The drawback of star-shaped systems is the low bandwidth they are capable of achieving on dedicated radio channels. Their throughput depends on the number of WETs within the farm and on the shared medium assigned to the entire system. Additionally, it may be the case (especially in large farms) that radio waves emitted by individual WETs will interfere with each other and, therefore, communication may be difficult.

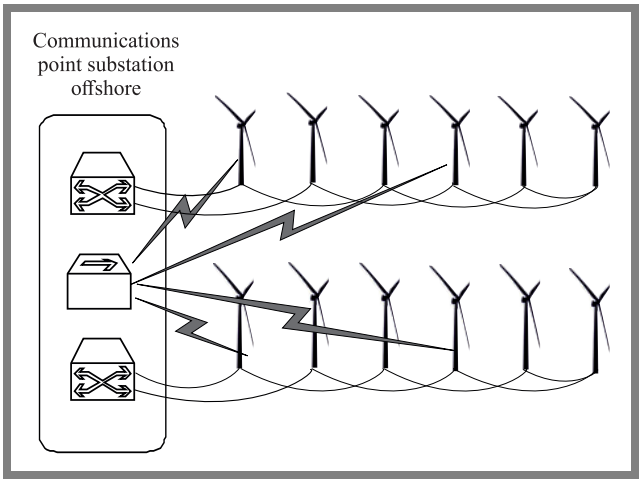



**Fig. 7.** Example of a star topology radio network.

The star topology seems to be a suitable solution for offshore wind farms. It is easy to implement and offers good reliability of the entire system. Radio networks may be based on a number of well-known standards, such as IEEE 802.11h (up to 54 Mbps in the 5 GHz band [12]) and IEEE 802.11g (up to 54 Mbps in the 2.4 GHz band [13]). Upcoming standards, such as IEEE 802.11n (248 Mbps gross bitrate in the 2.4 and 5 GHz [14] bands) and IEEE 802.11s (for radio mesh networks [15]) should be taken into account as well. When using these types of systems, one should bear in mind that the total capacity of the system is shared among all installed stations. In addition, data transfer speeds may be significantly reduced due to interference (e.g. generated by radar working in the 5 GHz band). Weather conditions also exert a negative impact on radio channels, causing interference and, consequently, degrading the range of the networks. Finally, it should be emphasized that in the case of radio networks, the transmitted data must be secured by technologies such as wired equivalent privacy (WEP) [16] and Wi-Fi protected access (WPA) [17].

## 4. Requirements for Telematics Systems in Wind Farms

Based on the information presented in the previous chapters, one may specify the requirements applicable to telematics systems used in offshore wind farms. The highlights of such a specification include the following:

– IT components must be adapted to the prevailing environmental and operational conditions (i.e. temperature, air chemistry, mechanical impacts, humidity, etc.). The most stringent ISO standards have to be respected;
– Components of telematics systems must be implemented redundantly. Parallel connections between individual elements of the telematics system should be used;
– The communication networks used should be redundant, i.e. should be based on main and subring topologies;

– When designing and building telematics systems, attention should be paid to the use of various types of transmission media (cables, radio links). It should be noted here that media redundancy involves the construction and operation of communication systems operating independently;
– State-of-the-art data security solutions must be implemented;
– When designing and building telematic systems, international standards applicable to IT systems need to be observed.

## 5. Summary and Outlook

This paper is devoted to the reliability of telematics systems used in offshore wind farms. After a short introduction, the main characteristics of such systems and the data network structures most frequently used in wind farms are described. Next, the general concept of a telematics system used for communication in offshore wind farms is presented, and techniques boosting the reliability of communication are shown. The notion of redundancy and its important role are described as ell. At the end of the article, a catalogue of safety- and reliability-related requirements applicable to electronic communication systems used in offshore wind farms is created.

A summary of this paper was presented at the 12th International Scientific Conference EXPLO-SHIP 2022 "Problems Concerning the Operation of Vessels and Port Facilities" [18].

The reliability of telematics systems used in wind farms, as analyzed in this study, is only one of the three important aspects guaranteeing the security of offshore wind farms. The remaining crucial aspects include the safety of navigation and personnel. There is still much room for improvement in these fields. Research concerned with the safety of navigation in offshore wind farms is facilitated by maritime navigation simulators. They allow to create conflict situations and identify the potential solutions. The Flensburg University of Applied Science operates a very modern Nautical Center equipped with five captain bridge simulators. Similar simulators are also available at the Maritime University in Szczecin. It would be advisable to establish cooperation in this particular field of research. Further research focusing on this topic is planned in the future.

## References


[1] –, Wind turbine manufacturers, http://en.wikipedia.org/wiki/List_of_wind_turbine_manufacturers.
[2] –, List of offshore wind farms, http://en.wikipedia.org/wiki/List_of_offshore_wind_farms.
[3] N. Struve, "Planung- und Konzeptionierung eines Kommunikationsnetzes für Offshore Windpark Thornton Bank (Belgien)", (in German) *Internal technical report, Flensburg University of Applied Sciences*, 2010.
[4] –, Offshore-projects by company BARD, http://www.windkraft-journal.de/2011/10/20/tuv-sud-zertifiziert-520-mw-offshore-projekte-von-bard.
[5] –, Standard ISO 20653, http://en.wikipedia.org/wiki/IP_Code.
[6] –, High availability stages, http://en.wikipedia.org/wiki/High_availability.
[7] Hirtschmann: Anwender-Handbuch (in German). Hirschman Automation and Control GmbH, Neckartenzlingen, 2010.
[8] –, Standard IEC 62439-2, http://en.wikipedia.org/wiki/Media_Redundancy_Protocol.
[9] –, Standard IEEE 802.1w, http://www.ieee802.org/1/pages/802.1w.html.
[10] K. Nowicki and T. Uhl, "Ethernet End-to-End", *1th Edition, Shaker-Publisher*, Germany, 2008 (ISBN: 978383832271404).
[11] –, Standard IEEE 802.1AX, http://ieee802.org/3/axay.
[12] –, Standard IEEE 802.11h, http://en.wikipedia.org/wiki/IEEE_802.11h-2003.
[13] –, Standard IEEE 802.11g, http://en.wikipedia.org/wiki/IEEE_802.11g.
[14] –, Standard IEEE 802.11n, http://en.wikipedia.org/wiki/IEEE_802.11n.
[15] –, Standard IEEE 802.11s, http://en.wikipedia.org/wiki/IEEE_802.11s.
[16] –, Wired Equivalent Privacy, http://en.wikipedia.org/wiki/Wired_Equivalent_Privacy.
[17] –, Wi-Fi Protected Access, http://en.wikipedia.org/wiki/Wi-Fi_Protected_Access.
[18] –, "Problems of vessels and port facilities operation", *12th International Scientific Conference EXPLO-SHIP*, 2022.



**Tadeus Uhl** received his M.Sc. in Telecommunications from the Academy of Technology and Agriculture in Bydgoszcz, Poland in 1975, his Ph.D. from the Gdańsk University of Technology, Poland in 1982 and his D.Sc. from Dortmund University, Germany in 1990. Since 1992, he has been working as a Professor at the Institute of Communications Technology, Flensburg University of Applied Sciences, Germany and, in addition, since 2013, as a Professor at the Institute of Transport Engineering and Economics, Maritime University of Szczecin, Poland. His activities focus on the following areas: traffic engineering, performance analysis of communications systems, measurement and evaluation of communications protocols, QoS and QoE by triple play services, Ethernet and IP technology. He is an author or co-author of five books and 130 papers dealing with LAN, WAN and NGN.

https://orcid.org/0000-0001-6849-9168
E-mail: t.uhl@pm-szczecin.pl, tadeus.uhl@hs-flensburg.de
Maritime University of Szczecin, Szczecin, Poland
Flensburg University of Applied Sciences, Flensburg, Germany

Paper

# High-level and Low-level Feature Set for Image Caption Generation with Optimized Convolutional Neural Network

Roshni Padate, Amit Jain, Mukesh Kalla, and Arvind Sharma

*Department of Computer Science and Engineering, Sir Padampat Singhania University, India*

https://doi.org/10.26636/jtit.2022.164222

**Abstract — Automatic creation of image descriptions, i.e. captioning of images, is an important topic in artificial intelligence (AI) that bridges the gap between computer vision (CV) and natural language processing (NLP). Currently, neural networks are becoming increasingly popular in captioning images and researchers are looking for more efficient models for CV and sequence-sequence systems. This study focuses on a new image caption generation model that is divided into two stages. Initially, low-level features, such as contrast, sharpness, color and their high-level counterparts, such as motion and facial impact score, are extracted. Then, an optimized convolutional neural network (CNN) is harnessed to generate the captions from images. To enhance the accuracy of the process, the weights of CNN are optimally tuned via spider monkey optimization with sine chaotic map evaluation (SMO-SCME). The development of the proposed method is evaluated with a diversity of metrics.**

***Keywords — CNN, image caption, proposed contrast, sharpness, SMO-SCME algorithm***

**Tab. 1.** List of terms and abbreviations.

| | |
|---|---|
| AACR | Algebraic amalgamation-based composed representation |
| BI-LSTM | Bidirectional LSTM |
| BI-GRU | Bidirectional gated recurrent units |
| Bleu_1 | Bilingual evaluation understudy |
| CNN | Convolutional neural network |
| CV | Computer vision |
| CMBO | Cat mouse-based optimization |
| DG-GAN | Dual generator GAN |
| DL | Deep learning |
| DBN | Deep belief network |
| FCA | Face centrality attribute |
| GLP | Global leader phase |
| GLL | Global leader learning phase |
| GLD | Global leader decision phase |
| gLSTM | Guiding LSTM |
| GAN | Generative adversarial networks |
| IGGAN | Interactions guided GAN |
| LLD | Local leader decision phase |
| LSTM | Long short-term memory |
| LLL | Local leader learning phase |
| LP | Learning percentage |
| LLP | Local leader phase |
| MK-KDES | Multiple kernel-kernel descriptors |
| MFO | Moth flame optimization |
| NLD | Natural language description |
| NLP | Natural language processing |
| NLG | Natural language generation |
| NN | Neural network |
| RMCNet | ResNet with multi-scale module and adaptive channel attention |
| SSA | Salp swarm algorithm |
| SSO | Shark smell optimization |
| SMO-SCME | SMO with sine chaotic map evaluation |
| sLSTM | Stacked LSTM |
| SMO | Spider monkey optimization |
| TR | Tag refinement |
| WHO | Wild horse optimizer |

## 1. Introduction

Automated image captioning entails image capturing, examining its content, and generating a linguistic description [1]–[3]. Automated portrayal of image contents, with an appropriate level of expressiveness and accuracy, is a demanding task. In order for the captioning of an image to be effective, three fundamental requirements need to be satisfied, namely identifying attributes, objects, and their mutual associations [4]–[7]. However, identification of objects and their associations with images is insufficient for producing informative textual descriptions. Therefore, in order to recognize the content of a given image and to generate a text, an automated tool is necessary which would be capable of create various inferences from these identified objects and their associations [8]–[10]. An effective image captioning model aids in multimodal machine conversion and word sensing disambiguation by reducing the level of uncertainty in words [7], [11]–[13].

Three key architectures of DL-oriented image captioning schemes may be distinguished, namely end-to-end, attention-based, and compositional. The schemes deployed for captioning images are classified as those relying on the template-based matching and retrieval-based approaches. In the former, all attributes, actions, and objects visible in the image are derived and packed into a stiff sentence pattern [14], [15]. The yield of the template-oriented matching model is not always sufficient and clear. In the latter approach, a visually identical image is recovered from a larger database and the recovered image's captions are matched with the query image [16], [17]. This model is characterized by a low level of flexibility in terms of adjusting the recovered captions and lacks in expressiveness and fluency [7], [18].

The contribution of the presented work is as follows. The authors proposed a new model for extracting both low-level and high-level data from image captions. The said data is then used for caption generation with enhanced CNN and SMO-SCME model weight adjustments.

The paper is arranged as follows. Section 2 presents related works. Section 3 introduces the model concerned and Section 4 explains the extraction of the proposed features. Section 5 describes the optimal CNN with SMO-SCME-aided optimization. Sections 6 and 7 illustrate the outcome and offer conclusions, respectively. Table 1 summarizes the abbreviations used.

## 2. Literature Review

Singh *et al.* [7] proposed an encoder-decoder oriented framework in which CNN was deployed to encode visual features of an image and sLSTM was relied upon, in combination with bi-directional LSTM and unidirectional LSTM, to generate captions in Hindi. To encode the ocular features of the image, sLSTM and "V GG19" designs were deployed in the process of generating captions at the decoder side. The results have shown that the adopted method achieved better results.

Ye *et al.* [19] proposed an image caption generation scheme depending on the optimized BI-LSTM model. A variant of MFO was developed for optimization purposes. The performance of the adopted approach was proven using varied datasets, such as Flicker 8k, Flicker 30k, VizWik and Coco datasets, using renowned metrics, such as Cider, Bleu, Spice and Rough. Performance simulations have proven that B-LSTM achieved better results over the remaining schemes. Zhang *et al.* [20] presented a system for creating a novel feature, known as MK-KDES-1, with the extraction of 3 KDES features and MKL scheme fusing. MK-KDES-1 features contribute to enhancing the Bleu score of captions. The subsequent issue was resolved by a new, effective two-layer tag refinement (TR) approach incorporated into the NLG scheme. Analyses have proven the system to be suitable for creating image captions.

Sur *et al.* [21] defined AACR as a simplifying, language structuring and modeling linguistic attribute (associated with grammar and language) that improved the grammatical structure and corrected the sentences. AACR allowed structure and represent feature spaces in a more accurate and unique manner. Shan *et al.* [22] developed the IGGAN approach for captioning images in an unsupervised mode. This technique combined object-to-object communications with multi-scale feature representation. Images were encrypted using RMC Net to gain a robust representation of features. The created sentence and the image were deployed to reconstruct one another in IGGAN. The approach produced sentences with no manual labeled image caption pairs. Yiwei *et al.* [23] presented a multi-attention method by deploying both non-local and local evidence for more effective image caption analysis. This scheme, known as Magan, includes a discriminator and a generator. The adopted generator aided in generating more exact sentences. At the same time, the discriminator was utilized to determine if the sentences created were machine-generated or case-specific.

Yang *et al.* [24] presented a capable approach relying on the EnsCaption model. The solution intended to improve an ensemble of generation-oriented and retrieval-oriented image captioning schemes via an innovative DG-GAN network. By relying on the adversary training procedure, caption re-ranking and caption generation, it offered better retrieved captions with superior scores and allowed lower ranking scores to be assigned to the retrieved and created image captions, thus offering improved effectiveness over other schemes.

Zhao *et al.* [25] developed a multimodal fusion scheme to produce descriptions portraying image contents. The developed technique involved four networks: CNN for feature extraction, image attribute extraction, sentence modeling, and recurrent network. Unlike extant models that anticipate the specific words based upon hidden states, the presented approach deployed the attributes of the image and designed recorded words.

Ding *et al.* [41] deployed high-level visual characteristics to serve as a foundation of a novel image captioning model. To identify areas of an image that need attention, low-level data, such as image value, was blended high-level elements. Tests relying on MSCOCO and Flickr data sets proved that the model performed well. Table 2 shows the comparison of different image captioning schemes.

## 3. Proposed Image Caption Generation Model

The proposed image captioning approach comprises three stages. Initially, feature extraction is performed, deriving low-level (contrast, sharpness, color) and high-level (motion and facial impact score) features. The said features are then fed to CNN that generates captions based on images. The CNN weights are optimized via the SMO-SCME model, resulting in precise image captioning. The SMO-SCME model is depicted in Fig. 1.

**Tab. 2.** Comparison of conventional image captioning systems.

| Reference | Deployed schemes | Features | Notes |
|---|---|---|---|
| [7] | CNN | • high efficiency<br>• improved Bleu score | Not capable of captivating the alphanumerical image content |
| [19] | B-LSTM | • high Bleu score<br>• enhanced Cider score | Occlusion may occur |
| [20] | MK-KDES-1 | • improved Bleu score<br>• creates more coherent features | Needs attention when dealing with visual image contents |
| [21] | AACR | • high Meteor score<br>• high Cider score | Other constructive elements should be included |
| [22] | IGGAN | • robust representation of features<br>• generates realistic sentences | Suffers from insufficient common sensical reasoning |
| [23] | Magan | • high Cider score<br>• high Meteor score | Additional discriminators increase the memory usage |
| [24] | EnsCaption model | • high Rouge score<br>• high Cider score | Ranking procedure needs to be improved |
| [25] | CNN | • chooses optimal captions<br>• enhanced accuracy | No deliberation on denser image captions |
| [41] | High- and low-level features | • efficiency is high<br>• accuracy is high | It is still difficult to describe visuals that have several objects in them |

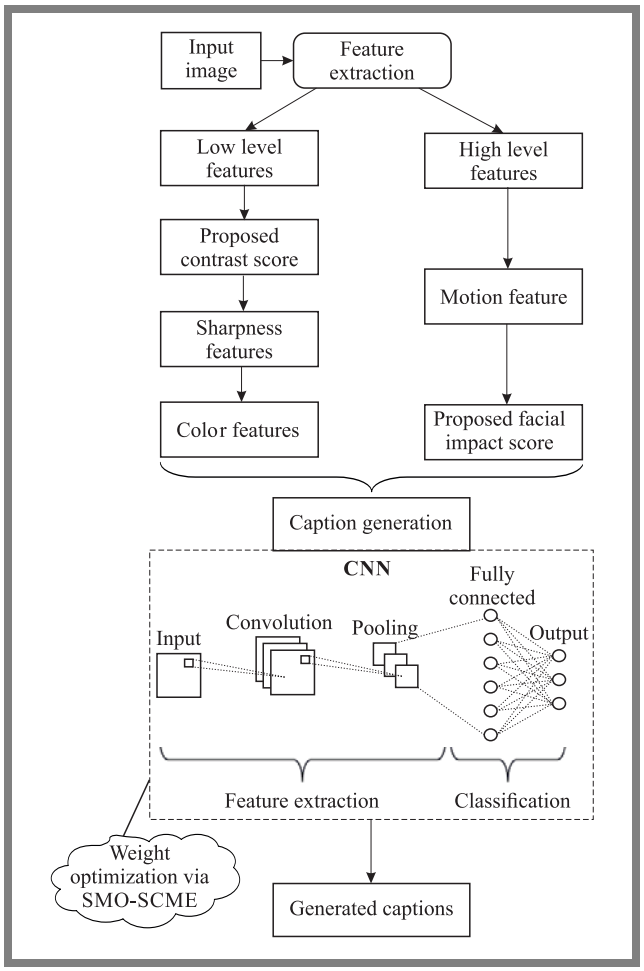



**Fig. 1.** Proposed image caption generation scheme.

# 4. Extraction of High-level and Low-level Features

Contrast [26] is defined as the ratio between the black and white parts of the image, representing the gradual change from black to white. The greater the contrast ratio, the more gradients there are, and the richer color in a given area. Initially, a gamma-corrected image is generated from the input image as:

$$R = CX^{\delta}, \tag{1}$$

where $C$ is a constant positive constraint for controlling brightness, $X$ refers to the original image (its values should be 0 and 1) and $\delta$ refers to a constant positive parameter which indicates the gamma value.

The formula for calculating is:

$$C = \sum_{y} \gamma(i,j)^2 P_{\gamma}(i,j), \tag{2}$$

where $P_{\gamma}(i,j)$ is probability of pixel distribution of gradation variation of $\gamma$ between neighboring pixels.

The $\gamma(i,j)$ parameter in Eq. (2) is calculated:

$$\gamma(i,j) = |i-j|. \tag{3}$$

However, for the proposed contrast score $\gamma(i,j)$ is:

$$\gamma(i,j) = \frac{|\vec{\mu} - i|}{|\vec{\mu} - j|}, \tag{4}$$

which represents the difference between neighboring pixels. $\vec{\mu}$ implies the average of $\gamma$ for neighboring pixels.

The candidate areas in the image for computing contrast score $C$ are initially transformed to luminance, and are then filtered and re-sampled for obtaining width and height [26].

Color is a significant feature for representing an image. It is invariant, regardless of the image's rotation, translation, and scaling. Color spacing, quantification, and similarity computations are the key elements of extracting color features [27].

Image sharpness is an important parameter for evaluating image quality. Sharpness quality directly impacts the subjective feelings of humans [26]. In an image, sharpness refers to its overall clarity in terms of both focus and contrast [28]. The derived low-level features (sharpness, color, and contrast) are together represented as $FT_{\text{low}}$.

These features offer temporal data and are generally attained using optical flow for detection the inappropriate motion at the backdrop [29]. A motion processing analysis may be performed for detecting motion, i.e. finding those locations where anything is moving within the image [30].

Facial information plays a significant role in expressing image semantics [26]. The presence of a face in every candidate region is determined by the $f$ score. Initially, a face denoted by $f_s$ is standardized to a frame by:

$$FS = \frac{Wi_f^2}{Ht_{im} \times Wi_m}, \tag{5}$$

in which $Wi_f^2$ refers to the width of the face bound box in pixels, $Ht_{im}$ refers to the height of the image and $Wi_m$ refers to the width of the image.

In the next step, FCA is formulated based on the Gaussian formula as:

$$\text{FCA} = \frac{1}{2\pi\sigma_y\sigma_x} \text{e}^{-\frac{\left|\frac{\Delta y^2}{\sigma_y^2}+\frac{\Delta y^2}{\sigma_y^2}\right|}{2}}, \tag{6}$$

where $\text{Cen}Y_f$ and $\text{Cen}X_f$ are the centroid column and the centroid row of the facial area, respectively, and:

$$\sigma_x = \frac{2Wi_m}{3}, \tag{7}$$

$$\sigma_y = \frac{Hy_{im}}{2}, \tag{8}$$

$$\Delta y = \left|\text{Cen}Y_f - \frac{wi_m}{2}\right|, \tag{9}$$

$$\Delta x = \left|\text{Cen}X_f - \frac{Ht_{im}}{5}\right|. \tag{10}$$

The facial impact score is:

$$f = \alpha FS \times \beta\,\text{FCA}, \tag{11}$$

where $\alpha$ and $\beta$ refer to scalars evaluated using logistic maps.

The derived high-level features (motion and proposed facial impact score) are indicated as $FT_{\text{high}}$.

# 5. CNN with SMO-SCME for Caption Generation.

The derived low-level and high-level features are together sent to CNN for caption generation. In CNN [26], every neuron gets connected to nearby neurons located in the preceding layer. At position in the $l$-th layer of the linked $w$-th feature map, the features are computed as:

$$B_{r,t,w}^l = W_w^{l^T} PI_{r,t}^l + D_w^l, \tag{12}$$

where $W_w^l$ is the weight, $D_w^l$ stands for the bias of the $w$-th filter linked to the $l$-th layer. At the center location $(r,t)$ of the $l$-th layer, the patch input is indicated as $PI_{r,t}^l$. The activation value $\text{act}_{r,t,w}^l$ related with convolutional features $B_{r,t,w}^l$ is:

$$\text{act}_{r,t,w}^l = \text{act}(B_{r,t,w}^l). \tag{13}$$

**In the pooling layer** the value of $C_{r,t,w}^l$ is:

$$C_{r,t,w}^l = \text{pool}(\text{act}_{m,h,w}^l), \quad \forall(m_h) \in NN_{r,t}, \tag{14}$$

where $NN_{r,t}$ is the neighbor's nearer position $(r,t)$.

The prediction result factor occurs at output layer of CNN. The CNN loss is:

$$\text{Loss} = \frac{1}{wn}\sum_{h=1}^{wn} l\left[\theta; C^{(h)}, F^{(h)}\right]. \tag{15}$$

The general constraint related with $W_w^l$ and $D_w^l$ is denoted as $\theta$, the counts of output-input relation is $PI^{(h)}, C^{(h)}; h \in [1,\ldots,wn]$. The $h$-th input, the label and output are denoted as $PI^{(h)}, C^{(h)}$ and $F^{(h)}$, respectively.

**For solution encoding**, CNN weights $W$ are optimally elected via the SMO-SCME scheme. Representation of the solution is shown in Fig. 2, where symbolizes the entire count of CNN weights. The objective Obj of this research is:

$$\text{Obj} = \text{Min}(Er), \tag{16}$$

where $Er$ implies an error.

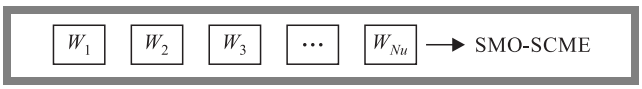



**Fig. 2.** Solution encoding scheme.

### *5.1. Proposed SMO-SCME Algorithm*

The existing SMO model [31] offers a low level of accuracy. Hence, to overcome this disadvantage, self-improvements are proposed [32]–[39]. The proposed SMO-SCME model includes four steps:

1) The monkey distance to the food calculation as kind of fitness.
2) The positions are updated and fitness is re-computed for each individual.
3) The local leader updates the position if the group is distributed as per perturbation rate.
4) When the group exceeds the bound of a maximum group, a parent group is formed.

The initialization of all $i$-th spider monkeys $SP_i$ takes place by:

$$SP_{ij} = SP_{\min j} + D_u(0,1) \times (SP_{\max j} - SP_{\min j}), \tag{17}$$

where $SP_{\max j}$ and $SP_{\min j}$ are the upper and lower limits of $SP_i$ in the $j$-th dimension, and $D_u(0,1)$ is a random number.

In the LLP phase, the $i$-th $SP$ of the $k$-th subgroup is specified using:

$$SP_{\text{new}\,ij} = SP_{\min j} + D_U(0,1) \times (ll_{kj} - SP_{ij}) + D_u(-1,1) \times (SP_{rj} - SP_{ij}), \tag{18}$$

where $SP_{ij}$ is the $i$-th $SP$ in the $j$-th dimension, $ll_{kj}$ is the $j$-th dimension of the $k$-th local group leader location, $gl$ is the group leader and $SP_{rj}$ is the $r$-th $SP$ chosen from a random $k$-th group in which $r \neq i$. In the proposed SMO-SCME, the update occurs as:

$$SP_{\text{new}\,ij} = SP_{\min j} + D_u(0,1) \times (ll_{kj} - SP_{ij}) + D_u(-1,1) \times (SP_{rj} - SP_{ij}) \times \left(w_{\max} - \frac{t}{t_{\max}}(w_{\max} - w_{\min})\right), \quad (19)$$

where implies the weight factor from 0 to 9.

During the GLP phase, the spider monkey position gets updated as:

$$SP_{\text{new}\,ij} = SP_{\min j} + D_u(0,1) \times (gl_j - SP_{ij}) + D_u(-1,1) \times (SP_{rj} - SP_{ij}), \quad (20)$$

where $gl_j$ is $GL$ position and $j \in 1, 2, \ldots, I$, $V$ is a random index. In the proposed SMO-SCME, the update occurs as:

$$SP_{\text{new}\,ij} = SP_{\min j} + D_u(0,1) - \text{prob}_i \times (gl_j - SP_{ij}) + D_u(-1,1) - \text{prob}_i \times (SP_{rj} - SP_{ij}), \quad (21)$$

where, $\text{prob}_i$ is a random probability factor created using the logistic map, $D_u(0,1)$ is computed based on a chaotic sine map.

The $\text{pro}_i$ factor is:

$$\text{pro}_i = y \times \frac{\text{fit}_i}{\max - \text{fit}} + x, \quad (22)$$

where $fit_i$ is fitness of the $i$-th monkey. The best outcomes are gained when $y = 0.9$ and $x = 0.1$.

The $GL$ position is updated by deploying a greedy selection procedure and $SP$ with the best fitness value is elected to determine the new position of $GL$. If $GL$ position is similar to the old one, the global limit count is increased by 1.

Similarly, the $ll$ position of entire groups is updated using a greedy selection procedure, and then $SP$ with the best fitness is selected. If the $ll$ position is similar, then the local limit count is increased by 1. If the position of $ll$ is not updated, i.e. the local leader limit is met, then the first step is repeated using:

$$SP_{\text{new}\,ij} = SP_{ij} + D_u(0,1) \times (gl_j - SP_{ij}) + D_u(0,1) \times (SP_{ij} - ll_{kj}). \quad (23)$$

The decision is taken based on the $gl$ position. The population is separated into subgroups if the $gl$ position is not updated. The groups are separated until the group count reaches the highest allowed number of groups $mg$ and, next, they are fused to form a group.

## 6. Simulation Results

The proposed CNN + SMO-SCME method for generating image captions was implemented in Python. The analysis was performed using the "Flickr Dataset" [40] and the performance of the CNN + SMO-SCME approach was evaluated in comparison with the faster R-CNN [40], GLSTM [34], LSTM, DBN, BI-GRU, CNN + CMBO, CNN + SSA, CNN + WHO, and CNN + SSO, with a wide variety of metrics benchmarked. Accordingly, the analysis was carried out by varying the LPs and different scores to represent the effectiveness of CNN + SMO-SCME. The results are shown in Fig. 3.


**Fig. 3.** Sample image representation, DBN, and CNN + SMO-SCME results.

### *6.1. Analysis of Varied Scores*

Meteor, Cider, and Rouge scores achieved by the CNN + SMO-SCME scheme are revealed in Fig. 4. Rouge provides a set of metrics used to assess the performance of machine translation and automatic summarization software in natural language processing. The metrics contrast is an automatically generated summary, a translation with a summary reference, translation or a collection of references. The Meteor measure has demonstrated stronger connection with human participants in terms of assessing the overall quality of the description. The remaining metrics rely on caption rankings and are unable to assess novel content.

It has been found that the correlation between these measurements and human judgement is rather poor. Our metric compares a machine-created sentence with a group of sentences that were written by humans for reference. The concepts of grammaticality, saliency, relevance, and accuracy (precision and recall) are essentially captured by our metric using sentence similarity.

The presented model was compared with Faster R-CNN [40], gLSTM [34], LSTM, DBN, BI-GRU, CNN + CMBO, CNN + SSA, CNN + WHO, and CNN + SSO schemes. The Meteor score attained by the CNN + SMO-SCME scheme, in comparison with Faster R-CNN [40], gLSTM [34], LSTM, DBN, BI-GRU, CNN + CMBO, CNN + SSA, CNN + WHO, LSTM [41] and CNN + SSO, regarding diverse LPs, is shown in Fig. 4a. The proposed model achieves a higher Meteor

score than the schemes it is compared with. The CNN + SMO-SCME model offers enhanced values at the 90th LP, compared with the Meteor score outputs at other LPs. Similarly, the Cider score of CNN + SMO-SCME method reached a higher value of 1.609 at the 90-th LP, whereas the conventional models achieved lower values.

Figure 4c shows that the CNN + SMO-SCME technique achieved a better Rouge score (0.70) than other models at the 90th LP.

In the next step, a CNN that has been optimized to produce captions based on an image is used to create the captions. The weights of CNN are tuned using SMO with SCME to increase caption generating accuracy.

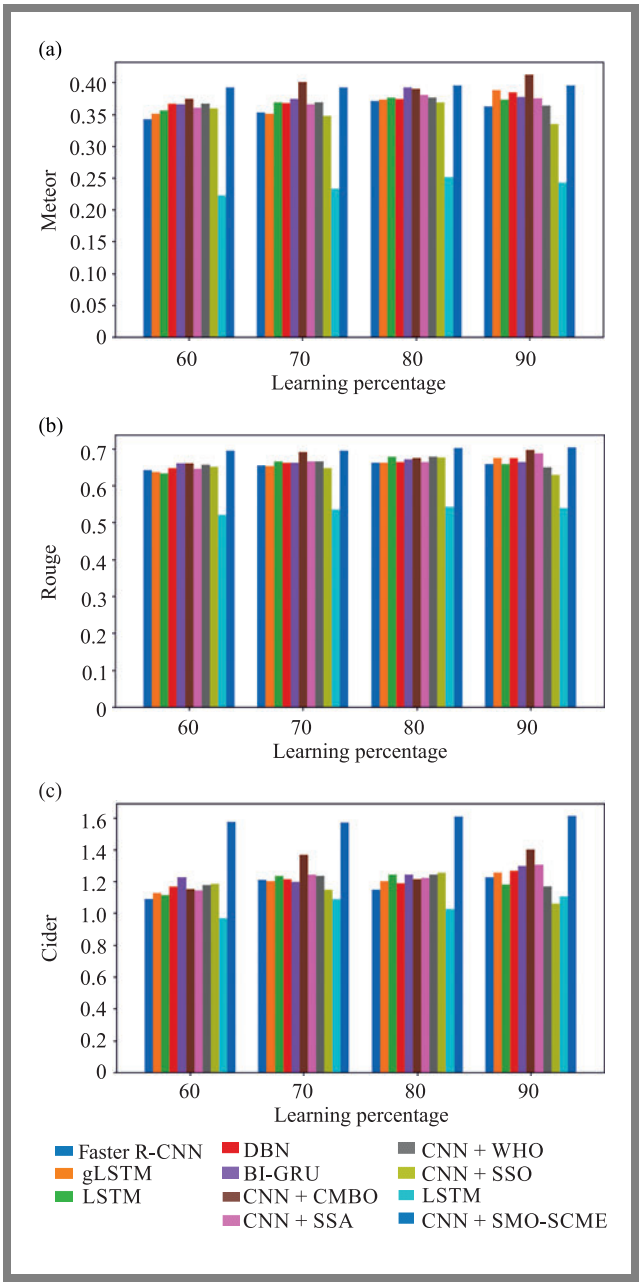



**Fig. 4.** Comparison of the proposed approach with other schemes for: (a) Meteor, (b) Cider, and (c) Rouge scores.

### *6.2. Analysis Using Bleu Score*

Bleu is a bilingual evaluation understudy algorithm which assesses accuracy of text in such a way that quality is defined as the similarity between machine-generated output and a professional human translation. Today, Bleu still remains one of the most widely used automated metrics. The Bleu score of the proposed CNN + SMO-SCME scheme is evaluated in comparison with Faster R-CNN [40], gLSTM [34], LSTM [41], DBN, BI-GRU, CNN + CMBO, CNN + SSA, CNN + WHO solutions, as well as with the CNN + SSO scheme for various LPs. The evaluation was performed with datasets from [40], and the relevant results are summarized in Table 3. As shown in the table, the CNN + SMO-SCME scheme provides better outputs at the 90th LP than other LPS. In the Blue 1 case, the DBN model achieved the worst results. Thus, the advantage of the developed model is proven.

### *6.3. Convergence Analysis*

The cost function of the SMO-SCME scheme, compared with CMBO, SSO, SSA, WHO, and SMO, is illustrated in Fig. 5. The SMO-SCME has generated lower costs in the 5–25 iteration range and higher cost values at lower iteration values. However, the cost factor still remains better than that of the compared schemes.

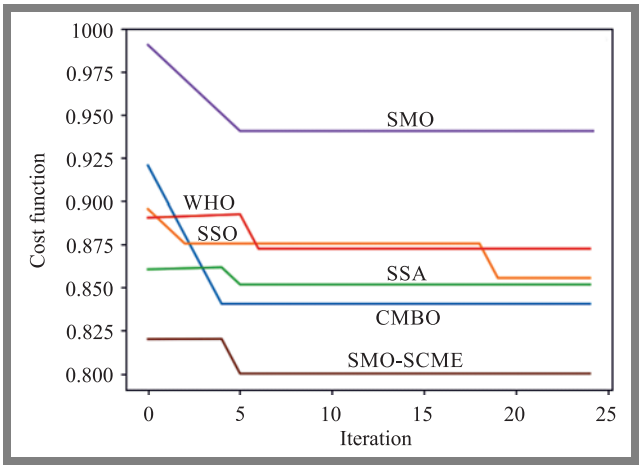



**Fig. 5.** Convergence analysis: SMO-SCME scheme vs. other models.

### *6.4. Time Analysis*

Table 4 presents the comparison of computational time typical of the SMO-SCME method and other methods (CMBO, SSO, SSA, WHO, and SMO) and allows to evaluate performance of the system. The SMO-SCME system achieved lower values than CMBO, SSA, WHO, and SMO methods. The SSO is characterized by the lowest value among all techniques assessed. However, since the remaining scores attained by the SMO-SCME are better than those of SSO, this difference can be disregarded. Next to SSO, the WHO model has gained the time over CMBO, SSA, and WHO. Thus, the computational time improvement of the SMO-SCME scheme is proven.

**Tab. 3.** Comparison of CNN + SMO-SCME model with other schemes using the Bleu score.

| LP | | Faster R – CNN [40] | gLSTM [34] | LSTM [41] | DBN | BI-GRU | CNN + CMBO | CNN + SSA | CNN + WHO | CNN + SSO | CNN + SMO-SCME |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 60 | Bleu-1 | 0.78 | 0.76 | 0.76 | 0.77 | 0.79 | 0.77 | 0.75 | 0.79 | 0.78 | 0.80 |
| 60 | Bleu-2 | 0.61 | 0.60 | 0.59 | 0.61 | 0.63 | 0.59 | 0.58 | 0.59 | 0.58 | 0.63 |
| 60 | Bleu-3 | 0.47 | 0.46 | 0.45 | 0.47 | 0.49 | 0.46 | 0.44 | 0.48 | 0.48 | 0.49 |
| 60 | Bleu-4 | 0.24 | 0.27 | 0.27 | 0.29 | 0.30 | 0.27 | 0.24 | 0.30 | 0.30 | 0.36 |
| 70 | Bleu-1 | 0.81 | 0.80 | 0.79 | 0.79 | 0.79 | 0.81 | 0.79 | 0.80 | 0.79 | 0.80 |
| 70 | Bleu-2 | 0.64 | 0.64 | 0.63 | 0.61 | 0.62 | 0.65 | 0.62 | 0.63 | 0.61 | 0.63 |
| 70 | Bleu-3 | 0.49 | 0.50 | 0.49 | 0.48 | 0.48 | 0.52 | 0.48 | 0.49 | 0.46 | 0.49 |
| 70 | Bleu-4 | 0.27 | 0.30 | 0.27 | 0.24 | 0.28 | 0.33 | 0.26 | 0.26 | 0.22 | 0.36 |
| 80 | Bleu-1 | 0.79 | 0.77 | 0.79 | 0.79 | 0.80 | 0.78 | 0.78 | 0.78 | 0.78 | 0.80 |
| 80 | Bleu-2 | 0.62 | 0.61 | 0.62 | 0.62 | 0.64 | 0.63 | 0.62 | 0.62 | 0.62 | 0.64 |
| 80 | Bleu-3 | 0.47 | 0.47 | 0.49 | 0.49 | 0.49 | 0.48 | 0.48 | 0.49 | 0.49 | 0.50 |
| 80 | Bleu-4 | 0.28 | 0.32 | 0.34 | 0.32 | 0.30 | 0.24 | 0.28 | 0.34 | 0.38 | 0.36 |
| 90 | Bleu-1 | 0.79 | 0.80 | 0.78 | 0.77 | 0.79 | 0.81 | 0.79 | 0.78 | 0.76 | 0.80 |
| 90 | Bleu-2 | 0.62 | 0.64 | 0.62 | 0.61 | 0.63 | 0.65 | 0.63 | 0.61 | 0.59 | 0.64 |
| 90 | Bleu-3 | 0.48 | 0.50 | 0.49 | 0.48 | 0.50 | 0.51 | 0.49 | 0.48 | 0.46 | 0.50 |
| 90 | Bleu-4 | 0.31 | 0.35 | 0.34 | 0.33 | 0.32 | 0.36 | 0.28 | 0.33 | 0.29 | 0.37 |

**Tab. 4.** Time analysis (in seconds) using SMO-SCME vs. conventional models.

| LP | CMBO | SSO | SSA | WHO | SMO | SMO-SCME |
|---|---|---|---|---|---|---|
| 60 | 186 | 63 | 435 | 178 | 116 | 99 |
| 70 | 185 | 93 | 415 | 180 | 116 | 99 |
| 80 | 195 | 93 | 415 | 180 | 116 | 100 |
| 90 | 195 | 93 | 415 | 180 | 116 | 100 |

### *6.5. Feature Analysis*

Table 5 presents the result of analysis of the developed CNN + SMO-SCME scheme and its comparison with the SMO-SCME model without lower-level features, SMO-SCME without higher-level features, and the proposed model without optimization. The study covers such metrics as Bleu, Meteor, Cider, and Rouge scores. The CNN + SMO-SCME achieved better values than SMO-SCME without lower-level features, SMO-SCME without higher-level features, and the proposed model without optimization. Moreover, the SMO-SCME model without lower-level features was characterized by relatively lower Bleu, Meteor, Cider, and Rouge scores than those of the SMO-SCME scheme without higher-level features and those of the proposed model without optimization.

Table 6 summarizes the performance of the former EC + SI-EFO scheme and the proposed CNN + SMO-SCME model. From the results, one may notice that the CNN + SMO-SCME model achieved better results than EC + SI-EFO. This is due to enhancements in the proposed methodology.

**Tab. 5.** Feature analysis and summary.

| Metrics | CNN + SMO-SCME | SMO-SCME without lower features | SMO-SCME without higher features | Proposed without optimization |
|---|---|---|---|---|
| Bleu-1 | 0.80 | 0.62 | 0.76 | 0.77 |
| Bleu-2 | 0.63 | 0.55 | 0.50 | 0.60 |
| Bleu-3 | 0.49 | 0.36 | 0.36 | 0.46 |
| Bleu-4 | 0.36 | 0.14 | 0.17 | 0.27 |
| Cider | 1.57 | 1.09 | 1.11 | 1.15 |
| Rouge | 0.69 | 0.64 | 0.54 | 0.66 |
| Meteor | 0.39 | 0.34 | 0.25 | 0.37 |

**Tab. 6.** Analysis of the EC + SI-EFO model and the proposed CNN + SMO-SCME scheme.

| Metrics | CNN + SMO-SCME | EC + SI-EFO |
|---|---|---|
| Bleu-1 | 0.80 | 0.77 |
| Bleu-2 | 0.63 | 0.58 |
| Bleu-3 | 0.49 | 0.44 |
| Bleu-4 | 0.36 | 0.26 |
| Cider | 1.57 | 1.43 |
| Rouge | 0.69 | 0.63 |
| Meteor | 0.39 | 0.37 |

## 7. Conclusion

The proposed image caption-generating model relying on low-level and high-level features provides improved caption

generation accuracy with the CNN weights tuned via SMO-SCME. The better performance of the offered scheme was proven in comparison with other methods. Specifically, the CNN + SMO-SCME scheme achieved better estimation results for all Bleu scores at the 90th LP. In the case of Bleu 1, the DBN model was characterized by outputs that were worse than those of other schemes, e.g. Faster R-CNN, gLSTM, LSTM, BI-GRU, CNN + CMBO, CNN + SSA, CNN + WHO, and CNN + SSO at the 90th LP. The SMO-SCME model achieves the minimum calculation cost value of 99.24 at the the 70th LP, which is negligible compared to CMBO, SSA, WHO, and SMO approaches.

## References


[1] Z. Deng, Z. Jiang, R. Lan, W. Huang, and X. Luo, "Image captioning using DenseNet network and adaptive attention", *Signal Processing: Image Communication*, vol. 85, 2020 (DOI: 10.1016/j.image.2020.115836).

[2] J. Su, J. Tang, Z. Lu, X. Han, and H. Zhang, "A neural image captioning model with caption-to-images semantic constructor", *Neurocomputing*, vol. 367, 2019, pp. 144–151 (DOI: 10.1016/j.neucom.2019.08.012).

[3] S. Bang and H. Kim, "Context-based information generation for managing UAV-acquired data using image captioning", *Automation in Construction*, vol. 112, 2020 (DOI: 10.1016/j.autcon.2020.103116).

[4] H. Wang, H. Wang, and K. Xu, "Evolutionary recurrent neural network for image captioning", *Neurocomputing*, vol. 401, pp. 249–256, 2020 (DOI: 10.1016/j.neucom.2020.03.087).

[5] R. Li, H. Liang, Y. Shi, F. Feng, and X. Wang, "Dual-CNN: A convolutional language decoder for paragraph image captioning", *Neurocomputing*, vol. 396, pp. 92–101, 2020 (DOI: 10.1016/j.neucom.2020.02.041).

[6] J. Guan and E. Wang, "Repeated review based image captioning for image evidence review", *Signal Processing: Image Communication*, vol. 63, pp. 141–148, 2018 (DOI: 10.1016/j.image.2018.02.005).

[7] A. Singh, T.D. Singh, and S. Bandyopadhyay, "An encoder-decoder based framework for hindi image caption generation", *Multimed. Tools Appl 80*, pp. 35721–35740, 2021 (DOI: 10.1007/s11042-021-11106-5).

[8] Ph. Kinghorn, L. Zhang, and L. Shao, "A region-based image caption generator with refined descriptions", *Neurocomputing*, vol. 272, pp. 416–424, 2018 (DOI: 10.1016/j.neucom.2017.07.014).

[9] Q. Liu, Y. Chen, J. Wang, and S. Zhang, "Multi-view pedestrian captioning with an attention topic CNN model", *Computers in Industry*, vol. 97, pp. 47–53, 2018 (DOI: 10.1016/j.compind.2018.01.015).

[10] G. Christie, A. Laddha, A. Agrawal, S. Antol, and D. Batra, "Resolving vision and language ambiguities together: Joint segmentation & prepositional attachment resolution in captioned scenes", *Computer Vision and Image Understanding*, vol. 163, pp. 101–112, 2017 (DOI: 10.1016/j.cviu.2017.09.001).

[11] F. Xiao, X. Gong, Y. Zhang, Y. Shen, and X. Gao, "DAA: Dual LSTMs with adaptive attention for image captioning", *Neurocomputing*, vol. 364, pp. 322–329, 2019 (DOI: 10.1016/j.neucom.2019.06.085).

[12] G. Huang and H. Hu, "c-RNN: A Fine-Grained Language Model for Image Captioning", *Neural Process Lett*, 2018 (DOI: 10.1007/s11063-018-9836-2).

[13] C. Wu, Y. Wei, X. Chu, F. Su, and L. Wang, "Modeling visual and word-conditional semantic attention for image captioning", *Signal Processing: Image Communication*, vol. 67, pp. 100–107, 2018 (DOI: 10.1016/j.image.2018.06.002).

[14] J. Yang, Y. Sun, J. Liang, B. Ren, and S. Lai, "Image captioning by incorporating affective concepts learned from both visual and textual components", *Neurocomputing*, 2018 (DOI: 10.1016/j.neucom.2018.03.078).

[15] T. Yinghua and C.S. Chee, "Phrase-based Image Caption Generator with Hierarchical LSTM Network", *Neurocomputing*, 2018 (DOI: 10.1016/j.neucom.2018.12.026).

[16] A. Yuan, X. Li, and X. Lu, "3G structure for image caption generation", *Neurocomputing*, 2018 (DOI: 10.1016/j.neucom.2018.10.059).

[17] Ch. Fan, Z. Zhang, and D.J. Crandall, "Deepdiary: Lifelogging image captioning and summarization", *Journal of Visual Communication and Image Representation*, vol. 55, pp. 40–55, 2018 (DOI: 10.1016/j.jvcir.2018.05.008).

[18] X. Chen, M. Zhang, Z. Wang, L. Zuo, and Y. Yang, "Leveraging Unpaired Out-of-Domain Data for Image Captioning", *Pattern Recognition Letters*, In press, accepted manuscript, 2018 (DOI: 10.1016/j.patrec.2018.12.018).

[19] Z. Ye, *et al.*, "A novel automatic image caption generation using bidirectional long-short term memory framework", *Multimed Tools Appl 80*, pp. 25557–25582, 2021 (DOI: 10.1007/s11042-021-10632-6).

[20] H. Zhang *et al.*, "Novel model to integrate word embeddings and syntactic trees for automatic caption generation from images", *Soft Comput 24*, pp. 1377–1397, 2020 (DOI: 10.1007/s00500-019-03973-w).

[21] C. Sur, "AACR: Feature Fusion Effects of Algebraic Amalgamation Composed Representation on (De)Compositional Network for Caption Generation for Images", *SN Comput. Sci. 1, 229*, 2020 (DOI: 10.1007/s42979-020-00238-4).

[22] C. Shan, A. Gaoyun, Z. Zhenxing, and R. Qiuqi, "Interactions guided generative adversarial network for unsupervised image captioning", *Neurocomputing*, vol. 417, pp. 419–431, 2020 (DOI: 10.1016/j.neucom.2020.08.019).

[23] Y. Wei, L. Wang, and C. Wu, "Multi-Attention Generative Adversarial Network for image captioning", *Neurocomputing*, vol. 387, pp. 91–99, 2019 (DOI: 10.1016/j.neucom.2019.12.073).

[24] M. Yang *et al.*, "An Ensemble of Generation- and Retrieval-Based Image Captioning With Dual Generator Generative Adversarial Network", *IEEE Transactions on Image Processing*, vol. 29, pp. 9627–9640, 2020 (DOI: 10.1109/TIP.2020.3028651).

[25] D. Zhao, Z. Chang, and S. Guo, "A multimodal fusion approach for image captioning", *Neurocomputing*, vol. 329, pp. 476–485, 2019 (DOI: 10.1016/j.neucom.2018.11.004).

[26] S. Ding, S. Qu, and S. Wan, "Image caption generation with high-level image features", *Pattern Recognition Letters*, vol. 123, pp. 89–95, 2019 (DOI: 10.1016/j.patrec.2019.03.021).

[27] S.R. Kodituwakku, "Comparison of Color Features for Image Retrieval", *Indian Journal of Computer Science and Engineering*, vol. 1, no. 3, pp. 207–211 (http://www.ijcse.com/docs/IJCSE10-01-03-06.pdf).

[28] –, https://photography.tutsplus.com/tutorials/what-is-image-sharpening--cms-26627.

[29] T. Bouwmans, C. Silva, C. Marghes, M.S. Zitouni, H. Bhaskar, and C. Frelicot, "On the role and the importance of features for background modeling and foreground detection", *Computer Science Review*, vol. 28, pp. 26–91, 2018 (ISSN 15740137, DOI: 10.1016/j.cosrev.2018.01.004).

[30] –, https://en.wikipedia.org/wiki/Motion_analysis.

[31] S. Harish, G. Hazrati, and J.C. Bansal, "Spider Monkey Optimization Algorithm", 2019 (DOI: 10.1007/978-3-319-91341-4_4).

[32] B.R. Rajakumar, "Impact of Static and Adaptive Mutation Techniques on Genetic Algorithm", *International Journal of Hybrid Intelligent Systems*, vol. 10, no. 1, pp. 11–22, 2013 (DOI: 10.3233/HIS-120161).

[33] B.R. Rajakumar, "Static and Adaptive Mutation Techniques for Genetic algorithm: A Systematic Comparative Analysis", *International Journal of Computational Science and Engineering*, vol. 8, no. 2, pp. 180–193, 2013 (DOI: 10.1504/IJCSE.2013.053087).

[34] S.M. Swamy, B.R. Rajakumar and I.R. Valarmathi, "Design of Hybrid Wind and Photovoltaic Power System using Opposition-based Genetic Algorithm with Cauchy Mutation", *IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013)*, 2013 (DOI: 10.1049/ic.2013.0361).

[35] A. George and B.R. Rajakumar, "APOGA: An Adaptive Population Pool Size based Genetic Algorithm", *AASRI Procedia – 2013 AASRI Conference on Intelligent Systems and Control (ISC 2013)*, vol. 4, pp. 288–296, 2013 (DOI: 10.1016/j.aasri.2013.10.043).

[36] B.R. Rajakumar and A. George, "A New Adaptive Mutation Tech-

nique for Genetic Algorithm", *In proceedings of IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–7, 2012 (DOI: 10.1109/ICCIC.2012.6510293).

[37] M.B. Wagh and N. Gomathi, "Improved GWO-CS Algorithm-Based Optimal Routing Strategy in VANET", *Journal of Networking and Communication Systems*, vol. 2, no. 1, pp. 34–42, 2019 (DOI: 10.46253/jnacs.v2i1.a4).

[38] S. Halbhavi, S.F. Kodad, S.K. Ambekar, and D. Manjunath, "Enhanced Invasive Weed Optimization Algorithm with Chaos Theory for Weightage based Combined Economic Emission Dispatch", *Journal of Computational Mechanics*, Power System and Control, vol. 2, no. 3, pp. 19–27, 2019 (DOI: 10.46253/jcmps.v2i3.a3).

[39] A.N. Jadhav and N. Gomathi, "DIGWO: Hybridization of Dragonfly Algorithm with Improved Grey Wolf Optimization Algorithm for Data Clustering", *Multimedia Research*, vol. 2, no. 3, pp. 1–11, 2019 (DOI: 10.46253/j.mr.v2i3.a1).

[40] –, https://www.kaggle.com/ming666/flicker8k-dataset.

[41] D. Songtao, *et al.*, "Image caption generation with high-level image features", *Pattern Recognition Letters 123*, pp. 89–95, 2019 (DOI: 10.1016/j.patrec.2019.03.021).

---

**Mukesh Kalla** is serving as Head & Assistant Professor in the Department of Computer Science and Engineering from the Sir Padampat Singhania University, Udaipur, India. He has received his Doctoral Degree in Computer Engineering in year 2017. He is having teaching and administrative experience of 19 years. He has about 24 research publications in international journals and conferences.

E-mail: mukesh.kalla@spsu.ac.in

Department of Computer Science and Engineering, Sir Padampat Singhania University, India

**Amit Jain** is currently working as Assistant Professor in Computer Science and Engineering Department, Sir Padampat Singhania University, Udaipur, India. He has completed his Doctoral Degree in Computer Engineering in 2016. He is having 25 years teaching and administrative experience and about 33 research publications in international journals and conferences.

E-mail: amit.jain@spsu.ac.in

Department of Computer Science and Engineering, Sir Padampat Singhania University, India

**Arvind Sharma** is presently working as Assistant Professor in Mathematics Department, Sir Padampat Singhania University, Udaipur, India. He has completed his Doctoral Degree in Mathematics in year 2017. He is having teaching and administrative experience of 17 years. He has about 10 research publications in international journals and conferences.

E-mail: sharma.arvind@spsu.ac.in

Department of Computer Science and Engineering, Sir Padampat Singhania University, India

**Roshni Padate** is Ph.D. scholar in Sir Padampat Singhania University, Udaipur, India and an Assistant Professor in the Department of Computer Engineering, Fr. Conceicao Rodrigues college of Engineering, Mumbai University. She has completed M.E. (Computer Engineering) from Mumbai University, India in 2010 and B.E. (Computer Science and Engineering) from Shri Sant Gajanan Maharaj College of Engineering Shegaon, Amravati University, India in 2000. Her area of research is image processing, artificial intelligence, machine learning, deep learning, and data mining.

E-mail: Roshni.padate@spsu.ac.in

Department of Computer Science and Engineering, Sir Padampat Singhania University, India

Paper

# An Efficient Hybrid Protocol Framework for DDoS Attack Detection and Mitigation Using Evolutionary Technique

T. Yerriswamy and Murtugudde Gururaj

*School of CSE, REVA University, India*

https://doi.org/10.26636/jtit.2022.165122

**Abstract — The ever-increasing use of the Internet has created massive amounts network traffic, causing problems related to its scalability, controllability, and manageability. Sophisticated network-based denial of service (DoS) and distributed denial of service (DDoS) attacks increasingly pose a future threat. The literature proposes various methods that may help stop all HTTP DoS/DDoS assaults, but no optimal solution has been identified so far. Therefore, this paper attempts to fill the gap by proposing an alternative solution known as an efficient hybrid protocol framework for distributed DoS attack detection and mitigation (E-HPFDDM). Such an architecture addresses all aspects of these assaults by relaying on a three-layer mechanism. Layer 1 uses the outer advanced blocking (OAB) scheme which blocks unauthorized IP sources using an advanced backlisted table. Layer 2 is a validation layer that relies on the inner service trackback (IST) scheme to help determine whether the inbound request has been initiated by a legitimate or an illegitimate user. Layer 3 (inner layer) uses the deep entropy based (DEB) scheme to identify, classify and mitigate high-rate DDoS (HR-DDoS) and flash crowd (FC) attacks. The research shows that in contrast to earlier studies, the structure of the proposed system offers effective defense against DoS/DDoS assaults for web applications.**

***Keywords — deep entropy based scheme, denial of service, distributed denial of service, flash crowd, high-rate DDoS, inner service trackback, outer advanced blocking***

## 1. Introduction

Availability of the network remains the most important security requirement. DDoS attacks may result in delaying access to resources, thus leading to network unavailability. The entire network or a single piece of network equipment (such as a switch) may be the target. On the other hand, it is the objective of DoS attacks to overwhelm the network with massive flows of data packets in order to render it inoperable or to degrade its performance [1]. A DoS/DDoS assault is both straightforward and effective, as the attack packets typically lack any precise characteristics that would help identify them.

Flash crowd (FC) is an increase in demand of a given service caused by authorized users making concurrent requests. FC saturates the server, resulting in a denial of service (DoS) assault that causes delays or takes down the site. From the perspective of the user's requests for the service, regardless of whether they are valid or not, FC may not be considered as an assault. It is considered an assault from the victim's or server-side perspective, as it negatively impacts the server's performance.

An LR-DDoS attack floods the victim machine with packets that have a low data rate in order to avoid detection by the existing anomaly-based intrusion detection techniques. Since LR-DDoS traffic is identical to normal flow of data, it can be concealed. Low-rate DDoS attacks usually rely on multiple low-rate assaults, such as those executed with the use of botnets.

This paper proposes an efficient hybrid protocol framework for DDoS detection and mitigation (E-HPFDDM), addressing all issues related to HTTP-based DoS/DDoS attacks. FC assaults are gradually reduced in effect by using the E-HPFDDM framework, which also immediately blocks high-rate DDoS. Furthermore, it protects web servers from assaults at several locations where data packets have ganged. The E-HPFDDM architecture performs the above tasks using a three-layer structure that efficiently detects, classifies and prevents HTTP DoS/DDoS attacks.

E-HPFDDM's first layer uses the OAB scheme and is configured at the edge router from which the packets are forwarded to the inner layers using applicable security measures. This scheme starts by matching incoming packets with the router's updated blacklist table. A packet is dropped if the incoming packet address matches the one on the blacklist and a notification message is sent to the user from the outer layer scheme. In other cases, the packets are moved to the next layer for further processing. This layer helps identify all the illegitimate addresses and directs the same to the outer layer in order to update the router's blacklist table. The two interconnected components make up the outer advanced blocking (OAB) strategy. In order to keep track of illegitimate IP sources, the updated blacklist table is used by IST and DEB schemes. If IP sources fail to pass the procedures defined in the initial component, they are rejected. The second element is a signaling strategy used by IST and DAE schemes to alert the OAB of potentially attacking IP sources, update the blacklist and block illegitimate IP sources during subsequent requests.

The second layer of E-HPFDDM framework uses the inner service traceback (IST) scheme that helps identify whether a given request has been initiated by a trusted user. It also

tracks the real IP source. The third layer of the HPFDDM framework uses the deep entropy based (DEB) scheme that helps detect anomalies and classifies DDoS attacks (e.g. HR-DDoS and FC attacks [2]). The DEB scheme uses the deep entropy-based algorithm for DDoS anomaly detection, classification, and mitigation. The DEB technique prevents traffic from entering the local network if it is determined to be a high-rate HTTP DoS/DDoS (HR-DDoS) attack. If the traffic is found to be of the FC type, the scheme lowers the connection timeout value, as per user requests, until the timeout value reaches zero, and simultaneously disables the HTTP connection's KeepAlive feature.

The contributions of this article include an experimental setup used for launching DDoS attacks, as well as an implementation of an efficient model that includes the strategies described above, such as defense mechanisms capable of detecting, preventing and mitigating DDoS attacks. This paper is organized as follows. Section 1 offers an overview of the E-HPFDDM framework. Section 2 presents the related works focusing on security protective frameworks, evolutionary techniques, DDoS detection, prevention, and attack mitigation techniques. Section 3 contains detailed information on the proposed E-HPFDDM technique. Section 4 presents results of an evaluation of the proposed framework based on ideal requirements applicable to a security system. Section 5 concludes the paper by presenting conclusions.

## 2. Literature Review

Ujjan *et al*. in [2] proposed a statistical model based on the entropy technique to analyze the flow of network traffic. Estimation of maximum entropy was proposed as a method for distinguishing between normal and malicious network traffic. Using a statistical-based entropy model, similar approaches to detecting DDoS attacks have been proposed. Dong *et al*. [3] proposed and characterized the various DDoS attacks in cloud and SDN networks. Lu *et al*. [4] proposed an efficient defense mechanism for DDoS attacks that identifies the intruder packets based on the entropy values and classifies, accordingly, the type of attack to be defended against.

Imran *et al*. [5], Alhijawi *et al*. [6] proposed an in-depth analysis of DoS prevention and mitigation techniques and divided them into three categories based on how malicious traffic was handled. The authors planned to develop a safe self-adaptive framework using machine learning-based techniques to detect DDoS assaults and retrieved network traffic properties based on the collected data.

Dong *et al*. [3] defined and characterized various DDoS attacks in cloud and in SDN networks, while Lu *et al*. [4] proposed an efficient defense mechanism for DDoS attacks that identifies the intruder packets based on the entropy values and accordingly classifies the type of attack to be defended against.

Imran *et al*. [5] and Alhijawi *et al*. [6] presented an in-depth analysis of DoS prevention and mitigation techniques and divided them into three categories based on how malicious traffic was handled. The aim was to develop a safe self-adaptive framework that uses techniques based on machine learning to detect DDoS assaults and retrieves network traffic properties based on the data collected. Cui *et al*. [7] and Mique *et al*. researched, in [8], an SVM-trained defense mechanism, known as cognitive entropy technique, that helps defend against DDoS attacks. In order to calculate the entropy value for the data packets flowing between the source and destination addresses, the method relied on the data flow table of the switch. Another concept of Sahay *et al*. [9], known as the ArOMA act, is a DDoS defense system automatically identifying assaults on centralized networks without human intervention. A unique NB classifier model was proposed for intrusion detection systems which are implemented using multi-agents that monitor network traffic and separate abnormal data from typical traffic. Another learning-driven detection mitigation (LEDEM) model used for detecting DDoS attacks has been proposed in [10]. It also relies on the machine learning technique [11].

Yerriswamy *et al*. [12] proposed a DDoS protection model that helps classify packets based on signatures in order to avoid DDoS attacks. The process is divided into four stages, namely signature extraction, signature classification, anomaly mitigation and signature reduction. The proposed simulation model used for DDoS attack mitigation used programmable commodity switches and the enhanced grey wolf optimizer (EGWO) for detection intrusions [13], [14]. Long *et al*. [15] presented a hybrid SSAE-SVM entropy-based model for DDoS attack detection and mitigation that analyzes the number of incoming packets on a concurrent basis. If the packet request from a single source exceeds the threshold value, the model will categorize it as an attack and such data will be skipped.

## 3. E-HPFDDM Framework

The literature survey showed that all current research concentrates on developing an efficient security model to adequately protect against all types of HTTP DoS/DDoS attacks. In general, the majority of studies failed to create a security model that is capable of defending against all DDoS assaults. In order to fill this gap, the proposed research provides an ideal security model that can protect against all of the mentioned threats. Furthermore, it should be able to track down and block the illegitimate IP sources as well. The proposed solution uses a security model (E-HPFDDM) that is integrated with evolutionary techniques capable of detecting, classifying, and mitigating all types of HTTP DoS/DDoS assaults.

The E-HPFDDM framework automatically stops network traffic that is entering a local network if it is determined to be a HR-DDoS attack. Otherwise, when the traffic is found to be an FC attack, the algorithm tries to reduce the connection timeout value. With the KeepAlive feature of the HTTP connection disabled, different layers of the framework work together and help protect the server from these HTTP-based DDoS assaults by executing their tasks in an efficient

manner. If all protocol tests have been passed, the packets are transferred to the next layer. If not, the packets are dropped. This process is repeated until data packets are delivered to the last layer.

E-HPFDDM is a three-layer architecture. The OAB scheme at the edge router makes up the first layer of the E-HPFDDM framework, while inner service traceback (IST) makes up the second layer. The deep entropy-based (DEB) layer is the third layer, as shown in Fig. 1.

The packets that are moved to the first layer of E-HPFDDM are compared with packets in the blacklist table and are dropped if they match with any of the packets in the blacklist table. Additionally, this research created the OAB shield, a protective component for AntiDDoS systems, using the OAB scheme. It combines the alerting method used by the IST shield and AntiDDoS subsystems to notify the OAB shield subsystem about the attacker's IP sources, so that the blacklist table may be updated periodically. The OAB system uses the blacklist database to store data packets identified by the IST scheme or DEB scheme, in the case of their failure. The other element is a signaling method that the IST and DEB schemes utilize to alert the OAB scheme about the attacking IP sources. By updating the OAB scheme's blacklist table, IST and DEB schemes ban these IP sources from serving requests in the future.

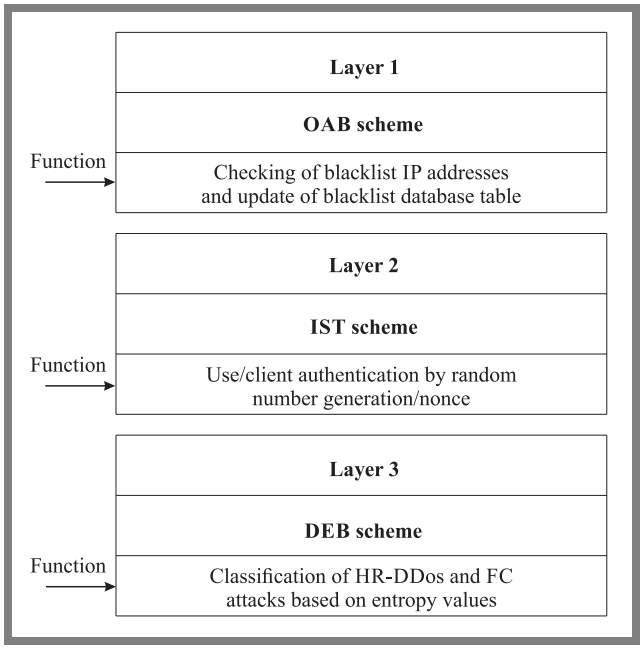



**Fig. 1.** Components of the E-HPFDDM framework.

E-HPFDDM's second layer uses the IST scheme that helps identify the whether the request has been initiated by a trusted user or an intruder, by validating the incoming requests. The first stage of the IST evaluates the incoming packets by scanning the header for distinct values which are only used by authorized requests. The IST scheme uses an authentication mechanism to produce a nonce value which is sent to the client or requester. The user must authenticate this nonce. If the authentication is successful, the data packet is moved to the next stage and the IP source is updated in the blacklist table.

If the intruder hacks the credentials of the victim and tries to enter the network as a trusted person, the model is trained to identify the intruder by relying on the particle swarm optimization (PSO) technique [16]–[20]. This technique allows each particle to learn its position and velocity from its neighbors.

Equation (1) represents the initial velocity, while Eqs. (2)-(5) represent the velocity of each particle that is updated based on the values of $X_{kd}(n)$ and $X_{jd}(n)$ for $n$ iterations, until the best solution is generated for both personal best solution $P_{kd}(n)$ and the global solution $g_d(n)$:

$$R_i d(n) = X_{\mathrm{kd}}(n) - X_{\mathrm{jd}}(n), \tag{1}$$

$$R_i d(n+1) = \mathrm{CV}_{\mathrm{id}}(n) + C[P_{\mathrm{id}}(n) - X_{\mathrm{id}}(n)], \tag{2}$$

$$R_i d(n+1) = \omega V_{\mathrm{id}}(n) + C_(1) r_(1)[X_{\mathrm{id}}(n) - P_{\mathrm{id}}(n)] + C_(2) r_(2)[X_{\mathrm{id}}(n) - g_d(n)], \tag{3}$$

$$R_i d(n+1) = R_{\mathrm{id}}(n) + C_r[P_{\mathrm{kd}}(n) - X_{\mathrm{id}}(n)], \tag{4}$$

$$R_i d(n+1) = R_{\mathrm{id}}(n) + 0.5C[P_{\mathrm{kd}}(n) - X_{\mathrm{id}}(n) + P_{\mathrm{id}}(n) - X_{\mathrm{id}}(n)], \tag{5}$$

where $C$ is the random value between 0 and 1, $X_{\mathrm{kd}}(n)$ and $X_{\mathrm{jd}}(n)$ are the particles that are randomly selected during $n$ iterations, and $P_{\mathrm{kd}}(n)$ and $g_d(n)$ are the personal and global best solutions, respectively.

The number of source and destination IP addresses and port numbers are used as input in the initial attack detection module that is based on entropy. The values used, associated with the number of source IP addresses $H_{\mathrm{sip}}$, destination IP addresses $H_{\mathrm{dip}}$, source port numbers $H_{\mathrm{sp}}$, and destination port numbers $H_{\mathrm{dp}}$, were calculated for the first consecutive z packets using Eqs. (6)-(8). Then, the next adjacent $z$ data packets are calculated using Algorithm 1.

**Algorithm 1:** Deep entropy-based algorithm for DDoS attack detection.

```
1: Input: user request's output: detection of DDoS
                                              attacks
2: Define the size of the window, Z
3: Collect sequence of data packets of size Z
4: H = {H_sip, H_dip,H_sp, H_dp}
5: If any data packet value of H exceeds
          the threshold value T, then indication
                                    of DDoS attack
6: End
```

Algorithm 1 provides a description of the module's computation procedure, while Algorithm 2 represents the deep entropy-based (DEB) technique. The DEB technique automatically stops network traffic that is entering the local network if it is determined to be HR-DDoS. Else, if the traffic is found to be an FC attack, the mechanism reduces the maximum connection timeout value, as per user requests, until the timeout value reaches zero and simultaneously disables the HTTP connection's KeepAlive feature.

The DEB scheme inspects the incoming data packet to detect HR-DDoS and FC assaults or, under regular conditions, by computing the entropy of all request as:

$$E = a \cdot \log_2 \frac{\text{number of uri counts}}{\text{total number of counts}}. \tag{6}$$

The specific source IP (sip) entropy values can be calculated using information entropy as:

$$E(\text{sip}_n) = \sum_{i=1}^{k} -p \cdot \text{sip}_n^i \cdot \log_2 p \cdot \text{sip}_n^i, \tag{7}$$

where $k$ is the number of different source IP addresses.

The number of packets $p_n$ with specific source address $n(\text{sip})$ for a given time $t$ can be calculated as:

$$E(\text{sip}_n) = \sum_{i=1}^{k} -\frac{n \cdot \text{sip}_n^i}{p_n} \cdot \log_2 n \cdot \frac{\text{sip}_n^i}{p_n}. \tag{8}$$

The DEB scheme uses the Algorithm 2 that compares the obtained entropy values to the maximum threshold for HR-DDoS and FC assaults. All the incoming packets request will be blocked that have the threshold values exceeding the maximum calculated entropy values, then signaling strategy used by the AST and DAE schemes to alert the OAB scheme by updating of blacklist database and blocking of illegitimate IP sources on subsequent requests. As a result, the incoming requests are treated as legitimate.

**Algorithm 2:** Deep entropy-based algorithm for DDoS attack classification and mitigation.

```
1: Input: user request's, output: type of DDoS
          attack detection,
          classification and mitigation
2: Wait for user request
3: If request is NULL return declined
4:   Enable AntiDDoS service of DEB scheme at
           layer 3 of E-HPFDDM model
5:   Enable modAntiDDoS service for DDoS assaults
           detection
6:   If requested IP address is from
          whitelisted table return OK
7:   If requested remote IP address is blacklisted
          reject the remote IP from edge router
8:     Set entropy_result = ok
9:   If entropy threshold matches with the
           threshold of AntiDDoS HR_DDoS
           return blacklisted table to search
           for remote IP and reject the remote
           IP from edge router
10:  If entropy threshold matches with
           the threshold of
                  AntiDDosFlashCrowd
11:    Repeat the steps 10 to 16
12:    Decease timeout and MaxKeepAlive requests
13:    If (timeout == 0 && KeepAlive == off)
                                    return
14:    update remote_ip blacklist table
15:    reject the remote_ip from edge router
16:  End if
```

The DEB scheme uses the Algorithm 2 that compares the obtained entropy values to the maximum threshold for HR-DDoS and FC assaults. All the incoming packets request will be blocked that have the threshold values exceeding the maximum calculated entropy values, then signaling strategy used by the AST and DAE schemes to alert the OAB scheme by updating of blacklist database and blocking of illegitimate IP sources on subsequent requests. As a result, the incoming requests are treated as legitimate.

### *3.1. Evaluation of the E-HPFDDM Framework*

Evaluation of the E-HPFDDM framework is performed during four different experimental tests of the AntiDDoS shield system. The first experiment covers the entry stage and relies on the OAB scheme. This will allow to check if illegitimate IP sources are blocked and if the blacklist is updated. The second experiment deals with the IST scheme that helps evaluate user authentication by deploying the nonce mechanism. The third and fourth experimental tests are done using the DEB scheme that helps detect and prevent HR-DDoS and FC assaults. The first and second experiments are conducted by simulating a large number of data packets that are both legitimate and illegitimate. An analysis has been performed to verify how effectively the model is classifying the data packets. The third and fourth experiments are conducted by launching a very large quantity of user requests, comparing the entropy threshold values and thereby analyzing how effectively the model is categorizing HR-DDoS and FC assaults. The experimental analysis has rendered positive results compared with existing security framework models. The simulation testbed is implemented using virtualization techniques allowing to share the resources during normal network usage [19], as illustrated in Fig. 2.

One of the following methods may be used to determine if an attack has occurred:

– Incoming data packets failing the validation tests performed by IST and DEB schemes. The incoming request must function according to its HTTP header parameters for the validation tests;
– The attacker launches an FC assault by flooding the web server with incoming data packets;
– The attacker launches a high-rate HR-DDoS assault by flooding the web application's cold web pages with many incoming requests;
– The incoming data packet requests are monitored by the edge router, so that the attacker's IP source can be blocked.

The simulation workflow is as follows:

– The client launches incoming data packet requests that are interpreted by the OAB scheme as requiring that the attacking source IP be blocked. The IST scheme authenticates the user by generation of nonce, and finally the DEB scheme detects and prevents HR-DDoS and FC assaults;
– The server responds to the particular incoming user requests after validating the distributed test plans;
– The OAB shield's router subsystem initially detects and analyses the IP source of incoming requests using the black-

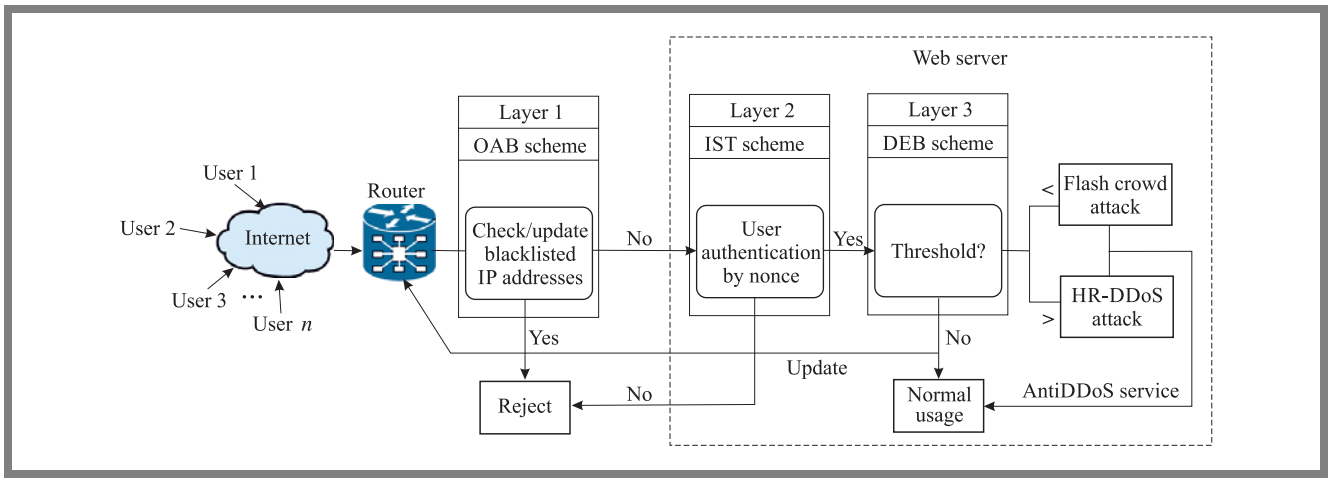



**Fig. 2.** Experimental setup and simulation model.

list table. The OAB shield immediately bans an attacking IP source by providing the "host unreachable" message to the requester. If not, a notification is sent to the server;

– The IST AntiDDoS system authenticates the incoming requests using the IST scheme and confirms that the request does not originate from an illegitimate user. This inbound request will move on to undergo a subsequent test within the IST shield subsystem, if it passes this human-launched test. If not, the IST shield subsystem instantly prevents it by sending the "HTTP forbidden" message to the requester. It must utilize the nonce that the IST shield subsystem gives back to address the issue appropriately.

The following AntiDDoS subsystem test will be run if the requester passes both tests. If not, the IST shield subsystem instantly prevents it by sending the "HTTP forbidden" message to the requester. Additionally, when the IST shield subsystem returns the "HTTP forbidden" message to the requester, it notifies the AOB shield subsystem, via signaling of the assaulting IP sources, thus helping update the blacklist table.

## 4. Results and Analysis

Evaluation of the E-HPFDDM framework is based on ideal requirements that a security system must comply with to defend against all categories of HTTP-based DoS and DDoS assaults. Based on the simulations, we may conclude that the E-HPFDDM framework's DEB scheme helps in categorizing HR-DDoS and FC assaults and, hence, helps protect the web server through the deep entropy-based algorithm capable of detecting, classifying and mitigating DDoS attacks. The experimental analysis has shown that the AntiDDoS module using the DEB scheme has detected and prevented 389,766 out of 410,000 FC attacks, thus being more efficient than existing security frameworks.

The E-HPFDDM framework's outer advanced blocking (OAB) scheme is quite capable of blocking attacker IP sources at the point of entry from the edge router. The experimental analysis has shown that the OAB scheme was successful in detecting and preventing all 410,000 attacking IP sources.

The HPFDDM framework's IST scheme is capable of validating an incoming request. The IST scheme determines whether a packet is legitimate and then passes the former while blocking the latter. It also includes a mechanism for tracing back and determining the true attacker IP source. The IST scheme uses the mechanism of nonce randomization and sends the packet to a user (rather than a bot) for authentication. The browser will authenticate the user using the nonce without any human action.

The HPFDDM framework moves the incoming data packets to the various layers and in each layer the packets have to be validated with various schemes (procedures) to defend against DDoS assaults, making the framework a collaborative, layered DDoS prevention solution. The layers of the framework collaborate and different schemes validate incoming data packets with special tests. Layer 2 helps provide nonce and authenticate the user. Similarly, the subsequent layer's DEB scheme performs the validation process with the use of special tests, such as entropy calculation, and accordingly classifies the type of DDoS assaults and attempts to mitigate the assaults accordingly.

The HPFDDM framework is easy to design and implement (this applies to all three layers) and the functionalities of each layer have been efficiently and independently distributed with proper coordination among the layers. The HPFDDM framework utilizes less bandwidth overhead when compared to existing network security solutions.

The HPFDDM framework supports the functionalities of all the layers, e.g. user authentication is performed in the application layer, using the nonce technique, and the detection and classification of DDoS assaults is performed in the network layer (e.g. HR-DDoS FC attacks).

The HPFDDM framework is integrated with evolutionary techniques, making it easier for the solution to adapt update dynamically, if needed. The IST scheme helps in updating the blacklist table when a non-legitimate IP address is detected and the DEB scheme helps in entropy calculations, based on which the DDoS attack table is updated if the DEB scheme detects HR-DDoS and FC attacks.

The designed HPFDDM framework that can support the evolutionary techniques like PSO that are integrated with DEB schemes which helps in DDoS attack detection and mitigating.

The individual layers of the E-HPFDDM framework (OAB, IST, and DEB) use very little memory for storage. The OAB layer of the E-HPFDDM framework uses much less memory for storing the blacklist table, whereas the IST layer of the E-HPFDDM framework consumes no memory at all, as all its functionalities are performed in real time. In addition, E-HPFDDM framework's DEB layer uses very little memory for storing web page-related information.

E-HPFDDM framework is resistant to spoofing assaults, as the IST scheme examines the incoming request's headers and nonce values to determine whether the requester is legitimate or illegitimate. If the requester fails these two tests, the user will be blocked immediately and the blacklist table will be updated.

## 5. Conclusion and Future Outlook

This study has proposed and developed an efficient hybrid protocol framework for DDoS attack detection and mitigation (E-HPFDDM), utilizing evolutionary techniques, and a unique efficient hybrid protection framework for defending against HTTP-based DoS/DDoS attacks. The unique architecture of the proposed framework eliminates all the flaws of the previous similar studies. It offers an innovative defense mechanism to safeguard online applications against all types of HTTP DoS/DDoS assaults, including high-rate HR-DDoS and FC. Additionally, it has a good ability to verify and track attacker IP sources and block them immediately, simultaneously updating the blacklist table. The E-HPFDDM framework is then assessed using the best guidelines applicable to defense systems that shield against all types of HTTP-based DoS and DDoS assaults. The simulations performed demonstrated that the E-HPFDDM framework was effective in fully meeting the applicable requirements. As the framework was unable to identify and defend against all FC assaults, a low rate of false negatives was experienced.

## References


[1] A. Saravanan, S.S. Bama, S. Kadry, and L.K. Ramasamy, "A new framework to alleviate DDoS vulnerabilities in cloud computing", *International Journal of Electrical & Computer Engineering*, vol. 9, no. 5, pp. 4163–4175, 2019 (DOI: 10.11591/ijece.v9i5.pp4163-4175).

[2] R.M. Ujjan, Z. Pervez, K. Dahal, W.A. Khan, A.M. Khattak, and B. Hayat, "Entropy based features distribution for anti-DDoS model in SDN", *Sustainability*, vol. 13, no. 3, pp. 1–27, 2021 (DOI: 10.3390/su13031522).

[3] S. Dong, R. Jain, and K. Abbas, "A Survey on Distributed Denial of Service (DDoS) Attacks in SDN and Cloud Computing Environments", *IEEE Access*, vol. 7, pp. 80813–80828, pp. 1–1, 2019 (DOI: 10.1109/ACCESS.2019.2922196).

[4] A. Alshamrani, A. Chowdhary, S. Pisharody, D. Lu, and D. Huang, "A defense system for defeating DDoS attacks in SDN based networks", *In proceedings of the ACM International Symposium on Mobility Management and Wireless Access*, pp. 83–92, 2017 (DOI: 10.1145/3132062.3132074).

[5] M. Imran, M.H. Durad, F.A. Khan, and A. Derhab, "Toward an optimal solution against denial of service attacks in software defined networks", *Future Gener. Comput. Syst.*, vol. 92, pp. 444–453, 2019 (DOI: 10.1016/j.future.2018.09.022).

[6] A. Bushra, A. Sufyan, E. Hany, B.S. Haythem, and A. Moussa, "A survey on DoS/DDoS mitigation techniques in SDNs: Classification, comparison, solutions, testing tools and datasets", *Computers & Electrical Engineering*, vol. 99, 2022 (DOI: 10.1016/j.compeleceng.2022.107706).

[7] J. Cui, M. Wang, Y. Luo, and H. Zhong, "DDoS detection and defense mechanism based on cognitive-inspired computing in SDN", *Future Generation Computer Systems*, vol. 97, 2019 (DOI: 10.1016/j.future.2019.02.037).

[8] M.A. Naagas, E.L. Mique, T.D. Palaoag, and J.S.D. Cruz, "Defense-through-deception network security model: Securing university campus network from DOS/DDOS attack", *Bulletin of Electrical Engineering and Informatics*, vol. 7, no. 4, pp. 593–600, 2018 (DOI: 10.11591/eei.v7i4.1349).

[9] R. Sahay, G. Blanc, Z. Zhang, H. Debar, "ArOMA: an SDN based autonomic DDoS mitigation framework", *Computers & Security*, vol. 70, pp. 482–499, 2017 (DOI: 10.1016/j.cose.2017.07.008).

[10] A. Mehmood, M. Mukherjee, S.H. Ahmed, H. Song, and K.M. Malik, "NBC-MAIDS: Naive Bayesian classification technique in multi-agent system-enriched IDS for securing IoT against DDoS attacks", *The Journal of Supercomputing*, vol. 74, no. 10, pp. 5156–5170 2018 (DOI: 10.1007/s11227-018-2413-7).

[11] N. Ravi and S.M. Shalinie, "Learning-driven detection and mitigation of DDoS attack in IoT via SDN-cloud", *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3559–3570, 2020 (DOI: 10.1109/JIOT.2020.2973176).

[12] T. Yerriswamy and M. Gururaj, "Signature-based Traffic Classification for DDoS Attack Detection and Analysis of Mitigation for DDoS Attacks using Programmable Commodity Switches", *International Journal of Performability Engineering*, vol. 18, no. 7, pp. 529–536, 2022 (DOI: 10.23940/ijpe.22.07.p8.529536).

[13] T. Yerriswamy and M. Gururaj, "An Efficient Algorithm for Anomaly Intrusion Detection in a Network", *Global Transitions Proceedings*, vol. 2, 2021 (DOI: 10.1016/j.gltp.2021.08.066).

[14] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques", *International Journal of Distributed Sensor Networks*, vol. 13, 155014771774146, 2017 (DOI: 10.1177/1550147717741463).

[15] L. Zhang and J. Wang, "A hybrid method of entropy and SSAE-SVM based DDoS detection and mitigation mechanism in SDN", *Computers & Security*, vol. 115, 102604, 2022 (DOI: 10.1016/j.cose.2022.102604).

[16] Kamel Hasan and Abdullah Mahmood, "Distributed denial of service attacks detection for software defined networks based on evolutionary decision tree model", *Bulletin of Electrical Engineering and Informatics*, vol. 11, pp. 2322–2330, 2022, (DOI: 10.11591/eei.v11i4.3835).

[17] T. Islam, *et al.*, "A Socio-Technical and Co-evolutionary Framework for Reducing Human-Related Risks in Cyber Security and Cybercrime Ecosystems", *G. Wang, M.Z.A. Bhuiyan S. De Capitani di Vimercati, Y. Ren (eds), Dependability in Sensor, Cloud, and Big Data Systems and Applications. DependSys 2019. Communications in Computer and Information Science*, vol. 1123, 2019 (DOI: 10.1007/978-981-15-1304-6_22).

[18] T. Yerriswamy and M. Gururaj, "Study of evolutionary techniques in the field of network security", pp. 594–598, 2020 (DOI: 10.1109/ICSTCEE49637.2020.9277 082).

[19] S. Supreeth and K.K. Patil, "Hybrid Genetic Algorithm and Modified-Particle Swarm Optimization Algorithm (GA-MPSO) for Predicting Scheduling Virtual Machines in Educational Cloud Platforms", *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 7, pp. 208–225, 2022 (DOI: 10.3991/ijet.v17i07.29223).

[20] A. Pradhan, S.K. Bisoy, and A. Das, "A Survey on PSO Based Meta-Heuristic Scheduling Mechanism in Cloud Computing Environment", *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 8, pp. 4888–4901, 2021 (DOI: 10.1016/j.jksuci.2021.01.003).

**T. Yerriswamy** received B.E. degree in Information Science Engineering from VTU, Belagavi Karnataka, India in 2008 and M. Tech. degree in CNE from VTU in 2011. Now he is pursuing Ph.D. under VTU and working as Assistant Professor in school of CSE, REVA University, Bengaluru, India. His research interests include information and network security and soft computing.

E-mail: yssvce2123@gmail.com

School of CSE, REVA University, India

**Murtugudde Gururaj** received B.E. degree in Computer Science Engineering in 2000, M.Tech degree in IT in 2011 and Ph.D. in CSE in 2014. Currently working as Professor in school of CSE, REVA University, Bengaluru, India. His research interests include information and network security, big data analytics.

E-mail: gururajmurtu@gmail.com

School of CSE, REVA University, India

Paper

# Convolutional Neural Networks for *C. Elegans* Muscle Age Classification Using Only Self-learned Features

Bartosz Czaplewski[1], Mariusz Dzwonkowski[1,2], and Damian Panas[3]

[1]*Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Department of Teleinformation Networks, Poland,*
[2]*Medical University of Gdańsk, Faculty of Health Sciences, Department of Radiology Informatics and Statistics, Poland,*
[3]*Institute of Animal Reproduction and Food Research of the Polish Academy of Sciences, Molecular Biology Laboratory, Poland*

https://doi.org/10.26636/jtit.2022.165322

**Abstract — Nematodes *Caenorhabditis elegans* (*C. elegans*) have been used as model organisms in a wide variety of biological studies, especially those intended to obtain a better understanding of aging and age-associated diseases. This paper focuses on automating the analysis of *C. elegans* imagery to classify the muscle age of nematodes based on the known and well established IICBU dataset. Unlike many modern classification methods, the proposed approach relies on deep learning techniques, specifically on convolutional neural networks (CNNs), to solve the problem and achieve high classification accuracy by focusing on non-handcrafted self-learned features. Various networks known from the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) have been investigated and adapted for the purposes of the *C. elegans* muscle aging dataset by applying transfer learning and data augmentation techniques. The proposed approach of unfreezing different numbers of convolutional layers at the feature extraction stage and introducing different structures of newly trained fully connected layers at the classification stage, enable to better fine-tune the selected networks. The adjusted CNNs, as featured in this paper, have been compared with other state-of-art methods. In anti-aging drug research, the proposed CNNs would serve as a very fast and effective age determination method, thus leading to reductions in time and costs of laboratory research.**

***Keywords — biomedical imaging, C. elegans muscle aging, convolutional neural networks, deep learning, machine learning***

## 1. Introduction

Aging is a complex process in which many changes take place simultaneously at every level of a biological organization - from organelles, cells, tissues, to the entire systemic environment. Naturally, aging and age-related diseases are studied primarily with human health and longevity in mind, but model organisms, e.g. *Caenorhabditis elegans* (*C. elegans*), may be successfully employed for those purposes as well [1]–[5]. *C. elegans*, as used in this study, is a free-living, 1 mm long nematode that feeds on bacteria. Its clear, age-dependent physiological changes, short lifespan of approx. 17–20 days, ease of maintenance in a laboratory setting and high genetic homology with humans make the worm a widely acknowledged system for aging research [6], [7].

*C. elegans* exhibits many age-associated changes [8], however deterioration of its muscle tissue draws particular attention. Since the nematode's body-wall muscle is analogous to human skeletal muscle, some disorders, especially those of particular concern to the public health, can be effectively studied with the use of this model. For example, due to its progressive loss of muscle mass, *C. elegans*, just like some elderly humans, suffers from sarcopenia. This condition has dire consequences for many activities, such as locomotion and ingestion, but surprisingly little is known regarding its cures and actual causes [9]–[11].

Extensive research has been performed on anti-aging drugs based on measuring the age of *C. elegans* and determining the effects that medication has at early stages of the aging process. The weakness of such studies has always been that the age of *C. elegans* was determined by human experts and, thus, the results were subjective and strongly dependent on the expert's experience. Gaining new insights in this domain is therefore potentially of great value, although succeeding advances might require increasingly sophisticated, state-of-the-art tools, such as convolutional neural networks (CNNs) [12]–[14]. This machine learning technique has recently shown some remarkable achievements and great potential in solving problems related to image classification, visual signal processing, biological and medical imaging, computer vision, and image recognition years [15]–[19]. In addition to being capable of outperforming medical diagnosis experts, machine learning-based tools, such as CNNs, may also significantly reduce the time and costs of such an evaluation. Hence, further explorations in this field are essential.

This paper shows an innovative solution to the problem of classifying *C. elegans* muscle aging using CNNs. The contributions of this paper are as follows.

– An innovative approach consisting in using a convolutional neural network for the purpose of *C. elegans* muscle aging classification was presented. New CNN models were obtained through fine-tuning at the feature selection stage,

and the introduction of new structures at the classification stage;

– A fully CNN-based solution was proposed which achieved the highest level of classification accuracy. The proposed solution relies solely on non-handcrafted, self-learned features, in contrast to the existing literature;
– The ability to use various networks known from the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) was investigated. New and different transfer learning strategies were applied. The effect that unfreezing of different numbers of convolutional layers during the transfer learning phase has on the overall classification accuracy was investigated. This step is usually ignored in other publications. The effect of different classifier structures was investigated and the fully connected layers were modified accordingly;
– Comparison of the proposed method with other methods used in tackling the same problem was conducted using the same or similar datasets.

The article is structured as follows. Related works are presented in Section 2. The proposed methodology is explained in Section 3. Section 4 describes experimental results, including the CNN structures, learning parameters, classification accuracy level obtained, and a comparison with other methods. Conclusions are given in Section 5.

## 2. Related Work

Image processing technologies are commonly incorporated into many studies to automate the measurement of age of *C. elegans*. A training-based approach relying on tracking the swimming motion of the worm was examined by Restif and Metaxas [20]. Johnston *et al.* devised a pattern recognition tool that assesses age using the pharynx pumping rate [21]. Machine learning techniques, more specifically a CNN trained for regression, that utilized the relationship between age and level of body bend, were proposed by Lin *et al.* in [22].

Many research programs described in the literature that investigate the topic rely on experiments with self-generated datasets and, therefore, preclude comparative analysis. However, some studies (as listed below) report their results based on data obtained from the IICBU Biological Image Repository, including, inter alia, a dataset concerned with *C. elegans* muscle aging.

An open-source utility classifier for biological image analysis called Wndchrm was proposed by Shamir *et al.* in [23], [24]. The classifier extracts image content descriptors (1025 image features including polynomial decompositions, high contrast features, pixel statistics, and textures) from the raw image. Feature extraction is also performed on image transforms, such as Fourier, wavelet and Chebyshev, as well as on compound transforms being different combinations of subsequent image transforms. The most informative features are selected and the feature vector of each image is used for classification and similarity measurement.

A BIOimage Classification and Annotation Tool (BIOCAT) was proposed by Zhou *et al. in* [25]. It allows automatic classification and annotation of entire 2- and 3-dimensional biological images or their specific regions of interest by utilizing pattern recognition algorithms. BIOCAT is equipped with approximately 20 built-in modularized algorithms of feature extractors that can be chained in a customizable way to form an appropriate model for image classification. The chains may be compared via BIOCAT to determine the most suitable model for the specific dataset.

Siji *et al.* enhanced the performance of a content-based biological image retrieval system by selecting discriminative feature sets from the set of canonical features [26]. Canonical features are extracted using Wndchrm [23] and are assigned to 4 separate feature sets. Each set is optimized using the principal component analysis (PCA) process and the Fisher score for the purpose of feature selection. The selected features are then used for training the support vector machine (SVM) and Bayesian classifiers.

Bioimage classification utilizing the SVM classifier was proposed by Song *et al.* [27]. In this approach, a high-dimensional multi-modal descriptor was introduced to combine multiple texture features. To enhance the discriminative power of the descriptors, a subcategory discriminant transform (SDT) was designed to transform the descriptors before performing SVM classification. Similarly to linear discriminant analysis (LDA), the descriptor transform obtained with the use of the SDT algorithm aims to minimize the within-class variation and maximize the between-class difference.

Sekhar and Mohan proposed another SVM approach based on distance metric learning techniques for semi-supervised pattern classification [28]. The authors assigned label information to the selected group of images, creating a sparingly labeled training set. The optimal kernel gram matrix was determined by using pairwise similar/dissimilar constraints derived from the available examples in the labeled training set. Label information for the remaining unlabeled examples in the entire dataset is determined in an iterative self-training manner, utilizing SVM. The method uses a confidence measure in the decision-making process.

Nanni *et al.* proposed a method that combines a CNN with an SVM classifier [29] to obtain a structure capable of synthesizing a large number of different image classification tasks. Features extracted from the deep layers of the CNNs are mixed with more traditional hand-crafted features. The output of each layer is treated as a feature vector for which a dimensionality reduction method, such as PCA or discrete cosine transform (DCT), can be applied, depending on the size of the vector. All of the vectors are then processed by SVM to provide the final output. This approach was further improved in [30], where a combination of local features, dense sampling features, and deep learning features was proposed. Each descriptor was used to train a different SVM, which were then combined with the sum rule. In 2020, Nanni and Ghidoni proposed an alternative approach [31] that boosts the performance of trained CNNs by composing multiple CNNs into an

ensemble and combining scores based on the sum rule. The method, dedicated in general for bioimage datasets, features the combination of multiple descriptors based on different feature types, both of the learned and handcrafted variety.

In addition to the above-mentioned approaches relied upon in the classification of nematodes, there are also numerous novel works that demonstrate the feasibility of incorporating deep learning technology into radiology practice. Ahn *et al.* proposed a mammographic density estimation [32] process based on CNN trained with local and global statistics extracted from the image set, whereas Li *et al.* showed that pretraining may improve the effectiveness of deep learning-based tissue microstructure estimation [33]. This property is particularly evident with abundant, publicly available high-quality datasets, as is the case with the dMRI scans used by the authors.

In this paper, we propose a novel approach to classifying muscle aging in *C. elegans* with the use of a CNN. To the best of our knowledge, this paper is the first work to achieve classification accuracy at a level comparable with (or, in some cases, better than) state-of-the-art methods relying solely on different fine-tuned CNN models, without the need to introduce additional machine learning classifiers, such as linear a SVM (as proposed, for instance, in [29] and [30]), additional metadata or a description of the input images based on expert knowledge.

## 3. Proposed Methodology

The goal of this paper is to classify *C. elegans* muscle aging based on microscope imaging, using self-learned features only. CNNs are a perfect tool for the considered classification problem, due to their appropriable characteristics. For example, there is no need to design an analytical feature selection model which usually has to be based on expert knowledge. On the other hand, however, a CNN analyzes local spatial properties in the patches of meaningful pixels, which reflects the human interpretation of input data.

The experiments were conducted based on the *C. elegans* muscle aging dataset [34], [35] compiled by the Laboratory of Genetics at the U.S. National Institute on Aging. The dataset consists of images of muscles taken with a fluorescence microscope, with the said images representing various ages of nematodes. The individuals are genetically identical and live in a controlled laboratory environment. The images are saved in the TIFF format, in grayscale, with a resolution of $1600 \times 1200$ pixels and a depth of 16 bits. The size of the dataset is 237 observations. The images are unevenly divided into 4 classes: "day 1" (52 observations), "day 2" (48 observations), "day 4" (95 observations), and "day 8" (42 observations). Examples of such images are presented in Fig. 1.

In the proposed solution, the input has the form of a grayscale image of a *C. elegans* muscle with resolution of $1600 \times 1200$ pixels and a depth of 16 bits. Due to the structure of the CNNs used for transfer learning, the pixel values from a single-color channel are duplicated to create a 3-channel image. In general, the image can be of any resolution, as it is scaled down to a fixed size corresponding to the input layer of the CNN. The output is a decision about the age of the muscle shown in the image, i.e. the image is classified into one of the four age classes: "day 1", "day 2", "day 4", "day 8". The flowchart for the proposed method is shown in Fig. 2.

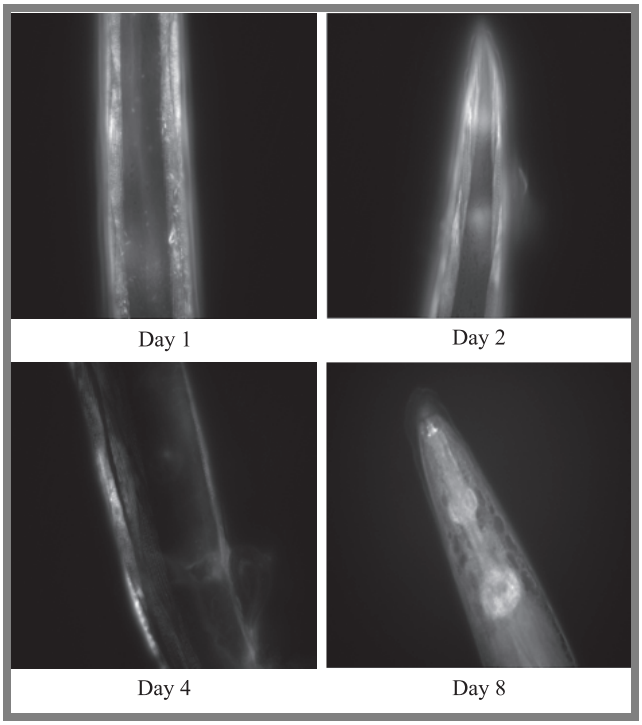



**Fig. 1.** Examples from the *C. elegans* muscle aging dataset [34], [35].

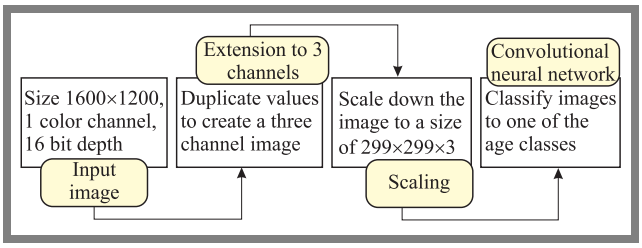



**Fig. 2.** Flowchart of the proposed approach.

The research presented in this paper can be divided into the following tasks. The first part consisted in selecting the best base network for transfer learning. Designing a new CNN for the classification problem under consideration from scratch would be very time-consuming and ultimately unnecessary. The literature has shown that it is possible to retrain an image classification CNN that has already been trained to extract powerful and informative features from natural images (usually based on ImageNet dataset) to become a classification CNN for a specific medical problem. Based on the results shown in the references [15], [22], three candidates for the base (starting point) network for transfer learning were selected: VGG16, InceptionV3, and InceptionResNetv2.

The second task was to select the best strategy for unfreezing the layers of the base network during transfer learning. The classical approach used in transfer learning is to freeze the weights and biases of the convolutional layers of the

base network, so that they are not retrained at all and only allow the layers responsible for classification to be retrained. During the research, it was found that this is not optimal. Instead, unfreezing a number of layers, both of the convolutional and fully connected variety, significantly increased the classification accuracy of the learned network.

The third task was to investigate how the base network's architecture could be modified to achieve even better results. Since the part of the network corresponding to feature selection should not be changed, as this would mean the rejection of all of the benefits from transfer learning, the main focus was on changes in the part responsible for the classification. Several new classification stage models were considered and the results achieved by the best of them are presented below.

The fourth part was to compare the proposed CNN with other prominent solutions available in the literature. The accuracy of the classification was a comparative criterion.

# 4. Experimental Results

## *4.1. Network Structures*

The authors decided to rely on the transfer learning technique in their research, as it is much faster and easier than training a network from scratch. It involves the use of an already pre-trained network for general image classification and, relying on its ability to recognize powerful and informative features from an image, training it for a new, much more specific classification task. For this purpose, three CNNs were selected: VGG16 [36], InceptionV3 [37], and InceptionResNetV2 [38]. In this paper, they are referred to as base networks. All three networks had been pretrained on the ImageNet dataset [39] which contains over a million images of natural objects, and are capable of classifying the said images into one of 1000 labels, such as goldfish, coffee mug, pencil, etc. These networks have already learned an effective set of image features, but the goal is to fine-tune the base network, so it can learn features specific to the new dataset and the new set of labels. A comparison of the base networks is given in Tab. 1.

**Tab. 1.** Comparison of the base networks for transfer learning.

| Network | Depth (layers) | No. of parameters | Image size |
|---|---|---|---|
| VGG16 | 16 | $144 \cdot 10^6$ | $224 \times 224 \times 3$ |
| InceptionV3 | 48 | $23.9 \cdot 10^6$ | $299 \times 299 \times 3$ |
| InceptionResNetV2 | 164 | $55.0 \cdot 10^6$ | $299 \times 299 \times 3$ |

The easiest approach to transfer learning is as follows. One should adopt the architecture of the base network, set the initial pre-trained values of weights and biases, remove the last layer of the classifier and replace it with a new, fully connected layer with the number of neurons equal to the new number of classes, freeze all convolution layers (set the learning rate factors to 0), speed up learning for the fully connected layers (set the learning rate factors to e.g. 10), and then start retraining using the new dataset. In this way, the deep network designer assumes that the starting network has already acquired adequate knowledge for the feature selection stage, suitable for solving the new problem, so the learned values in the convolutional layers would not be changed. A designer could make this decision if the new dataset for the new problem is very similar to the original dataset for which the underlying network was trained. Changing the last fully connected layer at the classification stage is the minimum condition that need to be fulfilled to teach the network to solve a new problem. Alternatively, instead of freezing the layers completely, one can just slow down their learning process, i.e. set the learning rate factor to 1, while setting a very low initial learning rate in the training options (e.g. 0.0001). In this way, the developer allows for minor changes to the feature selection stage.

Unfortunately, such an approach produces results that are not satisfactory for the purpose of our research. This is mainly because the dataset differs significantly from the ImageNet dataset. Unfortunately, such a straightforward approach can be found throughout the literature. Therefore, fine-tuning is essential. In this research, two approaches to fine-tuning were applied: firstly, we unfroze a number of the convolutional layers at the feature extraction stage, and secondly, we introduced different structures of newly trained, fully connected layers at the classification stage.

When selecting layers to unfreeze, one should not select any of the initial layers. The initial convolutional layers are used to identify simple, general patterns in an image, e.g. edges. Even if the new dataset used for transfer learning to solve the new problem is significantly different from the original dataset used for the base network, we still want the network to be able to recognize such elementary features in the image. Moreover, allowing the re-learning of the initial layers will render the knowledge in the subsequent layers useless due to the alternation of the output from the initial layers. On the other hand, deeper convolutional layers are used to identify complex features specific to a given dataset, based on previously identified simple patterns. In the case concerned, both the images and

**Tab. 2.** General information concerning network structures used.

| | CNN#1 | CNN#2 | CNN#3 |
|---|---|---|---|
| Depth | 17 | 49 | 165 |
| Base CNN | VGG16 | InceptionV3 | Inception-ResNetV2 |
| Input size | $224 \times 224 \times 3$ | $299 \times 299 \times 3$ | $299 \times 299 \times 3$ |
| Frozen conv. layers | 12 | 89 | 238 |
| Unfrozen conv. layers | 1 | 9 | 6 |
| Fully connected layers | 4 | 2 | 2 |
| No. of parameters | 150 527 244 | 30 211 936 | 60 648 676 |

**Tab. 3.** Detailed network structures in relation to base networks.

| | CNN#1 | CNN#2 | CNN#3 |
|---|---|---|---|
| Frozen feature extraction part | – Convolutional layers from "conv1_1" to "conv5_2":<br>• number and sizes of filters as in VGG16,<br>• bias and weight learn rate factors set to 1,<br>• pre-trained biases and weights initialized<br>– ReLU activation and max pooling layers as in VGG16 | – Convolutional layers from "conv2d_1" to "conv2d_89":<br>• number and sizes of filters as in InceptionV3,<br>• bias and weight learn rate factors set to 0,<br>• pre-trained biases and weights initialized<br>– Batch normalization, ReLU activation and max pooling layers as in InceptionV3 | – Convolutional layers from "conv2d_1" to "conv2d_199", from "block17_1_conv" to "block17_20_conv", from "block35_1_conv" to "block35_10_conv" from "block8_1_conv" to "block8_9_conv":<br>• number and sizes of filters as in InceptionResNetV2,<br>• bias learn rate factor set to 0, weight learn rate factor set to 1,<br>• pre-trained biases and weights initialized<br>– Batch normalization, ReLU activation, max pooling, addition, scaling, and depth concatenation layers as in InceptionResNetV2 |
| Unfrozen feature extraction part | – Convolutional layer "conv5_3":<br>• number and sizes of filters as in VGG16,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– ReLU activation and max pooling layer as in VGG16 | – Convolutional layers from "conv2d_86" to "conv2d_94":<br>• number and sizes of filters as in InceptionV3,<br>• bias and weight learn rate factors set to 10,<br>• pre-trained biases and weights initialized.<br>– Batch normalization, ReLU activation and max pooling layer as in InceptionV3 | – Convolutional layers from "conv2d_200" to "conv2d_203", "block8_10_conv", and "conv_7b" layers:<br>• number and sizes of filters as in InceptionResNetV2,<br>• bias and weight learn rate factors set to 10,<br>• pre-trained biases and weights initialized.<br>– Batch normalization, ReLU activation and max pooling layer as in InceptionResNetV2 |
| Classification part | – Fully connected layers "fc6" and "fc7":<br>• 4096 neurons,<br>• bias and weight learn rate factors set to 1,<br>• pre-trained biases and weights initialized<br>– ReLU activation and 0.5 dropout layer<br>– Fully connected layers "new_fc8":<br>• 4096 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– ReLU activation and 0.5 dropout layer<br>– Fully connected layer "new_fc9":<br>• 4 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– Softmax and classification layer | – Global average pooling layer<br>– Fully connected layer "new_fc1":<br>• 4096 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– ReLU activation and 0.5 dropout layer<br>– Fully connected layer "new_fc2":<br>• 4 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– Softmax and classification layer | – Global average pooling layer<br>– Fully connected layer "new_fc1":<br>• 4096 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– ReLU activation and 0.5 dropout layer<br>– Fully connected layer "new_fc2":<br>• 4 neurons,<br>• bias and weight learn rate factors set to 10,<br>• biases initialized as zeros and weights initialized randomly<br>– Softmax and classification layer |

their associated labels are fundamentally different from the images and labels in ImageNet, so we expect that the network will learn to recognize completely different complex features. This means that many deep convolutional layers should be unfrozen and, furthermore, they should be able to change their values rapidly during the training process. This means that if we want to unfreeze convolutional layers (set the learning rate factors to 10, for instance) and allow them to be re-trained, we have to do it from the end of the network, starting with the layers immediately preceding the fully connected part. In this study, the number of unfrozen convolutional layers ranged from 1 to 9, depending on the network.

In terms of the fully connected part of the network, the starting point was always the classification stage model established for the considered base network. Furthermore, the CNNs were also tested using other classification stage models, e.g. the

five models proposed in [15] were tested. Modifications of fully connected layers at the classification stage are motivated by the fact that the set of labels is completely different from that of the core network, both in meaning and in number. However, it has been noticed that the best results are usually obtained by extending the fully connected part with just one new fully connected layer with a large number of neurons immediately before the last classification layer, which has to be also a new layer with the number of neurons corresponding to the new number of classes.

As the last step in selecting the network structures, the influence of other network parameters was tested, e.g. activation layers were changed from ReLU layers to clipped ReLU layers, leaky ReLU layers or tanh layers, the training parameters were changed, etc. The impact of these changes was negligible and is, therefore, not discussed in this paper.

For all three base networks, hundreds of training sessions were run with various numbers of unfrozen layers as well as with various classifier models. In Tabs. 2 and 3, and throughout the paper, only the best obtained network structures for each of the base networks are presented.

### *4.2. Training Options*

All of the CNNs were trained via supervised learning. For estimating performance of the trained models, the 5-fold cross-validation procedure was used. The dataset was randomly split in such a way that 60% of it was used for the training set, 20% for the validation set, and 20% for the test set. The training data was shuffled before each epoch, and the validation data was shuffled before each validation. Both shuffling and data augmentation were applied to reduce the variance in results and ensure that the model is general and not overfitted.

To avoid overfitting, CNNs rely heavily on the large datasets available. Overfitting occurs when the CNN learns to model the training data too perfectly, such that the learned model is characterized by a high variance for the test data. In analyzing medical images, big datasets are usually not available. With such a small dataset, the risk of overfitting is very high. In such a case, data augmentation [40] is essential to prevent overfitting by introducing some random geometrical transformations into the training data for each epoch. In this way, one randomly augmented version of each image is used during each training epoch. In the course of the research, the data augmentation settings were based on the data augmentation models presented in [15], which deals with a similar classification problem. The applied data augmentation consists of the following geometric transformations: random rotation in the 0–30 range, random uniform scaling with the 1–1.05 factor, random horizontal and vertical translations by 0 to 10 pixels, random horizontal and vertical reflections with a probability of 0.5 for each direction.

The goal of shuffling is to ensure that training set and validation sets are representative of the overall distribution of the data. This is important when using the minibatch gradient descent algorithm in which the minibatch must be representative of the overall dataset. Shuffling after each epoch helps avoid the risk of bad batches even further. This causes each image in the training set to have an independent impact on the model without being biased by the previous images.

All of the training options are presented in Tab. 4. They were chosen empirically in order to perform precise training in a reasonable time frame.

**Tab. 4.** Available training options.

| Training parameter | Value |
|---|---|
| Mini-batch options | |
| Max epochs | 30 |
| Mini-batch size | 10 |
| Shuffle | Every epoch |
| Validation options | |
| Validation frequency | 16 |
| Validation patience | Infinity |
| Solver options | |
| Solving algorithm | Stochastic gradient descent with momentum (SGDM) |
| Momentum | 0.9 |
| Initial learn rate | 0.0001 |
| Learn rate schedule | None |
| Learn rate drop factor | 0.1 |
| Learn rate drop period | 10 |
| L2 regularization | 0.0001 |

### *4.3. Classification Accuracy*

The effectiveness of CNNs can be measured and compared using the following parameters: classification accuracy (the average accuracy in the case of k-fold cross validation), prediction rates for particular pairs of classes (confusion matrix), training time, and decision time. The most important thing when implementing a CNN-based solution is the classification accuracy, and more specifically, low false prediction rates and high true prediction rates. Table 5 shows the obtained classification accuracy. The prediction rates are stored as confusion matrices in Tabs. 6, 7, and 8. The confusion matrix for a given CNN shows the overall classification efficiency for each class. The correct classification of each class is rep-

**Tab. 5.** Classification accuracy for the proposed networks.

| Network | Accuracy | Decision time | Training time |
|---|---|---|---|
| CNN#1 | 0.8174 | $\sim$ 0.09 s | $\sim$ 5–7 min |
| CNN#2 | 0.9783 | $\sim$ 0.13 s | $\sim$ 10–12 min |
| CNN#3 | 0.9304 | $\sim$ 0.15 s | $\sim$ 15–17 min |

**Tab. 6.** Confusion matrix for CNN#1.

| True class | Predicted class | | | | True positive | False negative |
|---|---|---|---|---|---|---|
| | day 1 | day 2 | day 4 | day 8 | | |
| Day 1 | 0.715 | 0.093 | 0.136 | 0.056 | 0.715 | 0.285 |
| Day 2 | 0.080 | 0.900 | 0.000 | 0.020 | 0.900 | 0.100 |
| Day 4 | 0.046 | 0.000 | 0.874 | 0.080 | 0.874 | 0.126 |
| Day 8 | 0.025 | 0.047 | 0.189 | 0.739 | 0.739 | 0.261 |

**Tab. 7.** Confusion matrix for CNN#2.

| True class | Predicted class | | | | True positive | False negative |
|---|---|---|---|---|---|---|
| | day 1 | day 2 | day 4 | day 8 | | |
| Day 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| Day 2 | 0 | 1 | 0 | 0 | 1 | 0 |
| Day 4 | 0.024 | 0.011 | 0.954 | 0.011 | 0.954 | 0.046 |
| Day 8 | 0 | 0.022 | 0 | 0.978 | 0.978 | 0.022 |

**Tab. 8.** Confusion matrix for CNN#3.

| True class | Predicted class | | | | True positive | False negative |
|---|---|---|---|---|---|---|
| | day 1 | day 2 | day 4 | day 8 | | |
| Day 1 | 0.904 | 0.018 | 0.078 | 0 | 0.904 | 0.096 |
| Day 2 | 0.020 | 0.920 | 0.040 | 0.02 | 0.920 | 0.080 |
| Day 4 | 0.023 | 0 | 0.965 | 0.012 | 0.965 | 0.035 |
| Day 8 | 0.025 | 0.047 | 0.022 | 0.906 | 0.906 | 0.094 |

resented by the diagonal cells, while elements beyond the diagonal indicate a misclassification. The last two columns contain the true positive rate and the false negative rate for each true class, respectively.

Figure 3 presents receiver operating characteristic (ROC) curves, being the plots of the true positive rate against the false positive rate for a given class. The ROC curve for the best classifier is near the upper left corner of the graph. On the other hand, a diagonal line across the whole plot would represent random guessing. The classifier's overall quality is measured by the area under the ROC curve (AUC). The larger the area under the plot, the better the classification performance. It can be seen that the ROCs for CNN#2 are close to the ROCs of the ideal classifier.

It should be noted that the training time and the number of learnable parameters have very little significance in this research. Furthermore, the training time depends strongly on the hardware setup and the dataset used. In the case of decision time, the differences for the different considered CNN architectures are negligible. The application under consideration is not a real-time solution, so all of the results are very satisfactory. The training time and decision time are shown in Tab. 5.

### *4.4. Proposed Network*

The experimental results show that CNN#2 is the best of the presented CNNs for solving the given problem. The overall classification accuracy of this network is equal to 0.9783. The accuracy for the "day 1" class is 1, for the "day 2" class is 1, for the "day 4" class is 0.954, and for the "day 8" class is 0.978. The learning time was about 12 minutes and the decision time was approximately 0.13 s. CNN#2 has a depth of 165, consists of 244 convolutional layers and 2 fully connected layers, and the number of learnable parameters is 60,648,676. The proposed network is shown in Fig. 4.

Deep learning techniques, in particular convolutional neural networks, are characterized by very high predictive power, but are not easily interpretable by humans. This is sometimes pointed out as a weakness of CNN compared to methods based on expert knowledge. Interpreting a non-linear classifier is important to gain trust into the prediction and to identify potential data selection mistakes. In order to solve this challenge, one can use the so-called importance maps or heatmaps.

Figure 5 shows exemplary heatmaps, calculated using the Grad-CAM (gradient-weighted class activation mapping) algorithm, which correspond to the sample data from the dataset presented in Fig. 1. These importance maps visualize the regions of the analyzed image that activated, the most, the outputs from the last convolution layer while calculating classification scores for specific class labels. Areas on a map with a large value are those that impact the network score for that class the most. There exists a direct analogy to magnetic resonance imaging (MRI) of the human brain, in which different regions activate for different stimulations. If the pixels on a heatmap are red, it means that the last convolutional layer has been strongly activated at a given location. The network considers this feature to be important and valid for a given class label. In this way, one can see that the network

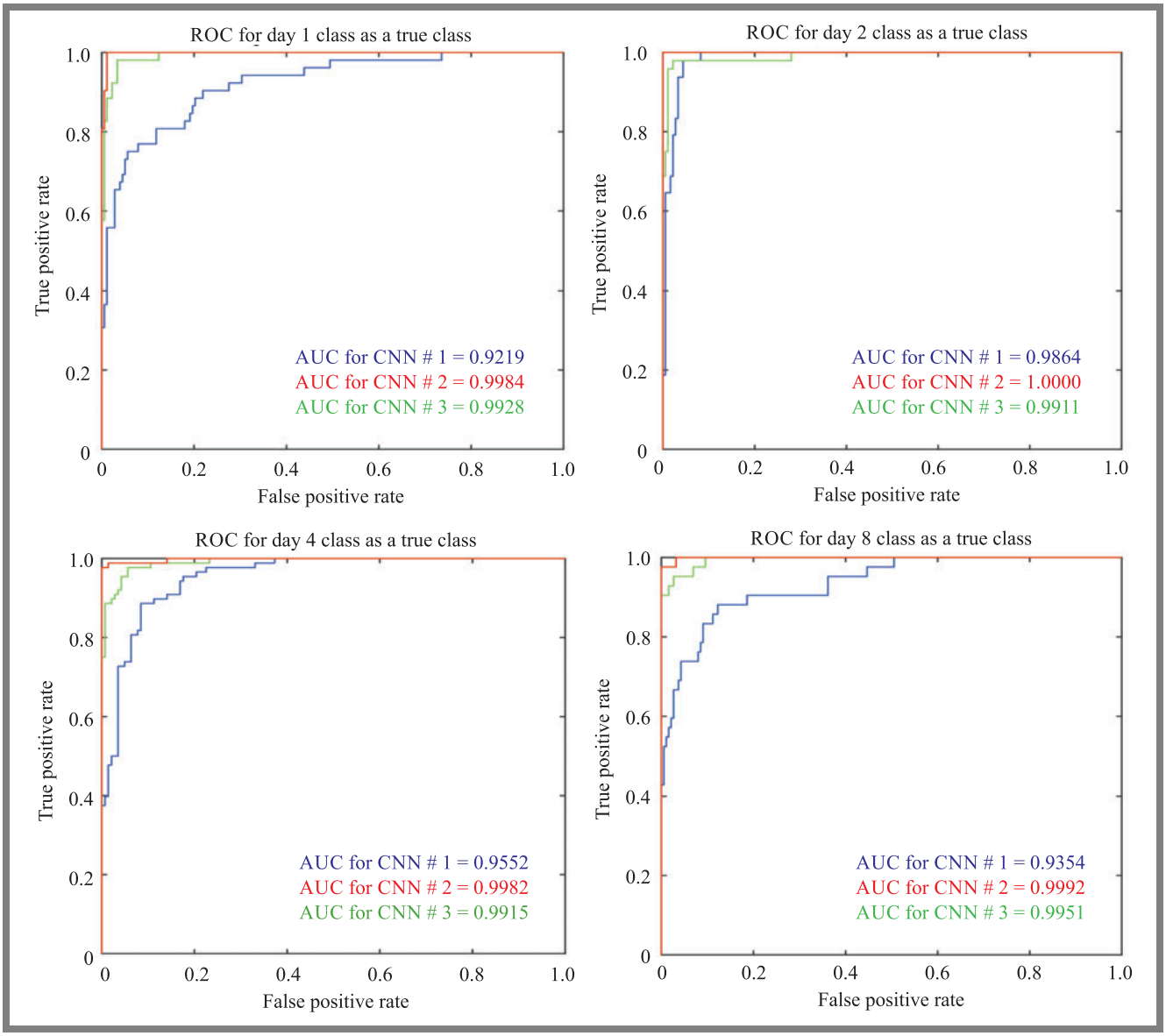



**Fig. 3.** ROC curves for the considered CNNs.

focuses on what is essential for distinguishing the age of *C. elegans*. By analyzing individual heatmaps, we can see with the naked eye what parts of the data are recognized and taken into consideration by the CNN.

As expected, the network extensively utilizes information from the rim of the nematode's body. This is the very region where the progressively deteriorating body-wall muscles are located. However, quite interestingly, the network also accommodates morphological changes in the pharynx, which is evident in Fig. 5, day 8, but also noticeable on day 2. The CNN appears to include even more regions (day 4). Unfortunately, the organs and tissues from such regions (excluding body-wall muscles) are difficult to determine.

### *4.5. Comparison of Methods*

The three CNNs featured in this paper have been compared with other state-of-the-art methods in terms of classification performance. The methods used for the comparison were tested on the same dataset as the proposed method. The results are presented in Tab. 9. It can be seen that the proposed CNN#2 offers nearly the highest accuracy currently available in the literature. However, when comparing the methods, one should take into account that the proposed CNN#2 comes with all of the advantages of a fully self-learned CNN-based solution without the handcrafted feature selection stage requiring expert knowledge.

The only method that has achieved a similar level of accuracy was [27], but their direct comparison is very difficult here, because it represents a completely different approach to solving the problem. The method applied in [27] uses separate feature selection and classification tools. Feature selection is based on a very sophisticated, hand-crafted, high-dimensional descriptor set, which is then assisted by a specialized SDT algorithm. Afterwards, classification is performed via SVM. On the other hand, the methodology proposed in this paper

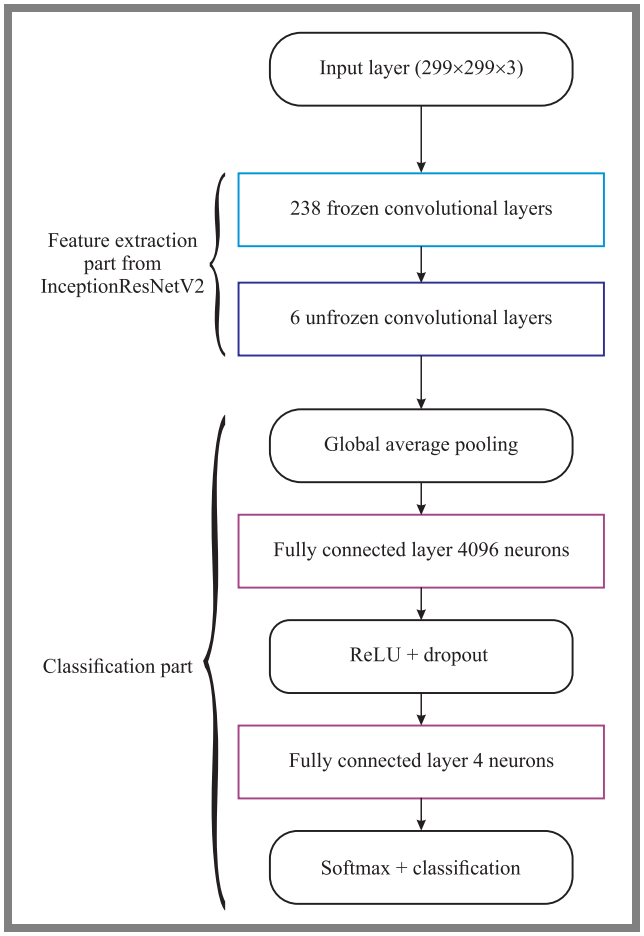



**Fig. 4.** The proposed CNN for *C. elegans* muscle aging classification.

**Tab. 9.** Classification accuracy for different state-of-the-art methods.

| Method | Accuracy |
|---|---|
| Song *et al.* [27] | 0.979 |
| CNN#2 | 0.9783 |
| CNN#3 | 0.9304 |
| Nanni *et al.* [30] | 0.9375 |
| Nanni *et al.* [29] | 0.9292 |
| Zhou *et al.* [25] | 0.896 |
| Lin *et al.* [22] | 0.8478 |
| CNN#1 | 0.8174 |
| Mohan and Sekhar [28] | 0.75 |
| Siji *et al.* [26] | 0.75 |
| Shamir *et al.* [35] | 0.53 |

offers the same level of effectiveness by adopting a much simpler strategy. Firstly, feature extraction and classification are implemented using one coherent tool (CNN). Secondly, the CNN-based solution is more open should a need to add more classes or extend the dataset arise. Thirdly, the method presented in this paper uses date augmentation to protect against overfitting, which is ignored in [27]. Fourthly, the CNN-based solution is free from SVM flaws, such as high algorithmic complexity, extensive memory requirements, sensitivity to

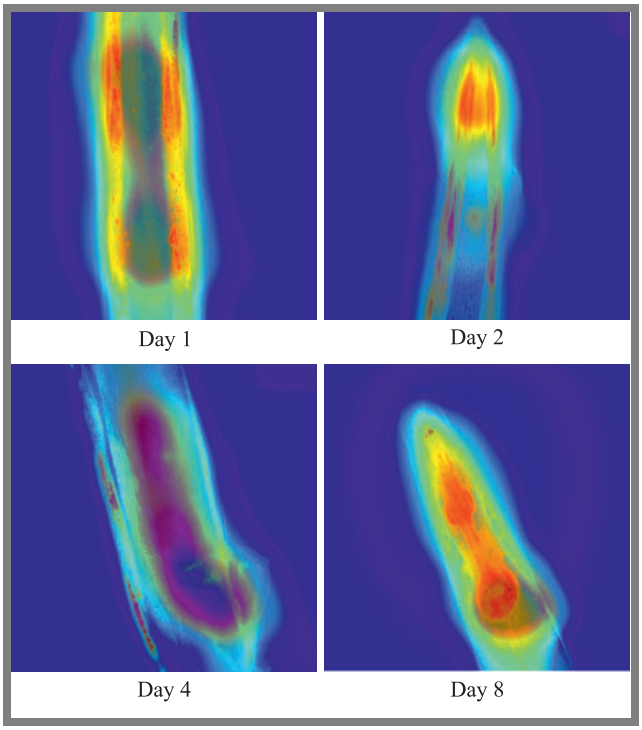



**Fig. 5.** Examples of heatmaps obtained with the use of CNN#2, in which the network accommodates not only body-wall muscles, but also pharynx, and regions that cannot be unambiguously determined.

outliers, lack of prediction scores (classification probabilities), very time-consuming training, unsuitability for large datasets, underperformance when the dataset is noisy, and underperformance if the number of features for each data point exceeds the number of training data samples. Ultimately, the biggest advantage of the proposed method over the method presented in [27] is that there is no need to develop a hand-crafted feature selection stage.

The next two methods from [29] and [30] produce slightly worse results. In both cases, CNNs are only used as a supplement to handcrafted features. Feature selection consists of local features, dense sampling features, and deep learning features which are then classified using SVM. Firstly, unfortunately, this still requires specialized expert knowledge to create a set of handcrafted features. Secondly, the CNNs presented in [29] and [30] have not been used in an optimal way, because the authors have not unfrozen (allowed re-learning) a single convolution layer and are using the deep learning features learned for a completely different, unrelated classification problem, completely relying on adaptation by SVM. As a result, the presented methodology for using CNNs, even when boosted by handcrafted features, does not provide as good results as the CNN#2 network presented in this paper. Thirdly, these methods also deal with the drawbacks of SVM outlined in the previous paragraph. In this context, the SVM acts as a bottleneck that blocks the useful classification properties of the CNN.

It should be noted that the method presented in [25], although less accurate than CNN#2, is intended for analyzing 3D images. In order for the proposed methodology to be able to

operate on 3D images, additional studies involving extrapolation to three dimensions would be necessary.

The last compared work that requires an additional comment is presented in [22]. It is a method that calculates *C. elegans* muscle age by regression. It is very difficult to fairly compare the regression method to the classification method, e.g. the results of the regression method can be interpreted as classification results with a class resolution of one day (i.e. many classes differing by one day). However, in [22], it is necessary to add some metadata or a description to the image, which implies additional work that needs to be supervised by a human. For the method proposed in this paper, only an image is entered as input, with no additional metadata or description.

# 5. Conclusion

This paper presents the results of research on *C. elegans* muscle age classification using CNNs with self-learned features only. The problem under consideration is very difficult for humans to solve. Therefore, machine learning-based solutions are being sought. In anti-aging drug research, the proposed CNN would be a very fast and effective age determination method, which would lead to reductions in time and costs.

The challenge in these studies was the small dataset, which made it difficult to train the CNNs effectively and posed the risk of overfitting. Transfer learning and data augmentation techniques were used to solve this problem. VGG16, InceptionV3, and InceptionResNetV2 networks were selected as the base networks for transfer learning. The impact of unfreezing successive convolutional layers on the classification accuracy was investigated, as well as the impact of extending the classifier model with additional fully connected layers. The overall classification accuracy of the proposed network is equal to 0.9783. The proposed network has a depth of 165, consists of 244 convolutional layers and 2 fully connected layers, and the number of learnable parameters is 60,648,676. The proposed network was compared with other solutions described in the literature.

The achieved results prove that CNNs can be successfully used for the classification of *C. elegans* muscle aging without handcrafted features. Moreover, given the fact that CNNs are free from the disadvantages of analytical feature selection and other machine learning methods, the results achieved are improvements in the research problem under consideration and are of great importance for anti-aging studies.

Our study shows that during the early stages of ageing, *C. elegans* exhibits morphological alterations not only in the muscle tissue but also in other parts of the nematode's body, e.g. its pharynx. The CNNs proposed by us are capable of detecting such changes with a high degree of accuracy, therefore providing reliable information regarding potentially new sources of ageing.

Future research may include the following. Retraining the network on an extended dataset with more classes, e.g. each class could represent a separate day from a 2-week span. Integrating the proposed CNN with region-based active contour models for medical image segmentation [41]. Application of the proposed CNN to solve other classification problems in diagnostics [34], [35].

### *Authors' Contributions*

Conceptualization: Damian Panas,
Data curation: Damian Panas,
Formal analysis: Mariusz Dzwonkowski,
Acquisition of funding: Bartosz Czaplewski and Mariusz Dzwonkowski,
Investigation: Bartosz Czaplewski, Mariusz Dzwonkowski and Damian Panas.
Methodology: Bartosz Czaplewski and Mariusz Dzwonkowski,
Project administration: Bartosz Czaplewski;
Resources: Damian Panas,
Software: Bartosz Czaplewski.
Supervision: Bartosz Czaplewski, Validation, Bartosz Czaplewski; Visualization: Bartosz Czaplewski,
Writing – original draft: Bartosz Czaplewski, Mariusz Dzwonkowski and Damian Panas. Writing – review and editing: Bartosz Czaplewski and Mariusz Dzwonkowski.

### *Funding*

This work was supported by a ministerial research subsidy awarded to Gdańsk University of Technology.

# References

[1] A.G. Alexander, *et al.*, "Use of *Caenorhabditis elegans* as a model to study Alzheimer's disease and other neurodegenerative diseases", *Front. Genet.*, vol. 5, 2014 (DOI: 10.3389/fgene.2014.00279).

[2] E. Moreno-Arriola, *et al.*, "*Caenorhabditis elegans*: A useful model for studying metabolic disorders in which oxidative stress is a contributing factor", *Oxid. Med. Cell. Longev.*, vol. 2014, 2014 (DOI: 10.1155/2014/705253).

[3] A.E. Apostolakou, *et al.*, "Exploring the conservation of Alzheimer-related pathways between H. sapiens and *C. elegans*: a network alignment approach", *Sci. Rep.*, vol. 11, no. 1, 2021 (DOI: 10.1038/s41598-021-83892-9).

[4] K. Podshivalova, *et al.*, "How a mutation that slows aging can also disproportionately extend end-of-life decrepitude", *Cell Rep.*, vol. 19, no. 3, pp. 441–450, 2017 (DOI: 10.1016/j.celrep.2017.03.062).

[5] M.M. Abdelsamea, "A semi-automated system based on level sets and invariant spatial interrelation shape features for *Caenorhabditis elegans* phenotypes", *J. Vis. Commun. Image Represent.*, vol. 41, pp. 314–323, 2016 (DOI: 10.1016/j.jvcir.2016.10.011).

[6] H.A. Tissenbaum, "Using *C. elegans* for aging research", *Invertebr. Reprod. Dev.*, vol. 59, no. sup1, pp. 59–63, 2015 (DOI: 10.1080/07924259.2014.940470).

[7] S. Zhang, F. Li, T. Zhou, G. Wang, and Z. Li, "*Caenorhabditis elegans* as a useful model for studying aging mutations", *Front. Endocrinol. (Lausanne)*, vol. 11, 2020 (DOI: 10.3389/fendo.2020.554994).

[8] J.J. Collins *et al.*, "The measurement and analysis of age-related changes in *Caenorhabditis elegans*", *WormBook: The Online Review of C. elegans Biology. Pasadena, CA, WormBook*, 2005–2018 (URL: https://www.ncbi.nlm.nih.gov/books/NBK116075/).

[9] D.S. Wilkinson, R.C. Taylor, and A. Dillin, "Analysis of aging in *Caenorhabditis elegans*", *Methods Cell Biol.*, vol. 107, pp. 353–381, 2012 (DOI: 10.1016/B978-0-12-394620-1.00012-6).

[10] L.A. Herndon, *et al.*, "Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*", *Nature*, vol. 419, no. 6909, pp. 808–814, 2002 (DOI: 10.1038/nature01135).

[11] H.G. Son, O. Altintas, E.J.E. Kim, S. Kwon, and S-J.V. Lee, "Age-dependent changes and biomarkers of aging in *Caenorhabditis elegans*", *Aging Cell*, vol. 18, no. 2, 2019 (DOI: 10.1111/acel.12853).

[12] G.J. Lithgow, "The Future of Worm Ageing", *Ageing: Lessons from C. elegans. Healthy Ageing and Longevity, A. Olsen, M. Gill, Eds., Switzerland, Cham: Springer International Publishing, A. Olsen, M. Gill (eds) Ageing: Lessons from C. elegans. Healthy Ageing and Longevity. Springer, Cham.*, pp. 431–435 2017 (DOI: 10.1007/978-3-319-44703-2_19).

[13] A. Olsen and M.S. Gill, "Introduction in Ageing: Lessons from *C. elegans*. Healthy Ageing and Longevity, *A. Olsen, M. Gill, Eds., Switzerland, Cham: Springer International Publishing*, pp. 1–7, 2017 (DOI: 10.1007/978-3-319-44703-2_1).

[14] D. Gems and L. Partridge, "Genetics of longevity in model organisms: debates and paradigm shifts", *Annu. Rev. Physiol.*, vol. 75, pp. 621–644, 2013 (DOI: 10.1146/annurev-physiol-030212-183712).

[15] S.S. Yadav and S.M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis", *J. Big Data*, vol. 6, no. 133, 2019 (DOI: 10.1186/s40537-019-0276-2).

[16] B. Czaplewski and M. Dzwonkowski, "A novel approach exploiting properties of convolutional neural networks for vessel movement anomaly detection and classification", *ISA Transactions*, vol. 119, pp. 1–16, 2022 (DOI: 10.1016/j.isatra.2021.02.030).

[17] B. Czaplewski, "An improved convolutional neural network for steganalysis in the scenario of reuse of the stego-key", *I. Tetko, V. Kůrková, P. Karpov, F. Theis (eds) Artificial Neural Networks and Machine Learning – ICANN 2019, Lecture Notes in Computer Science*, vol. 11729, Springer, Cham. 2019 (DOI: 10.1007/978-3-030-30508-6_7).

[18] J. Zhang, K. Shao, and X. Luo, "Small sample image recognition using improved Convolutional Neural Network", *J. Vis. Commun. Image Represent.*, vol. 55, pp. 640–647, 2018 (DOI: 10.1016/j.jvcir.2018.07.011).

[19] M. Claro, *et al.*, "An hybrid feature space from texture information and transfer learning for glaucoma classification", *J. Vis. Commun. Image Represent.*, vol. 64, 102597, 2019 (DOI: 10.1016/j.jvcir.2019.102597).

[20] C. Restif and D. Metaxas, "Tracking the Swimming Motions of *C. elegans* Worms with Applications in Aging Studies", *Med. Image Comput. Comput. Assist. Interv.*, D. Metaxas *et al.* Eds., Germany, Berlin, Heidelberg: Springer, vol. 5241, pp. 35–42, 2008 (DOI: 10.1007/978-3-540-85988-8_5).

[21] J. Johnston, W.B. Iser, D.K. Chow, I.G. Goldberg, and C.A. Wolkow, "Quantitative image analysis reveals distinct structural transitions during aging in *Caenorhabditis elegans* tissues", *PLoS One*, vol. 3, no. 7, 2008 (DOI: 10.1371/journal.pone.0002821).

[22] J-L. Lin, W.-L. Kuo, Y-H. Huang, T-L. Jong, A-L. Hsu, and W-H. Hsu, "Using convolutional neural networks to measure the physiological age of *Caenorhabditis elegans*", *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 18, no. 6, pp. 2724–2732, 2021 (DOI: 10.1109/TCBB.2020.2971992).

[23] L. Shamir, N. Orlov, D.M. Eckley, T. Macura, J. Johnston, and I.G. Goldberg, "Wndchrm – an open source utility for biological image analysis", *Source Code Biol. Med.*, vol. 3, no. 13, 2008 (DOI: 10.1186/1751-0473-3-13).

[24] N. Orlov, L. Shamir, T. Macura, J. Johnston, D.M. Eckley, and I.G. Goldberg, "WND-CHARM: Multi-purpose image classification using compound image transforms", *Pattern Recognit. Lett.*, vol. 29, no. 11, pp. 1684–1693, 2008 (DOI: 10.1016/j.patrec.2008.04.013).

[25] J. Zhou, S. Lamichhane, G. Sterne, B. Ye, and H. Peng, "BIOCAT: a pattern recognition platform for customizable biological image classification and annotation", *BMC Bioinformatics*, vol. 14, no. 291, 2013 (DOI: 10.1186/1471-2105-14-291).

[26] K.K. Siji, B.S. Mathew, R. Chandran, B.S. Shajeemohan, and K.S. Shanthini, "Feature selection, optimization and performance analysis of classifiers for biological images", *2014 IEEE National Conference on Communication, Signal Processing and Networking (NCCSN)*, pp. 1–5, 2014 (DOI: 10.1109/NCCSN.2014.7001150).

[27] Y. Song, W. Cai, H. Huang, D. Feng, Y. Wang, and M. Chen, "Bioimage classification with subcategory discriminant transform of high dimensional visual descriptors", *BMC Bioinformatics*, vol. 17, no. 465, 2016 (DOI: 10.1186/s12859-016-1318-9).

[28] B.S. Shajee Mohan and C.C. Sekhar, "Distance metric learnt kernel based SVMs for semi-supervised pattern classification", *2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR)*, pp. 1–6, 2017 (DOI: 10.1109/ICAPR.2017.8592956).

[29] L. Nanni, S. Ghidoni, and S. Brahnam, "Handcrafted vs. non-handcrafted features for computer vision classification", *Pattern Recognit.*, vol. 71, pp. 158–172, 2017 (DOI: 10.1016/j.patcog.2017.05.025).

[30] L. Nanni, S. Brahnam, S. Ghidoni, and A. Lumini, "Bioimage classification with handcrafted and learned features", *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 16, no. 3, pp. 874–885, 2019 (DOI: 10.1109/TCBB.2018.2821127).

[31] L. Nanni, S. Ghidoni, and S. Brahnam, "Ensemble of convolutional neural networks for bioimage classification", *Appl. Comput. Inform.*, vol. 17, no. 1, 2020 (DOI: 10.1016/j.aci.2018.06.002).

[32] C.K. Ahn, C. Heo, and H. Jin, "A novel deep learning-based approach to high accuracy breast density estimation in digital mammography", *Med. Imaging: Computer-Aided Diagnosis*, vol. 10134, SPIE, 2017 (DOI: 10.1117/12.2254264).

[33] X. Li, Y. Quin, Z. Liu, "Pretraining improves deep learning based tissue microstructure estimation", *Computational Diffusion MRI, Springer, Cham*, 173–185, 2021 (DOI: 10.1007/978-3-030-73018-5_14).

[34] –, *IICBU dataset website*, 2008 (URL: https://ome.irp.nia.nih.gov/iicbu2008/).

[35] L. Shamir, N. Orlov, D.M. Eckley, T. Macura, and I.G. Goldberg, "IICBU 2008: a proposed benchmark suite for biological image analysis", *Med. Biol. Eng. Comput.*, vol. 46, no. 9, pp. 943–947, 2008 (DOI: 10.1007/s11517-008-0380-5).

[36] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", 2014 (DOI: 10.48550/arXiv.1409.1556).

[37] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2016 (DOI: 10.1109/CVPR.2016.308).

[38] C. Szegedy, S. Ioffe, V. Vanhoucke, and A.A. Alemi., "Inception-v4, Inception-ResNet and the impact of residual connections on learning", *Proc. Conf. AAAI Artif. Intell.*, vol. 31, no. 1, pp. 4278–4284, 2017 (DOI: 10.1609/aaai.v31i1.11231).

[39] –, *ImageNet – image database*, 2021 (URL: http://www.image-net.org).

[40] 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 (DOI: 10.1186/s40537-019-0197-0).

[41] A. Pratondo, C-K. Chui, and S-H. Ong, "Integrating machine learning with region-based active contour models in medical image segmentation", *Journal of Visual Communication and Image Representation*, vol. 43, pp. 1–9, 2017 (DOI: 10.1016/j.jvcir.2016.11.019).

---

**Bartosz Czaplewski** received the M.Sc. and Ph.D. degrees from the Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, in 2011 and 2015, respectively. Currently, he works as an assistant professor at Gdańsk University of Technology, Poland. His research interests include cybersecurity, cryptography, digital fingerprinting, image processing, machine learning, and deep learning.

https://orcid.org/0000-0001-7904-5567

E-mail: bartosz.czaplewski@eti.pg.edu.pl

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Department of Teleinformation Networks, Poland

**Mariusz Dzwonkowski** received the M.Sc. and Ph.D. degrees from the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, in 2011 and 2017, respectively. He is an assistant professor with Medical University of Gdańsk, Poland and with Gdańsk University of Technology, Poland. His research interests include cryptography, hypercomplex algebra, information coding and reversible data hiding.

https://orcid.org/0000-0003-3580-7448

E-mail: mard@gumed.edu.pl

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Department of Teleinformation Networks, Poland

Medical University of Gdańsk, Faculty of Health Sciences, Department of Radiology Informatics and Statistics, Poland

**Damian Panas** is a bioinformatician at the Molecular Biology Laboratory at the Institute of Animal Reproduction and Food Research, Polish Academy of Sciences. He received his M.Sc. degree from the Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, Poland, and his Ph.D degree from the Faculty of Health Sciences, Medical University of Gdańsk, Poland. His areas of interest include mathematical modeling, statistics, machine learning, computational biology, and multi-omic data analysis.

https://orcid.org/0000-0003-2181-2414

E-mail: d.panas@pan.olsztyn.pl

Institute of Animal Reproduction and Food Research of the Polish Academy of Sciences, Molecular Biology Laboratory, Poland

# *Information for Authors*

***Journal of Telecommunications and Information Technology*** **(JTIT)** is published quarterly. 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.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

**Manuscript.** TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The authors JTIT LATEX style file is available:
https://www.il-pib.pl/pl/submission

Papers published should contain up to 10 printed pages in LATEX authors style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150–200 words long and the relevant keywords. 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 should not repeat the title of the manuscript. About four keywords or phrases in alphabetical order should be used, separated by commas.

The original files accompanied with pdf file should be submitted by e-mail: redakcja@il-pib.pl

editorial.office@il-pib.pl

**Figures, tables and photographs.** Original figures should be submitted. Drawings in Corel Draw and PostScript formats are preferred. Figure captions should be placed below the figures and can not be included as a part of the figure. Each figure should be submitted as a separated graphic file, in .cdr, .eps, .ps, .png or .tif format.

Tables and figures should be numbered consecutively with Arabic numerals.
Each photograph with minimum 300 dpi resolution should be delivered in electronic formats (TIFF, JPG or PNG) as a separated file.

**References.** All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:


[1] Y. Namihira, Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres, *Electron. Lett*., vol. 30, no. 3, pp. 262–264, 1994.

[2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

[3] S. Demri and E. Orłowska, Informational representability: Abstract models versus concrete models, in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314


**Biographies and photographs of authors.** A brief professional authors biography of up to 200 words and a photo of each author should be included with the manuscript.

**Galley proofs.** Authors should return proofs as a list of corrections as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within one week after receiving the offprint.

**Copyright.** Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the author(s) 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.

For copyright form see:
https://www.il-pib.pl/pl/submission

---

(*Contents Continued from Front Cover*)



National Institute of Telecommunications

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

tel. +48 22 512 81 83
fax: +48 22 512 84 00
e-mail: redakcja@il-pib.pl
editorial.office@il-pib.pl
http://www.nit.eu

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY 4/2022