JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2022

(Contents Continued on Back Cover)

Editor-in-Chief

Steering Editor

Editorial Advisory Board

Publications Staff

Content Editor:	**Robert Magdziak**
Managing Editor:	**Ewa Kapuściarek**
Technical Editor:	**Grażyna Woźnica**

on-line: ISSN 1899-8852

Paper

Modified PSO Based Channel Allocation Scheme for Interference Management in 5G Wireless Mesh Networks

Nirmalkumar S. Benni and Sunilkumar S. Manvi

Department of Electronics and Communication Engineering, REVA Institute of Technology and Management, Bangalore, Karnataka, India

https://doi.org/10.26636/jtit.2022.156621

Abstract—Efficient channel management is a challenge that next-generation wireless networks need to meet in order to satisfy increasing bandwidth demand and transmission rate requirements. Non-orthogonal multiple access (NOMA) is one of such efficient channel allocation methods used in 5G backhaul wireless mesh networks. In this paper, we propose a power demand-based channel allocation method for 5G backhaul wireless mesh networks by employing NOMA and considering traffic demands in small cells, thereby improving channel utility. In this scheme, we work with physical layer transmission. The foremost aim is to mutually optimize the uplink/downlink NOMA channel assignment in order to increase user fairness. The approach concerned may be divided into two steps. First, initial channel allocation is performed by employing the traveling salesman problem (TSP), due to its similarity to many-to-many double-side user-channel allocation. Second, the modified particle swarm optimization (PSO) method is applied for allocation updates, by introducing a decreasing coefficient which may have the form of a standard stochastic estimate algorithm. To enhance exploration capacity of modified the PSO, a random velocity is included to optimize the convergence rate and exploration behavior. The performance of the designed scheme is estimated through simulation, taking into account such parameters as throughput, spectral efficiency, sum-rate, outage probability, signal-to-interference plus noise ratio (SINR), and fairness. The proposed scheme maximizes network capacity and improves fairness between the individual stations. Experimental results show that the proposed technique performs better than existing solutions.

Keywords—channel allocation, co-channel interference, convex optimization, multicarrier NOMA, Rayleigh fading.

1. Introduction

5G is a trending communication technology that is capable of improving network performance in urban areas [1]. 5G enhances the collection of information and the framework measurement rate. The data transmission rate, thickness association, and inactivity of ultra-low signals are the other benefits of 5G development enjoyed in multiple input multiple output (MIMO) systems.

Association of the web with remote cell towers is known as the concept of backhaul. In multi-level media communication, the backhaul region consists of a system characterized by a specific orientation of its connections, e.g. spine organization, center system, and the edge of the progressive area [2]. Backhaul improves the speed at which information is exchanged. Genuinely, without backhaul, users would not be able to enjoy a web relationship in any way, shape or form. Therefore, the backhaul effect should be considered to provide a high priority information background.

Prerequisites of this type affect backhaul in a peculiar manner, as it may bear information with highly complex and flexible green contemplation [3]. These effects are very difficult to acknowledge and secure 5G communication-based protocols in the networking environment. Media communications organization handles expansive volume spilling information and uphold experiences from peculiarity recognition and prescient displaying to comprehend their systems and their clients [4], [5]. In this way, the presumption of security segment of existing cell frameworks, which rely upon making sure about the significant attainable quality and guard of end-users, the 5G cell structure is depended upon to ensure that a redesigned security instrument is set up in general framework to address issues of approval and support for various interconnected IoT devices.

Many methods are available for enhancing the uplink stream. Versatile quality of service (QoS)-related environments may affect the system of 5G communication-based portable hubs in terms of their collection ability [6], [7]. NOMA is a prominent access system for executing upgrades in cutting-edge cell interchanges. Contrasted with symmetrical recurrence division different access, which is a notable high-limit orthogonal multiple access scheme, NOMA offers a range of tempting advantages, including higher pro-ductivity. A variety of NOMA strategies exist, including power-area and code-domain [8].

The diverse and demanding characteristics of 5G require a move from the rigid systems of the past, towards more flexible and versatile networks. The use of recently developed radio network technologies results in an improvement of dimensional flexibility in 5G networks [9]. New hand-off solutions are likewise encouraging in the Internet of Things (IoT) applications [10]. As we gain ground about the 5G network of remote systems, the bit-per-joule energy efficiency turns into an imperative structure paradigm for practical advancement [11]. With that considered, MIMO-related innovations turn out to be one of the major empowering agents for 5G solutions in which BSs are furnished with adequate reception hardware to satisfy requirements related to phantom and energy efficiency increases over current LTE systems [12].

1.1. Problem Statement

5G wireless networks are expected to support very diverse applications and terminals. Massive connectivity with a large number of devices is an important requirement. In the 5G era, the evolution of heterogeneous networks (Het-Nets) results in densification of different sizes of cells. Due to the time- and space-dependent service requirements and traffic patterns, time-varying asymmetric traffic loads are expected in both uplink (UL) and downlink (DL) connections in different cells. Many optimization strategies have been designed to provide seamless coverage and QoS in DL and UL. However, the intractable nature of the channel selection problem motivates us to design an efficient channel allocation scheme through joint optimization of UL and DL streams. Performance of the designed scheme is estimated through Matlab, with such performance parameters as throughput, spectral efficiency, sum-rate, outage probability, signal-to-interference plus noise ratio (SINR) and fairness taken into consideration and compared with other recent optimization techniques.

1.2. Research Contributions

In this paper, we examine a 5G wireless mesh network that comprises multiple primary networks and subscribed users (SUs). At any instant, a different number of channels with different capacities is allocated by the primary networks to each SU.

The channel allocation problem is formulated as TSP, which is then associated with the many-to-many two-sided user-channel allocation. We acquaint a diminishing coefficient with the updating principle. Thus, the population-based artificial intelligence (AI) concept is used in our work, i.e. a modified PSO may be perceived as a standard stochastic estimate algorithm.

The modified PSO is then used to jointly optimize both uplink and downlink channels using NOMA for optimal channel allocation with proper interference management. Finally, we mutually reform the UL/DL channel allocation using NOMA to widen user fairness with proper interference management.

The related works are discussed in Section 2. Section 3 describes the system model. Section 4 elaborates on the proposed method. Section 5 evaluates performance of the channel allocation process. Section 6 presents the conclusions.

2. Related Works

Xia *et al.* [13] proposed a new kind of virtual channel optimization technique in NOMA for successful power balancing. By using the process presented, data may be separated through the power balancing effect. The minimum Euclidean distance for constellation points without estimating the channel is considered as the best method for the channel state estimation process. A closed-form of the optimization process is developed by maximizing the fixed optimal solution with 2 to 3 users. Also, the less complex effect of the maximum likelihood detector reduces computational intricacy without influencing the implementation of quadrature phase shift keying (QPSK).

Paper [14] proposed a calculation for the UL of huge MIMO frameworks to isolate the joined gotten flag of all clients at the BS into autonomous signs for every client class. While applying the proposed calculation, the computational expense of the flag handling process is diminished and it is conceivable to ensure adaptability on the location methods at the BS. A flag is shown for heterogeneous systems with various classes of clients. Discretionary design cell acquisition subframes (CAS) and distributed antenna systems (DAS) are presented in this paper as well. Total rate examination and computational multifaceted nature contemplated for the proposed decoupled signal detection (DSD) method are exhibited.

In [15], user transmission rate and interference are mainly considered for developing a fractional transmission power allocation (FTPA)-based channel allocation process. Greedy algorithm (GA) is applied to obtain the ideal solution for allocation purposes. Mathematical development of max-min energy effectiveness is developed in the intractable programming solution. A sequential programming approach is determined to obtain an optimal solution for power analysis. Focus is placed on energy-efficient power-based channel allocation to provide an enhanced version of optimal channel modulation.

These upgrades demonstrate that a quality-of-service-aware game theory-based power control (QoS-GTPC) plan can be obtained for 5G versatile frameworks.

Sarigiannidis *et al.* [16] presented a spatially unique power control answer for relieving cell-to-device-to-device (D2D) and D2D-to-cell impedance. The proposed D2D control arrangement is somewhat adaptable, including the exceptional instances of no D2D connections or using the greatest transmit control. Under the considered power control, a diagnostic methodology is produced to assess the pro-ductivity and proficiency of such systems. Investigation of the power control arrangement can productively alleviate obstruction between the cell and the D2D level.

In general, 5G wireless networks provide various facilities to assist assorted applications and terminals which require superb connectivity to connect an enormous number of devices. Han *et al.* [17] developed a security-related protocol for NOMA-based massive MIMO uplink communication. The power allocation model is given through the joint power and sub-channel allocation for secrecy capacity (JPSASC) method to get a sub-optimal solution to the joint issue. Especially, the power stint is developed as a non-pliable game with the perspective of a distribution system. The simulation outcome of JPSASC is developed to explain the secrecy capacity in the NOMA model.

5G networks can be configured as heterogeneous networks (HetNets), in which cell densification is the main feature, with cells of different sizes forming the network. Furthermore, 5G is relied upon to help alleviate time-shifting awry traffic load for both UL and DL connections in various cells. Several optimization strategies have been developed to support seamless coverage and QoS for both DL and UL. However, the intractable nature of the channel selection problem in 5G heterogeneous networks [18] motivates us to design an efficient channel allocation scheme through joint optimization of UL and DL streams.

3. System Model

Dense HetNets have been created based on the principle of diminishing the cell size and increasing the quantity of small cells (SCs) per unit of territory, as such a solution is capable of handling the traffic rates expected in 5G. Here, we consider a 5G heterogeneous network comprising N networks of any type, e.g. wired and wireless links. The network consists of a specified number of primary users (PUs) and subscribed users (SUs). Each network is allocated with the most extreme number of channels, where the channels allocated to SUs rely on the conduct of PUs [19]. 5G supports a heterogeneous network that consists of discrete elements, such as users, services, radio access networks (RAN), and backhaul networks. The backhaul network plays a major role in transferring data intended for the users from/to different base stations within the cellular network. The scheduling of backhaul transmission determines that performance be optimized relative to traffic demands placed on the small cells served. Normally, the traffic demands may change over time during longer transmissions, due to channel dynamics.

We will probably meet long-duration traffic requests over blocks of N diminishing slots. Next, the scheduling problem of more than one such block will be considered. On account of remote backhaul, the drawn-out requests will regularly be out of the ergodic rate locale of the backhaul remote channel. This persuades us to characterize the backhaul scheduling problem in order to decide on and work at the rate point in the backhaul ergodic rate area that is nearest, in some sense, to the traffic requests in the access network. An example of a 5G backhaul network and its components is depicted in Fig. 1.

With MIMO and millimeter-wave communication technologies, the small cell scheme is an inevitable solution for upcoming 5G networks. MIMO has arisen as an innovation catalyst for cutting edge mobile communications in 5G. Furthermore, the increase in channel allocation guaranteed by MIMO is forecast to overcome the capacity crunch experienced in current mobile networks and to allow for the

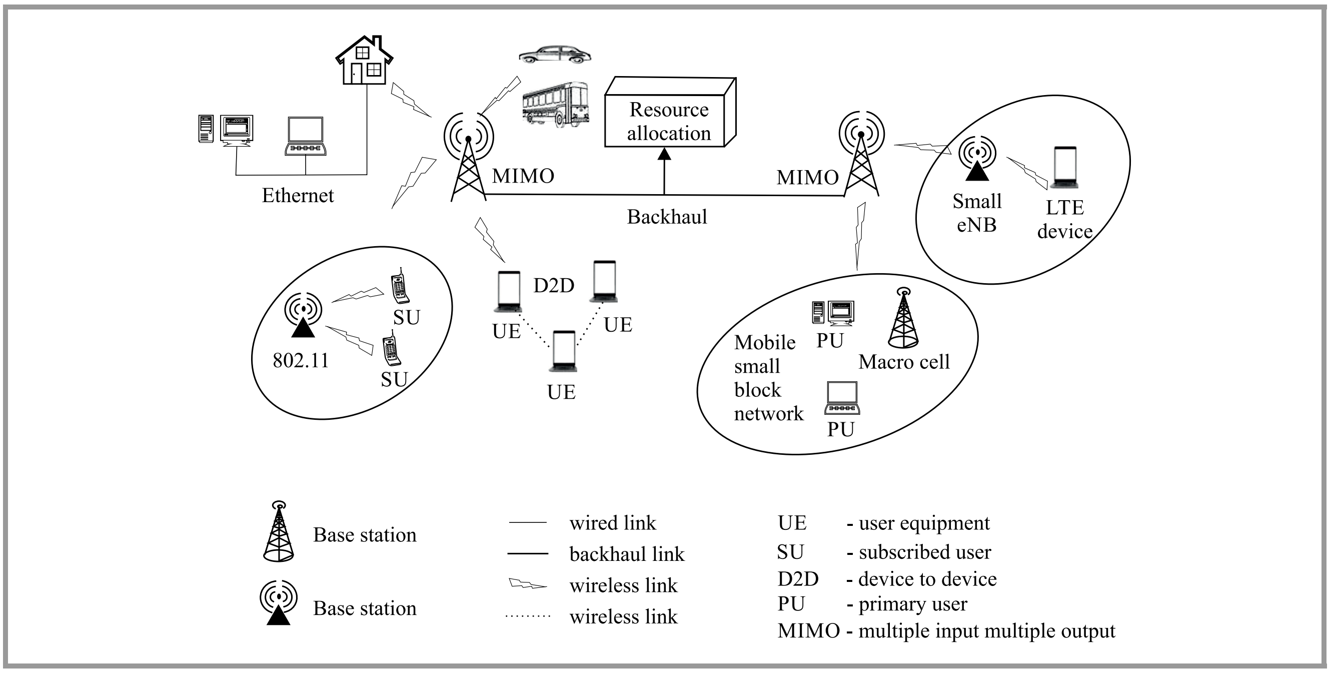


Fig. 1. A 5G wireless mesh backhaul network.

aggressive focus of 5G. The test to acknowledge MIMO for 5G is an effective and cost-proficient reconciliation in the whole network concept. This work features categorization and usage schemes for MIMO with a small cell 5G indoor framework taken into consideration.

In this regard, the MIMO innovation, where the BSs are furnished with countless antennas to satisfy different spectral and energy efficiency gain requirements, will be a key innovation empowering agent for 5G [20], [21].

4. Proposed Method

We initially portray the basics of UL and DL NOMA transmissions and underline their vital differences in terms of usage unpredictability, recognition, and unraveling at the successive interference cancelation (SIC) receiver(s), brought about intra-cell and between cell impedance. At that point, for joint DL and UL NOMA, we hypothetically infer the NOMA predominant condition for every individual client in a two-client NOMA group. The NOMA predominant condition refers to a condition under which the spectral efficiency gains of NOMA are ensured, in contrast with orthogonal frequency division multiple access (OFDMA) [22]. By and large, NOMA permits the superposition of definite message signals of clients within a NOMA group. The ideal message signal is then recognized and decoded at the receiver (client in the DL and BS in the UL) by applying SIC.

Table 1 shows the notations used in the following considerations.

Table 1
Notations used

Symbol	Description
x_n	Unit power message signal for user n
p_n^d	Power allocated for user n
N	Total number of users
Pt	Total power at base station
w_n	Gaussian noise at the receiver for user n
h_n	Channel gain between the BS and user n
v_n	Additive noise
v	Indicates the additive noise at the BS
I_n	Interfering signal
θ	Power splitting factor
w	Receiver noise
R_n^u	Uplink rate achieved by the n-th user
N_0	Noise power
P_I	Power of the interference received by the base station
γ_n^u	Uplink signal to noise ratio
γ_n^d	Downlink signal to noise ratio
P_m^i	Transmit power of downlink
q_r^i	Transmit power of uplink
β_m	Joint effect of path loss
β_n	Shadowing between DL users and UL BSs
x	Current number of iterations
α, β	Positive constant parameters
a	Acceleration factor
w	Inertia weight
ϕ	Path loss exponent
μ	Non-negative constant
H, G	Constant parameters

4.1. Downlink NOMA

In the DL NOMA, the BS communicates the superimposed sign:

$$x = \sum_{n=1}^{N} \sqrt{P_n^d x_n^d} \ ,$$

where x_n^d is the unit power of the message signal proposed for client n, P_n^d means the power allocated for client n, and N indicates the complete number of clients signified by U_N in a NOMA framework.

The power allocated to a client relies on the forces of different clients, because of the BS all out force requirement, $P_t = \sum_{n=1}^{N} P_n^d$, where P_t is the absolute BS power. The signal received by the n-th user is termed $y_n = h_n x + w_n$, where h_n indicates the channel gain between the user n and the BS and w_n represents the Gaussian noise at the receiver for user n [23], [24].

Downlink NOMA utilizes a power allocation mechanism, where high power transmission is utilized for clients with below-par channel conditions and vice versa. Accordingly, at a given client in the NOMA group, the strong interfering signals are caused primarily by the powerful message signals of generally frail channel clients. In that capacity, in order to separate the ideal signal, every client cancels the strong interference by SIC translating, demodulating, and by deducting them from the received signal y. Consequently, the highest channel gain client drops all intra-cluster interferences, while the least channel picks up a client gets the interferences from all clients within its group [25], [26]. Additionally, the transmitting power is subject to [27]:

$$\sum_{n=1}^{N} P_n^d \leqslant P \ . \tag{1}$$

We expect that the signal received by every client U_N is parted into two streams, and the power fraction with condition $0 \leqslant \theta_n \leqslant 1$ is utilized for data processing. Transmitters are the source of interference I. There is a simultaneous transmission of data signals and power from BS. The perception at the n-th client, which is utilized for data interpreting, is given by:

$$y_n = h_n \sqrt{\theta_n} \sum_{i=1}^{N} \sqrt{P_i^d s_i^d} + \sqrt{\theta_n I_n} + w_n \ , \tag{2}$$

where w_n indicates additive noise at U_n and I_n is the interfering signal. Noise is included in two parts of the recipient: the receiver antenna noise and the circuit noise. In any case, the antenna noise is ignored. Hence, we incorporate only one additive noise. Each client U_j completes SIC by identifying and eliminating the $U_N^{\prime}s$ message, for all $n < j$, from its perception. Consequently, the attainable rate at U_N, $n \in [1, 2, \ldots, N]$ is limited by:

$$R_n^d = \min(R^d_{n\to n}, R_{n\to n+1}, \ldots, R^d_{n\to N}) \ . \tag{3}$$

$R^d_{n\to j}$ indicates the amount at which user U_j finds the intended message for user U_N:

$$R^d_{n\to j} = \tau \log_2\left(1 + \frac{P_n^d \theta_j g_j}{\theta_j g_j \sum_{i=n+1}^{N} P_i^d + \theta_j P_I, j+1}\right) \ , \tag{4}$$

where $P_n^d = \frac{P_n^d}{N_0}$ and $P_{I,j} = \frac{P_{I,j}}{N_0}$ are the interference power received by U_j. We consider that $P_{I,j}$ is detected precisely by U_j and portrayed to the BS to guarantee and allocate the current resources. When n = N, Eq. (4) is:

$$R^d_{n\to j} = \log_2\left(1 + \frac{P_n^d \theta_N g_n}{\theta_N P_{I,N} + 1}\right) \ . \tag{5}$$

$P = P_1^d, \ldots, P_N^d$ shows the set of the power transmission values between the clients and $\theta = \theta_1, \ldots, \theta_N$ is the set of isolating power factors between the clients.

4.2. Uplink NOMA

In uplink NOMA, each client transmits its individual signal x_n^u with:

$$x = \sum_{n=1}^{N} \sqrt{P_n^u} x_n^u$$

and transmit power P_n^u in such a way that the BS received signal can be characterized as:

$$y = \sum_{n=1}^{M} \sqrt{P_n^u} h_n x_n^u + w \ ,$$

where w is the receiver noise (with a density of power spectral N_0) at the BS. The power transmitted per client is restricted by the client's maximum battery power [23].

Note that, for applying SIC and decrypting signals at the BS, it is essential to keep up the uniqueness of different message signals which are superimposed within y. Considering that the channels of various clients are diverse in the uplink, individual message signals encounter definite channel gains. Subsequently, the received signal power, compared with the most potent channel client, is likely the most potent at the BS. Accordingly, this signal is decrypted, first at the BS, and encounters interference from all clients in the group with generally more vulnerable channels. Hence, the communication of the most potent channel gain client encounters interference from all clients within its group, though the communication of the least channel gain client receives zero interference from the clients in its group. In this way, the perception at the BS is given by [27]:

$$y = \sum_{n=1}^{N} \bar{h}_n \sqrt{P_n^u s_n^u} + I + v \tag{6}$$

where I indicates the interfering signal and v is the additive noise at the BS. By utilizing SIC, the capacity region is limited by:

$$\sum_{n\in M_k} R_n^u \leqslant (1-\tau)\log_2\left(1 + \frac{\sum_{n\in M_k} P_n^u g_n}{P_I + 1}\right) \forall M_k : M_k \subseteq N \ , \tag{7}$$

with R_n^u being a quantum of UL accomplished by the n-th client, $P_n^u = \frac{P_n^u}{N_0}$, $P_1 = \frac{P_1}{N_0}$, N_0 is the noise power, and P_I is the interference power acquired by the BS. P_I is detected precisely by the BS. At last, M_k signifies any conceivable subset of the clients. τ is introduced to denote the effect of cross-correlation. The asymptotic Shannon capacities on the UL (C_{UL}) and the DL (C_{DL}) for MU-MIMO channels under convenient transmission are given by [21]:

$$C_{UL} = \sum_{n=1}^{N} \log_2(1 + \gamma_n^u M \beta \psi_n) \ , \tag{8}$$

$$C_{DL} = \max_{a_n \geqslant 0, \sum a_n \leqslant 1} \sum_{n=1}^{N} \log_2(1 + \gamma_n^d M a_n \psi_n) \ , \tag{9}$$

where γ_n^u and γ_n^d are the overall UL and DL SNR's, $[\psi_n], n = 1, 2, \ldots, N$ constitutes the coefficients of large-scale fading for the N UE's, and a_n is a group of variables which should be enhanced to get C_{DL}. At the point when suitable power control systems are utilized to standardize the impact of β_n, the UL capacity improves to $N\log_2(1 + M\gamma_n^{uSNR})$.

Corresponding considerations are given to DL NOMA. Hence, we accomplish multiplexing gains and cluster gains, under suitable transmission conditions, utilizing simple linear processing techniques at the BS, such as, for example, maximal ratio combining (MRC) and zero forcing (ZF) detection. This streamlines the computational burden and the hardware requirements related to the BSs, the BS's actualize complex signal processing techniques, for example, maximum likelihood (ML) recognition and successive interference cancelation (SIC), to accomplish optimal capacities.

The sum power utilization P, accumulated across UL and DL transmissions in a NOMA-MIMO framework, can be displayed as [21]:

$$P = P_{PA} + P_C + P_{sys} \ , \tag{10}$$

where P_{PA} shows the complete DL and UL used through the power amplifiers (PA's) at the UEs and the BS, P_C constitutes the absolute DL and UL circuit power utilized by different digital and analog signal processing circuits at the BS and the UEs and P_{sys} refers to the excess framework-dependent component in P.

While P_{PA} empowers for the sum power use on RF transmissions, P_C incorporates the sum power utilization from

RF chain components, for example, synthesizers and filters, additionally the tasks of the baseband, for example, digital up/down conversion, FFT/IFFT, recipient/precoding combining, channel deciphering/coding, and assessment of the channel. Here, P_C can not be planned according to the regular exercise as a stable term autonomous of (M,K) as per the requirement of hardware and the number of circuit processes in the framework created with K and M. P_{sys} will assume a critical role in describing energy efficiency of 5G networks, since many BS and UE types will co-exist in a multi-tier architecture with various cell sizes, power utilization levels and access technologies.

4.3. Full-Duplex-NOMA System

We consider a full duplex multicarrier NOMA (FD MC-NOMA) framework which includes full duplex base stations (FD-BSs), K DL clients, and J UL clients. All DL and UL clients are provided with two antennas. The FD-BSs are additionally furnished with two antennas for facilitating synchronized DL transmission and UL reception in a similar frequency band. We expect that the BSs and the DL clients are furnished along with successive interference cancelers. The whole frequency W band is apportioned into N_F subcarriers. In this article, the individual subcarriers are distributed between two DL clients and two UL clients at the most, to restrict multi-user interference (MUI) and the UL-to-DL co-channel interference (CCI) on an individual subcarrier and to guarantee low hardware complexity and low processing delays [28], [29].

Presuming that UL clients $r \in (1,\ldots,J)$, UL clients $t \in (1.\ldots,J)$, DL clients $m \in (1,\ldots,K)$ and DL clients $n \in (1,\ldots,K)$ are preferred and multiplexed on subcarrier $i \in (1,\ldots,N_B)$ likewise the required signals at DL client n, DL client m, and the BS are indicated likewise by:

$$Y^i_{DL_m} = \sqrt{P^i_m \beta_m} h^i_m x^i_{DL_m} + I^i_{MU_m} + I^i_{CC_m} + w^i_{DL_m} \ , \tag{11}$$

$$Y^i_{DL_n} = \sqrt{q^i_n \beta_n} h^i_n x^i_{DL_n} + I^i_{MU_n} + I^i_{CC_n} + w^i_{DL_n} \ , \tag{12}$$

$$Y^i_{BS} = \sqrt{q^i_r \bar{\omega}_r} g^i_r x^i_{UL_r} + \sqrt{q^i_t \bar{\omega}_t} g^i_t x^i_{UL_t} + I^i_{SI} + w^i_{BS} \ , \tag{13}$$

where $x^i_{UL_r}$ and $x^i_{DL_m}$ represent the transmission of signals from UL client r to the FD-BS and from the FD-BS to DL client m on subcarrier i. P^i_m and q^i_r are transmit powers of DL client m and UL BS r, separately. β_m and β_n are the joint impact of path loss and shadowing among DL clients and UL BS. h^i_r and h^i_m indicate the small scale fading coefficients for the link between UL BS r and the FD-BS and the link between the FD-BS and DL client m. $I^i_{MU_m}$ and $I^i_{CC_m}$ are MUI and CCI. The joint impact of path loss and shadowing between UL BS r and the FD BS and between DL client m and UL BS r is depicted by $\bar{\omega}_r$ and g_r. Finally, the complex additive white Gaussian noise (AWGN) on subcarrier i is depicted by $w^i_{DL_m}$, $w^i_{UL_r}$, and w^i_{BS}.

An instant subcarrier is assigned to the clients of two DL and UL in the FD MC-NOMA framework simultaneously. Commonly, the UL power of client signals is lower than that of signals released by the BS for DL clients, which makes it complex for the clients of DL to extract and remove the UL signal by accomplishing SIC. Because of their various constraints on the complexity of receiver hardware and QoS needs, different coding schemes and modulations are used in the DL and UL. As a result, DL clients cannot decode and demodulate the UL signals. So, every user signal is treated as noise, and to eliminate other DL clients, only the DL client achieves SIC. For example, we initially consider an individual policy for the SIC decoding order 4 and an allocation of subcarrier. UL BSs r, t and DL clients m, n are multiplexed on subcarrier i. In addition to decoding SIC and eliminating the DL client, *ms* signal is achieved by the DL client n. Before decrypting the UL client *ts* signal, the FD BS first decrypts the UL client *rs* signal and then eliminates it by SIC. Equation (14) is applied to represent the weighted sum throughput of subcarrier i in such an approach:

$$\begin{aligned} U^i_{m,n,r,t} = s^i_{m,n,r,t} \Bigg[& w_m \log_2 \left(1 + \frac{H^i_m P^i_m}{\alpha^i_m + 1}\right) \\ & + w_n \log 2 \left(1 + \frac{H^i_n P^i_n}{\alpha^i_n + 1}\right) + \mu_r \log_2 \left(1 + \frac{G^i_r q^i_r}{\phi I_s I^i \alpha^i_r + 1}\right) \\ & + \mu_t \log_2 \left(1 + \frac{G^i_t q^i_t}{\phi I_s I^i \alpha^i_t + 1}\right) \Bigg] \ , \end{aligned} \tag{14}$$

for the links between the DL, clients m and n and FD BSs r and t on subcarrier i are defined by the total small scale fading coefficients, such as α^i_m, α^i_n, α^i_r and α^i_t respectively. The subcarrier allocation indicator is represented by $s^i_{m,n,r,t} \in (0,1)$. If UL BSs t and r and DL clients n and m are multiplexed on subcarrier i, then $s^i_{m,n,r,t} = 1$. Before decrypting the BS *ts* signal, the FD BS first decrypts UL BS *rs* signal and eliminates it. Likewise, DL client n executes SIC of the DL client m signal. Another resource allocation policy is utilized when $s^i_{m,n,r,t} = 0$. To achieve a particular notation of fairness in resource allocation, the non-negative constants that are identified in the media access control (MAC) layer and $0 \leq w_m \leq 1$ and $0 \leq \mu_r \leq 1$ Eq. (14) mentions the preferences of DL client m and UL BS r respectively.

In practice, self-interference (SI) cannot be canceled completely, regardless of whether the SI channel is known at the FD-BS, because of the limited dynamic range of the receiver. Subsequently, we mold the surplus SI following elimination at the receiving antenna with autonomous zero-mean Gaussian distortion noise, for which change is relative to the received power of the antenna. NOMA frameworks utilize the power domain for multiple access, wherein various clients are provided with various power levels. Specifically, for a particular subcarrier, let us presume that the DL client n intends to decrypt and eliminate the CCI induced by DL client m employing SIC. Interference cancelation is fruitful if SINR received by client *ns* for client m signal is bigger than or equivalent to the SINR received received by client m for its signal. For instance, DL client n can only successfully decrypt and eliminate DL client m signal by SIC on subcarrier i when the accompanying disparity holds:

$$w_n \log_2\left(1+\frac{H_n^i P_n^i}{\alpha_n^i+1}\right) \geqslant w_m \log_2\left(1+\frac{H_m^i P_m^i}{\alpha_m^i+1}\right) , \quad (15)$$

on subcarrier i, the suddenly weighted sum throughput Eq. (14) for the instance of $r=t$ and $m=n$, turn out to be:

$$U_{m,n,r,t}^i = s_{m,n,r,t}^i \Big[w_m \log_2(1+\frac{H_m^i(P_m^i+P_n^i)}{\alpha_m^i+1}) + \mu_r \log_2\left(1+\frac{G_r^t(q_r^t+q_t^i)}{\phi I_s I^i(\alpha_r^i+1)}\right)\Big] . \quad (16)$$

Variable ϕ denotes path loss exponent, μ is the non-negative constant, the constant parameters H and G are defined as $H_m^i = \frac{\varpi_m |h_m^i|^2}{\sigma_{zDL_m}^2}$, $G_r^i = \frac{e_r |g_r^i|^2}{\sigma_{zBS}^2}$ in Eqs. (14)–(16).

The joint UL/DL NOMA channel allocation issue is that, allocating a path for a node either a BS or end client among various nodes in a wireless mesh network, to reduce the processing time and to expand the framework throughput. In our definition, the correspondence between nodes in the backhaul network in a 5G framework can be preoccupied as TSP.

Here, we consider a well-known TSP, in which we need to determine the shortest closed path between clients of both J UL and K DL, with at least one subcarrier allocated to each client. Suppose, $i=(1,2,\ldots,N)$ is the set of TSP clients and the weighted sum throughput of each client is given by $U_{m,n,r,t}^i$.

The system aims to increase the weighted sum throughput of the system. The best joint UL/DL NOMA distribution policy is obtained by mixed-integer linear programming, problem for TSP is represented as:

$$Q(x) = \text{maximize}_{p,q} \sum_{i=1}^{N_F}\sum_{m=1}^{K}\sum_{n=1}^{K}\sum_{r=1}^{J}\sum_{t=1}^{J} U_{m,n,r,t}^i , \quad (17)$$

subject to:

$$C1 : s_{m,n,r,t}^i \in [0,1], \forall_{i,m,n,r,t} , \quad (18)$$

$$C2 : \sum_{i=1}^{N_F}\sum_{m=1}^{K}\sum_{n=1}^{K}\sum_{r=1}^{J}\sum_{t=1}^{J} s_{m,n,r,t}^i (P_m^i + P_n^i) \leqslant P_{max}^{DL} , \quad (19)$$

$$C3 : \sum_{i=1}^{N_F}\sum_{m=1}^{K}\sum_{n=1}^{K}\sum_{r=1}^{J}\sum_{t=1}^{J} s_{m,n,r,t}^i (P_r^i + P_t^i) \leqslant P_{max}^{UL} , \quad (20)$$

$$C4 : \sum_{i=1}^{N_F}\sum_{m=1}^{K}\sum_{n=1}^{K}\sum_{r=1}^{J}\sum_{t=1}^{J} s_{m,n,r,t}^i \leqslant 1, \ \forall i , \quad (21)$$

$$C5 : P_m^i \geqslant 0, \ \forall \ i,m , \quad (22)$$

$$C6 : P_r^i \geqslant 0, \ \forall \ i,r . \quad (23)$$

If $s_{m,n,r,t}^i = (0,1)$, then C1 assures effective SIC at DL operator n. For the reception of UL, since the receiver for all UL signals is the FD-BS, it can accomplish SIC in any order. For the BS, the C2 constraint is the power constraint through an extreme allowance of power transmission P_{max}^{DL}. By using P_{max}^{UL}, C3 bounds the transfer power of UL user r. To guarantee that each subcarrier constraint C4 is imposed, it is allocated to the top two DL and UL clients. The client pairings of DL, UL-to-DL, and UL are accomplished on each subcarrier. For the UL and DL clients, C5 and C6 are said to be the non-negative power transmission constraints.

4.4. Modified Particle Swarm Optimization (MPSO) with Inertia Weight

Eberhart and Kennedy discovered a particle swarm optimization (PSO) algorithm relying on a population-based, cooperative search metaheuristic procedure. PSO particles are known as the population's candidate solutions in which it coincides and develops instantly according to the sharing of knowledge from neighboring particles. PSO is a modern optimization algorithm, but when the problem dimension arises, it normally requires some enhancements [31].

In this paper, an altered PSO algorithm is proposed. For planning a modern multi-stage exception expansion, MPSO is applied here. Also, for multi-stage planning, some groups of particles are separated from the population in the altered PSO algorithm. Here, x_{ij} represents the position vector of the j-th particle of the i-th group. An intermediate network is optimized by each group of particles during single stage iterations. Therefore, the number of planning stages and groups tends to be similar. The particles fly sequentially from dissimilar groups (i.e. dissimilar phases). From its own last position, a j-th particle of the first group starts to move in every single iteration, but for $i>1$, from the last position of a j-th particle of $i-1$ group, the j-th particle of the i-th group starts to move. Equation 16 is used to calculate the objective function of each particle for the i-th stage in the i-th group. The objective function $U_{m,n,r,t}^i$ (where $s=i$) is presented x_i^{LB} as a local best position and it is minimized by the better position vector of the i-th group. The overall explanations are gained from the positions of sequential particles of all groups. The number of particles of each group is equal to the number of newly created particles for every iteration.

Equation (17) is used to compute the objective function of every particle. Then, the better solution is predicted. The best particles are presented x_i^{GB} as a better global position. The velocity vector of a better existing position in conventional PSO is calculated as:

$$v_{i,n+1} = w * v_{i,n} + C1 * rd_1(P_{i,n} - x_{i,n}^{LB}) + C2 * rd_2(P_{i,n} - x_{i,n}^{GB}) , \quad (24)$$

$$v_{i,n} = \begin{bmatrix} (1-\frac{t}{T})V_m ; & \text{if } v_{i,n} > V_m \\ -(1-\frac{t}{T})V_m ; & \text{if } v_{i,n} < -V_m \end{bmatrix} , \quad (25)$$

$$x_{i,n+1} = x_{i,n} + v_{i,n+1} . \quad (26)$$

Apart from the scaling term $1-\frac{t}{T}$, an altered algorithm of PSO is almost similar to the original one in the scheme; and in Eq. (25) it is multiplied with maximum velocity V_m. The proceeded maximum number of generations processed is mentioned using T and the number of the current generation is symbolized using t.

A significant component inertia weight in particle swarm is used to overcome the local optimal problem and slow rate of convergence of PSO. Inertia weight is one of the important factors to influence the convergence speed and searching for outcomes. The global search ability is better for PSO and when the weight of inertia is higher, then the function of the concrete value is to improve the rate of convergence. Local search ability is also enhanced when the inertia weight is insignificant. This means that an enhanced solution may be achieved immediately after the local search algorithm is completed. Early convergence is processed quickly by the PSO algorithm as well. In the PSO algorithm, we initially fixed a large inertia weight value and in the global scope, to guess the series of the optimum value. To perform the optimal value search for the algorithm, we set a lower value of inertia weight w, so that the algorithm offers faster convergence and better search outcomes. The function of the modified inertia weight is:

$$w(x) = \frac{n.\alpha}{n + x^a} + \beta \ , \tag{27}$$

where the rate of change and acceleration factor is defined by a, the range of w is controlled by the threshold values (positive constant parameters) α and β, the current iteration number and the algorithm iteration number are defined by x and n, respectively.

The following equations are used to verify the validity of this function:

$$w(x)' = \frac{-n\alpha a x^{a-1}}{(n + x^a)^2} \ , \tag{28}$$

$$w(x)'' = \frac{n\alpha a x^{\gamma-2}(n + x^a)[(a+1)x^a - (a-1)n]}{(n + x^a)^4} \ . \tag{29}$$

When x is greater than $\sqrt{\frac{(\gamma-1)n}{\gamma+1}}$ then the significance of $w(x)''$ is in excess of 0 and $w(x)'$ is lower than 0 as per Eqs. (28) and (29).

The weight of the inertia function is a convex, as well as descending function besides it, is observed from the above formula that the descending speed steadily slows down when the iteration number increases. Without affecting the precision of convergence, it significantly raises the rate of convergence of this algorithm.

The values of α and β are set to 0.8 and 0.5, because the range of inertia weight w ranges from 0.5 to 1.2 in PSO. In an existing space, searching for the smaller value of w takes place but the bigger w can be searched in the new spaces. The inertia weight variable is suitably picked for balancing both the local and global search. In Eq. (27), a different value of a is displayed for the inertia weight function.

In the process of its execution, the algorithm preserves two superior variables named g-best and l-best position. Two comparisons are performed: to decide the g-best location of every creation in the entire population, each particle's fitness at its current position is related to the remaining particles' fitness. Then, to choose the l-best position for every particle, the current location of separate particles is contrasted with diverse visiting positions. Based on Eq. (24), the refining velocity of every particle in the species of particle group is realized by these two positions. To update the speed rate of the fresh particle, two stochastic variables outweigh the effect of two locations.

The pseudo-code with the M-PSO algorithm for task scheduling is presented as Algorithm 1.

Algorithm 1. Modified PSO for joint UL/DL channel allocation

1. **Initialization**. The population and iteration number are fixed as N and N_t, respectively. Within the predefined decision variable range, initialize velocity v_i and position x_i of the particles with random numbers. The upper bound of the decision variable is fixed at V_m. The fixed iteration count $t = 0$ and $p_i = x_i$ as personal better position.

2. **Estimation**. Every single particle in the current population is estimated. Set $t = t + 1$, $p_i = x_i$ when $Q(t) < Q(t-1)$. Find a corresponding position $x_{\min}$ and $Q_{\min} = \min Q(t)$. The global best is selected by using $x_{i,n}^{GB} = x_{\min}$.

3. **Generation of new particles**. Compute the objective function values for every single new particle and, depending upon the current x_i ($i = 1, 2, \ldots, N$), compute the new position x_i and velocity v_i. Associate new x_i ($2N$ particles) as well as all x_i together and collect them in a temporary list.

4. **Non-dominated sorting**. In *tempList* recognize dominated results and save them in a P_{front} (Pareto front) matrix. Fixed front number $k = 1$ and:

   (a) from the *tempList* the non-dominated particles are eliminated,

   (b) $k = k + 1$. The non-dominated results in the excepting *tempList* are recognized and are collected in a $P_{front}k$ (front k) matrix,

   (c) when all $2N$ particles get ranked into several fronts then stop the repeating steps b–c.

5. **For the next iteration, select particles**. From P_{front} randomly pick out N particles, if $P_{front}size > N$ and store them as next x_i. Or select random particles in next front (front k) and include them in next x_i until next x_i size becomes N.

6. For every particle in next x_i, compute objective function values and for the next iteration set the next x_i as the current position's x_i.

7. If all $v_i < 0.1V_m$, then implement subsequent steps or else go to step 8.

   (a) from the current population randomly pick 20% particles and modify their positions by 10% of the V_m. Finally, x_{temp} is used to store those results,

(b) x_{temp} gets estimated and all the dominant particles are identified. The particles in the current x_i are replaced by those dominating particles,

(c) it is definite that the number of x does not exceed N and repeat K times ($K = 1, 2, \ldots, 10$) steps a–b.

8. If $t < N_t$ go to step 2.

9. From the final population, the non-dominated solutions are stored and performance metric values are calculated.

The creation of randomized and legal influence of location is a major goal of these coefficients. So, once in a while it is very essential to find few examinations and at other times small exploitation stochastically. Depends upon a new speed with solving Eq. (26), this algorithm updates the particle's current position to a new value. Every single particle defines the PSO particle population's new state and reviews its position. Based on their new location, the fitness values are evaluated by the algorithm. The duplication of the processes is used to estimate the location of fresh particles and to predict the global and the local best positions which, in turn, are used to inform the position of the particles.

Table 2
Simulation parameters

Parameter	Description
System bandwidth	12 GHz
Number of subcarriers	5
Carrier frequency	3.6 GHz
Subcarrier separation	20 kHz
Sub-frame length	1.0 ms
Symbol duration	66.67 μs + cyclic prefix: 4.69 μs
Receiver type	MMSE+SIC
Number of users per cell	10
Number of PU per cell	3
Number of SU per cell	7
Inter-site distance	500 m
Maximum transmit power	46 dBm
Channel model	3GPP spatial channel model (SCM)
Path loss model	133.6 +35 log(d) [km]
Traffic model	Full buffer
Power factor	0.25
Size of swarm	30
Modulation technique	16 QAM
Maximum iteration	500
Encoding	Conventional

5. Performance Evaluation

The performance of the presented channel allocation method is examined using the simulation parameters from Table 2. Within the outer and inner boundaries, both the K and J users are uniformly and randomly dispensed. P_{max}^{DL} defines the extreme transmit power of the FD BS. Here, we incorporate the Rayleigh fading model in the UL NOMA for communication between users and the BS. Similarly, we have incorporated the Rician fading model in the DL NOMA for communication between the BS and users.

5.1. Throughput

Throughput is defined as the movement of data from one location to another, over a specific period of time. It is a key indicator of the effectiveness and quality of network connectivity. A high rate of failed message deliveries will eventually lead to low throughput and degraded performance. Decoding methods in NOMA systems, to decode various simultaneous transmissions, SIC which is a multiuser detection technique that uses the structured nature of interference. Separate signals are retrieved, one by one, from the composite signal in the following manner. It is questionable when the remaining signals are decoded if all of the signals fail to be decoded. Throughput depends on every single signal. The order of decoding also plays an essential role in the positive outcome of decoding operation.

The order of decoding the received superposition coded signals is not forced by the principle of NOMA. The system's advantages become visible when throughput decodes stronger signals ahead of their weaker counterparts.

However, due to the decoding difficulties of the SIC method and the changing nature of wireless channels, most users are allocated, undesirably, to subchannels which will intensify the system's throughput. Throughput of the presented system increases with the increasing number of users, as shown in Fig. 2 and is compared with some of the existing approaches [25].

Consistent throughput performance was obtained in the graph for the existing algorithms, where the number of users does not affect throughput to a considerable degree. Tree search-based transmission power allocation (TTPA) and fractional transmission power allocation (FTPA) are the conventional methods that work based on grouping, power allocation, and ordering. Figure 2 shows that the proposed system is affected by significant changes concerning the number of users. Typically, NOMA systems are intended for dense user networks. The proposed system is highly suitable here. Extreme throughput is achieved by users who gain the best of all subchannels. Similarly, the maximum throughput is achieved by the users who are nearer to the base station, as well as in overall subchannels, with the effect of Rayleigh fading being statistically parallel for all users.

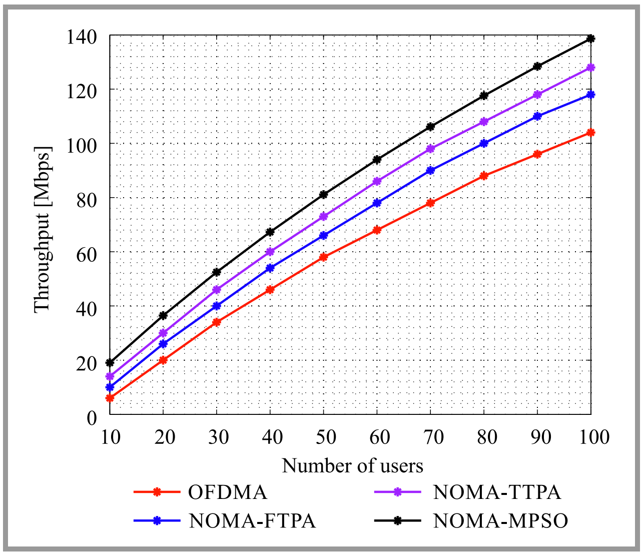


Fig. 2. Throughput vs. number of users.

5.2. Spectral Efficiency

Spectral efficiency is defined as the bits per second net data rate divided by bandwidth. Net data rate and index rate are associated with the raw data rate, including the payload and all overheads that can be used. Figure 3 shows the spectral efficiency versus the number of users. The NOMA-MPSO spectral efficiency exceeds the NOMA-TTPA, NOMA-FTPA, and OFDMA schemes and the presented methods of resource optimization tend to provide maximum spectral efficiency than the existing previous techniques [32], [33].

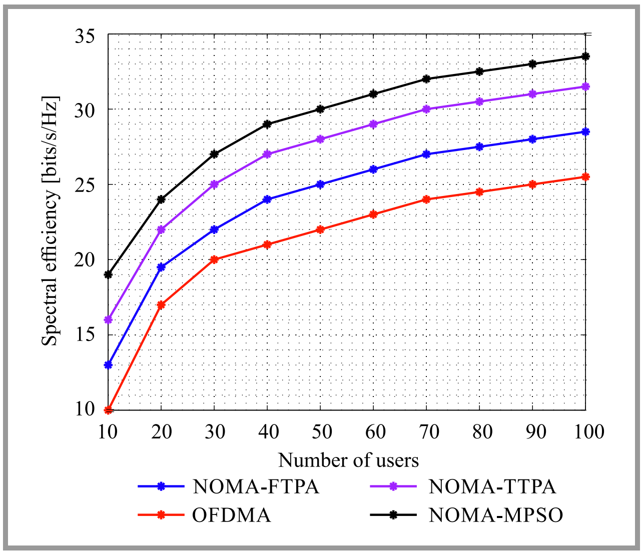


Fig. 3. Spectral efficiency vs. number of users.

Figure 3 shows that with the number of users, the efficiency of the spectrum rises and the growth of data rate becomes slower with the rise in the number of users. It endures rising in the total sum rate when the number of users is greater than the number of sub-channels as well as it grows at a smaller speed due to the gain of multiuser diversity. When the user number is insignificant then the effect of diversity of multiuser is more remarkable.

5.3. Sum-Rate

The sum rate is defined as the sum of all rates of communication between the nodes, taking place in a network:

$$R_T = \sum_{k=1}^{K} R_k \ , \tag{30}$$

where R_k is the k-th user equivalent sum-rate and R_T is the sum-rate obtained by using the proposed approach. The highest sum-rate for users will be achieved by conducting all communication at once, or there must be some scheduling between the different tasks. The sum-rate is maximized when the system is operating continuously, in the full-duplex mode.

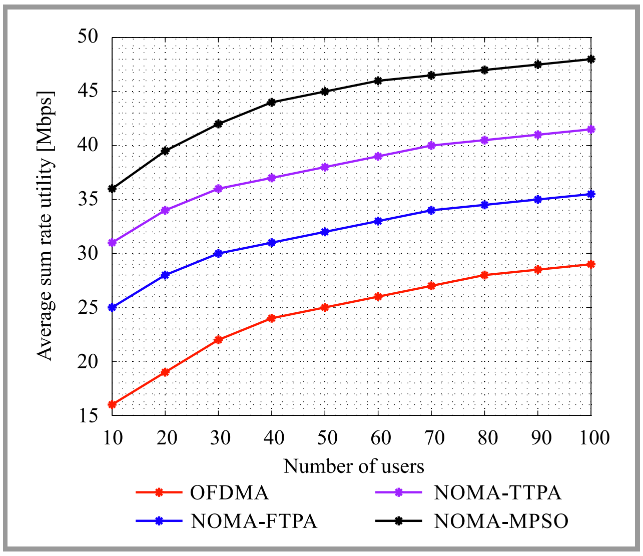


Fig. 4. Sum rate vs. number of users.

Figure 4 shows R_T, the total rate of the user, the number of UEs per cell function, is got through system-level simulations for the presented scheme of NOMA-MIMO using MPSO at the SIC receiver at both ends and transmitter side of BS. We put in a simple assessment model using the Shannon capacity [34]. Using NOMA, we assess a case accounting OFDMA, in which the transmission of a single-stream is applied per transmitter beam. In addition, MPSO channel allocation methods are compared with the exhaustive NOMA-FTPA and NOMA-TTPA searches. The NOMA-MPSO technique is suitable for finding the constraints of QoS and the sum-rate with weights in which the majority of the gap is less than 5% and it is very adjacent to the globally optimum value. Hence, the suggested joint UL/DL channel assignment technique is capable of attaining approximate optimal performance with fewer difficulties.

5.4. Outage Probability

Outage probability is the achieved data rate of an individual user which is lower than the predefined value. For common outage probability, an outage occurs when the user is deactivated. In individual outage probability, individual user outage events are considered. Outage probability is used as a performance measure, since it not only allows for the identification of the probability of errors, but may be also used to evaluate the outage capacity/rate. After 20 dB, the graph gets saturated due to the proper achievement of SNR in this region. The effect of the user's non-uniform locations and the interference is caught by applying stochastic geometry, and the order of diversity is computed to demonstrate efficient use of the channel's degree of freedom by the presented framework [35]. In Fig. 5, the scheme in question is compared with NOMA-TTPA, NOMA-FTPA, OFDMA. One may notice that their outage sum-rate performance is similar, up to a certain reduced number of users, but when the number of users increases, outage probability changes considerably. However, the NOMA-MPSO scheme is capable of offering much better reception reliability, certainly for maximum power transmissions.

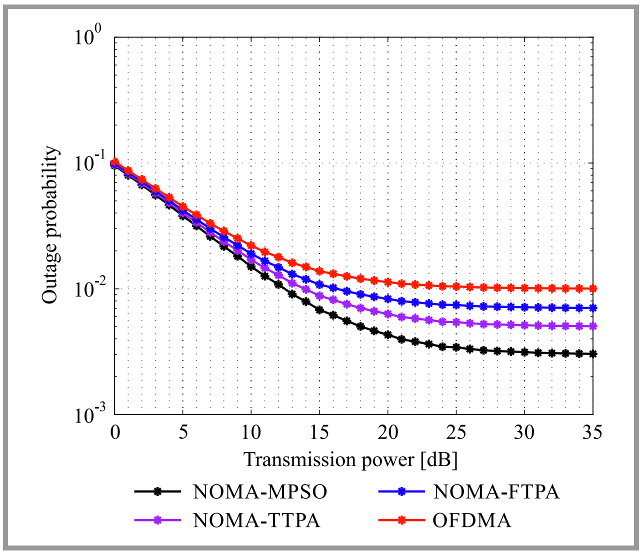


Fig. 5. Outage probability vs. transmission power.

5.5. SINR

SINR measures the quality of a transmission channel. SINR is generally defined for a specific user and is represented as:

$$SINR = \frac{P}{I+N} , \tag{31}$$

where incoming signal, interference signal, and noise are denoted as P, I, and N, respectively. Noise cancelation is impossible under separate constraints of power, but the possibility of the set of target SINRs may be under a sum power constraint. Because of the unavailability of power constraints, arbitrary target SINRs may be achieved. This is caused by the cascaded structure of interference formed by the successive decoding operations. To accomplish a certain set of SINRs, an equal amount of total power is essential for both links. Under a sum power constraint, both links have an equal achievable SINR range. Likewise, the same beamformers, accomplish the target [36]. Figure 6 illustrates the performance of the offered method BER unit and compares its performance with existing algorithms OFDMA, NOMA-FTPA, and NOMA-TTPA. BER performance is improved in the proposed system.

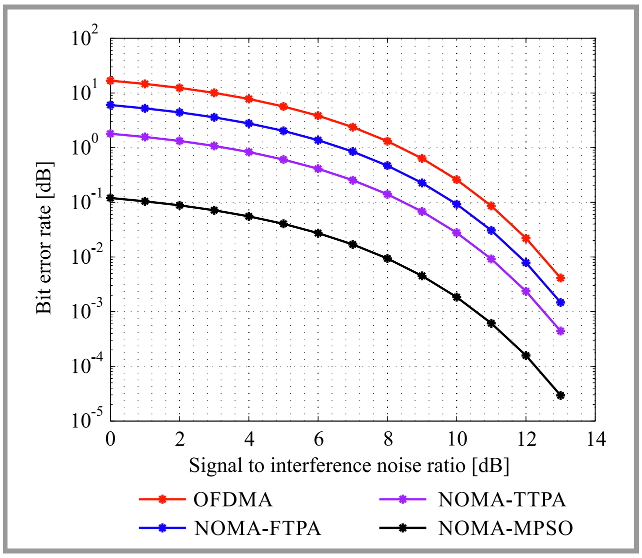


Fig. 6. BER vs. SINR.

5.6. Fairness

This is the most popular metric used in network engineering to determine if users or applications receive a fair share of he system's resources. Fairness is defined as:

$$Fairness(r_1, r_2, \ldots, r_n) = \frac{(\sum_i r_i)2}{n \sum_i r_i 2} , \tag{32}$$

where the throughput individual nodes is denoted as $r_1, r_2, \ldots, r_n$. Based on the scheduling period, the fairness of the proposed method is examined. The time domain of the scheduling process is slotted. The time slot index is denoted by t. With 20 slots, we interpret a scheduling frame. Channel state information is grouped once per frame. The Jain's fairness index is calculated by $\frac{(\sum_{k=1}^{K} \overline{r_k})2}{K \sum_{k=1}^{K} \overline{r_k} 2}$, when the average user's rates are $\overline{r_1}, \ldots, \overline{r_k}$, at the end of the period of schedule. In network communications, this index, developed in [37] and is utilized for user throughput as a measure of fairness value from $\frac{1}{K}$ and 1.0 is reached. Fairer throughput distribution is specified by higher values. This index will drop the index value but doesn't evade a user from being assisted with low throughput (or even zero throughputs) but it will bring down the index value.

The fairness index for the number of users is shown in Fig. 7. Thanks to the higher level of competition between the users, fairness degradation in all schemes is caused by an increase in the number of users [32]. But in the scheme we propose, the slope of the curve is decreasing slightly

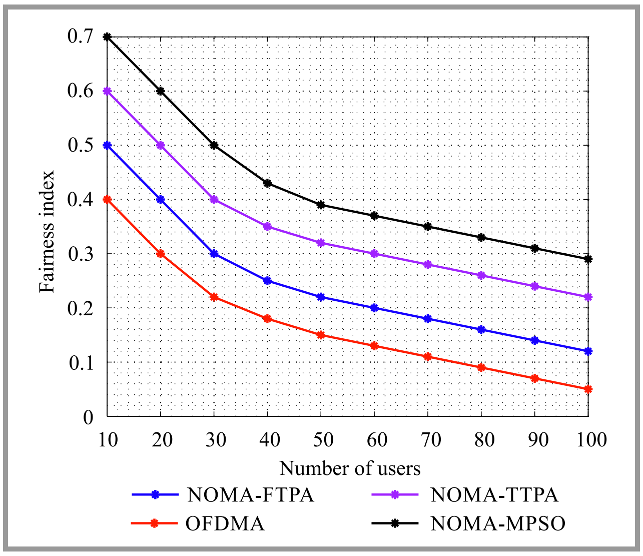


Fig. 7. Fairness index vs. number of users.

with an increase in the number of users, which means that fairness level is improved compared to that of the existing schemes, i.e. OFDMA, NOMA-TTPA, and NOMA- FTPA.

6. Conclusion

This paper discusses the design of FD BS for the MIMO-NOMA system channel allocation algorithm. For the enlargement of the weighted sum system throughput, the model of the algorithm is generated as a mixed combinatorial non-convex optimization issue. Based on various workloads and task scheduling approaches, an M-PSO algorithm is developed in this paper. The inner weight factor plays a vital role in M-PSO, where the higher value of inertia weight is performed as global search and the small weight of inertia value performed as local search. The M-PSO algorithm with a greater number of users achieves better results than the same algorithm with a few users. Moreover, the presented FD MC-NOMA approach was proven to offer a perfect balance between maintaining fairness and improving the system's throughput among users. The proposed FD MC-NOMA M-PSO scheme offers better performance in terms of throughput, fairness, sum-rate, and spectral efficiency for a given number of users.

References

[1] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the Internet of Things: Communication technologies and challenges", *IEEE Access*, vol. 6, pp. 3619–3647, 2018 (DOI: 10.1109/ACCESS.2017.2779844).

[2] A. Boulogeorgos *et al.*, "Wireless terahertz system architectures for networks beyond 5G", *arXiv preprint arXiv:1810.12260*, 2018.

[3] M. A. Khan *et al.*, "Understanding autonomic network management: A look into the past, a solution for the future", *Computer Commun.*, vol. 122, pp. 93–117, 2018 (DOI: 10.1016/j.comcom.2018.01.014).

[4] T. Shuminoski, S. Kitanov, and T. Janevski, "Advanced QoS provisioning and mobile fog computing for 5G", *Wirel. Commun. and Mobi. Comput.*, vol. 2018, article ID 2109394, 2018 (DOI: 10.1155/2018/5109394).

[5] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches", *IEEE Wirel. Commun.*, vol. 22, no. 5, pp. 22–31, 2015 (DOI: 10.1109/MWC.2015.7306534).

[6] I. Aldmour, "Wireless broadband tools and their evolution towards 5G networks", *Wirel. Personal Commun.*, vol. 95, no. 4, pp. 4185–4210, 2017 (DOI: 10.1007/s11277-017-4058-x).

[7] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges", *IEEE Commun. Surveys & Tutor.*, vol. 19, no. 2, pp. 721–742, 2016 (DOI: 10.1109/COMST.2016.2621116).

[8] C. Sexton *et al.*, "5G: Adaptable networks enabled by versatile radio access technologies", *IEEE Commun. Surveys & Tutor.*, vol. 19, no. 2, pp. 688–720, 2017 (DOI: 10.1109/COMST.2017.2652495).

[9] N. Bui *et al.*, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques", *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 3, pp. 1790–1821, 2017 (DOI: 10.1109/COMST.2017.2694140).

[10] T. H. Naveen and G. Vasanth, "Qualitative study of existing research techniques on wireless mesh network", *Int. J. of Adv. Comp. Sci. and Appl.*, vol. 8, no. 3, pp. 49–57, 2017 (DOI:10.14569/IJACSA.2017.080308).

[11] A. BenMimoune and M. Kadoch, "Relay technology for 5G networks and IoT applications", in *Internet of Things: Novel Advances and Envisioned Applications*, D. P. Acharjya and M. K. Geetha, Eds. *Studies in Big Data*, vol. 25, pp. 3–26. Springer, 2017 (DOI: 10.1007/978-3-319-53472-5_1).

[12] E. Ezhilarasan and M. Dinakaran, "A review on mobile technologies: 3G, 4G and 5G", in *Proc. 2nd Int. Conf. on Recent Trends and Challen. in Comput. Models ICRTCCM 2017*, Tindivanam, India, 2017, pp. 369–373 (DOI: 10.1109/ICRTCCM.2017.90).

[13] W. Xia, Y. Zhou, G. Yang, and R. T. Chen, "Power-balanced non-orthogonal multiple access based on virtual channel optimization", *IEEE Trans. on Circ. and Syst. II: Express Briefs*, vol. 67, no. 4, pp. 795–799, 2019 (DOI: 10.1109/TCSII.2019.2925270).

[14] M. Dighriri, G. M. Lee, and T. Baker, "Measurement and classification of smart systems data traffic over 5G mobile networks", in *Technology for Smart Futures*, M. Dastbaz, H. Arabnia, and B. Akhgar, Eds. Springer, 2018, pp. 195–217 (DOI: 10.1007/978-3-319-60137-3_9).

[15] Z. J. Ali, N. K. Noordin, A. Sali, and F. Hashim, "Fair energy-efficient resource allocation for downlink NOMA heterogeneous networks", *IEEE Access*, vol. 8, pp. 200129–200145, 2020 (DOI: 10.1109/ACCESS.2020.3035212).

[16] I. Ahmad, Z. Kaleem, R. Narmeen, L. D. Nguyen, and D.-B. Ha, "Quality-of-service aware game theory-based uplink power control for 5G heterogeneous networks", *Mob. Netw. and Appl.*, vol. 24, no. 2, pp. 556–563, 2019 (DOI: 10.1007/s11036-018-1156-2).

[17] S. Han, X. Xu, X. Tao, and P. Zhang, "Joint power and sub-channel allocation for secure transmission in NOMA-based mMTC networks", *IEEE Syst. J.*, vol. 13, no. 3, pp. 2476–2487, 2019 (DOI: 10.1109/JSYST.2018.2890039).

[18] Y. Shi, J. Zhang, W. Chen, and K. B. Letaief, "Generalized sparse and low-rank optimization for ultra-dense networks", *IEEE Commun. Magazine*, vol. 56, no. 6, pp. 42–48, 2018 (DOI: 10.1109/MCOM.2018.1700472).

[19] N. U. Hasan *et al.*, "Network selection and channel allocation for spectrum sharing in 5G heterogeneous networks", *IEEE Access*, vol. 4, pp. 980–992, 2016 (DOI: 10.1109/ACCESS.2016.2533394).

[20] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization", in *Proc. IEEE Globecom Worksh. GC Wkshps 2013*, Atlanta, GA, USA, 2013, pp. 129–134 (DOI: 10.1109/GLOCOMW.2013.6824974).

[21] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: Opportunities and challenges", *IEEE Wirel. Commun.*, vol. 24, no. 3, pp. 86–94, 2017 (DOI: 10.1109/MWC.2016.1500374WC).

[22] M. Hadi and R. Ghazizadeh, "Sub-channel assignment and power allocation in OFDMA-NOMA based heterogeneous cellular networks", *AEU – Int. J. of Electron. and Commun.*, vol. 120, article 153195, 2020 (DOI: 10.1016/j.aeue.2020.153195).

[23] H. Tabassum *et al.*, "Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques", arXiv preprint arXiv:1608.05783, 2016.

[24] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems", *IEEE J. on Selec. Areas in Commun.*, vol. 35, no. 12, pp. 2744–2757, 2017 (DOI: 10.1109/JSAC.2017.2725618).

[25] Y. Saito *et al.*, "Non-orthogonal multiple access (NOMA) for cellular future radio access", in *Proc. IEEE 77th Veh. Technol. Conf. VTC Spring 2013*, Dresden, Germany, 2013 (DOI: 10.1109/VTCSpring.2013.6692652).

[26] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama, and T. Nakamura, "Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials", in *Proc. Int. Conf. on Wirel. Netw. and Mob. Commun. WINCOM 2015*, Marrakech, Morocco, 2015 (DOI: 10.1109/WINCOM.2015.7381343).

[27] P. D. Diamantoulakis, K. N. Pappi, G. K. Karagiannidis, H. Xing, and A. Nallanathan, "Joint downlink/uplink design for wireless powered networks with interference", *IEEE Access*, vol. 5, pp. 1534–1547, 2017 (DOI: 10.1109/ACCESS.2017.2657801).

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

[29] A. Memarinejad, M. Mohammadi, and M. B. Tavakoli, "Full-duplex NOMA cellular networks: Beamforming design and user scheduling", *AEU – Int. J. of Electron. and Commun.*, vol. 126, article 153415, 2020 (DOI:10.1016/j.aeue.2020.153415).

[30] P. S. Shelokar, P. Siarry, V. K. Jayaraman, and B. D. Kulkarni, "Particle swarm and ant colony algorithms hybridized for improved continuous optimization", *Applied Mathem. and Comput.*, vol. 188, no. 1, pp. 129–142, 2007 (DOI: 10.1016/j.amc.2006.09.098).

[31] M. Masdari, F. Salehi, M. Jalali, and M. Bidaki, "A survey of PSO-based scheduling algorithms in cloud computing", *J. of Network and Sys. Manag.*, vol. 25, no. 1, pp. 122–158, 2017 (DOI: 10.1007/s10922-016-9385-9).

[32] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks", *IEEE Trans. on Wirel. Commun.*, vol. 15, no. 11, pp. 7686–7698, 2016 (DOI: 10.1109/TWC.2016.2606100).

[33] M. Jain, S. Soni, N. Sharma, and D. Rawal, "Performance analysis at far and near user in NOMA based system in presence of SIC error", *AEU – Int. J. of Electron. and Commun.*, vol. 114, article 152993, 2020 (DOI: 10.1016/j.aeue.2019.152993).

[34] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Power and channel allocation for non-orthogonal multiple access in 5G systems: Tractability and computation", *IEEE Trans. on Wirel. Commun.*, vol. 15, no. 12, pp. 8580–8594, 2016 (DOI: 10.1109/TWC.2016.2616310).

[35] Z. Ding, R. Schober, and H, V. Poor, "On the design of MIMO-NOMA downlink and uplink transmission", in *Proc. IEEE Int. Conf. on Commun. ICC 2016*, Kuala Lumpur, Malaysia, 2016 (DOI: 10.1109/ICC.2016.7510759).

[36] M. Aldababsa and O. Kucur, "BER performance of NOMA network with majority based JTRAS scheme in practical impairments", *AEU – Int. J. of Electron. and Commun.*, vol. 129, article 153523, 2021 (DOI: 10.1016/j.aeue.2020.153523).

[37] R. K. Jain *et al.*, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system", Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, 1984 [Online]. Available: https://www.cs.wustl.edu/~jain/papers/ftp/fairness.pdf

Nirmalkumar S. Benni is currently an Assistant Professor at the School of Electronics and Communication Engineering, REVA University, Bengaluru, India. He has completed an M.Tech. degree in Digital Communication & Networking, UBDT, Davangere, India in 2007, from Kuvempu University, Shimoga and a B.E. degree in Electronics and Communication Engineering, Hirasugar Institute of Technology, Nidasoshi, Belagavi, India, in 2005, from VTU, Belagavi. He is pursuing his Ph.D. in Electronics and Communication Engineering in the area of wireless communication and networks.

E-mail: nirmalkumar.benni@gmail.com
Department of Electronics and Communication Engineering
REVA Institute of Technology and Management
Rukimini Knowledge Park, Kattigenahalli, SH 104
Srinivasa Nagar, Bangalore
Karnataka 560064, India

Sunilkumar S. Manvi received his B.E. degree from Karnataka University in 1987, M.E. degree in Electronics from the University of Visweshwariah College of Engineering, Bangalore, in 1993 and Ph.D. degree in Electrical Communication Engineering, from the Indian Institute of Science, Bangalore, India in 2003. He is currently working as a Principal Investigator at the Wireless Information Systems Research Lab, Principal, REVA Institute of Technology and Management, and a Director of the School of Computing and Information Technology at REVA University. Bangalore, India. His research interests fokus on software agent-based network management, wireless networks, multimedia networking, underwater networks, wireless network security, grid and cloud computing, e-commerce, and mobile computing.

E-mail: ssmanvi@reva.edu.in
Department of Electronics and Communication Engineering
REVA Institute of Technology and Management
Rukimini Knowledge Park, Kattigenahalli, SH 104
Srinivasa Nagar, Bangalore
Karnataka 560064, India

Paper

Secrecy Rate Region Enhancement in Multiple Access Wiretap Channel

Shahid Mehraj Shah

Department of Electronics and Communication Engineering, National Institute of Technology, Srinagar, Jammu and Kashmir, India

https://doi.org/10.26636/jtit.2022.152021

Abstract—It is commonly known that physical layer security is achieved with a trade-off in terms of the achievable rate. Hence, security constraints generate rate losses in wiretap channels. To mitigate such rate losses in multi-user channels, we propose a coding/decoding scheme for multi-user multiple access wiretap channel (MAC-WT), where previously transmitted messages are used as a secret key to enhance the secrecy rates of the transmitting users, until the usual Shannon capacity region of a multiple access channel (MAC) is achieved without the secrecy constraint. With this coding scheme, all messages transmitted in the recent past are secure with respect to all the information of the eavesdropper till now. To achieve this goal, we introduce secret key buffers at both the users and the legitimate receiver. Finally, we consider a fading MAC-WT and show that with this coding/decoding scheme, we can achieve the capacity region of a fading MAC channel (in the ergodic sense).

Keywords—multiple access channel, physical layer security, strong secrecy, wiretap channel.

1. Introduction

In [1], Wyner proved, degraded wiretap channel, that by assigning multiple codewords to a single message, we can achieve reliability as well as security in a point-to-point channel communication and characterized secrecy capacity for this channel. After decades of work commenced after the wireless revolution had begun, researchers started extending Wyner's coding scheme (wiretap coding) in different directions. A single user fading wiretap channel was studied in [2], [3]. A secret key buffer was used in [4] to mitigate the fluctuations in the secrecy capacity due to the variations in the channel's gain over time.

A multiple access channel (MAC) with security constraints was studied in [5] and [6]. In [5], the transmitting users treat each other as eavesdroppers (Eve) and an achievable secrecy rate region is characterized. In some special cases the secrecy capacity region is also found. In [6], the authors consider the eavesdropper to be listening at the receiving end. The authors provide an achievable secrecy-rate region. The secrecy-capacity region is not known for such a MAC. The same authors also studied a fading MAC with full channel state information (CSI) of Eve known at the transmitters. Paper [7] is a research extension of a scenario in which the CSI of Eve is not known at the transmitters. For a detailed review of theoretical information theoretic security see [8], [9], and [10].

In all of the above mentioned papers, a notion of weak secrecy was used, i.e. if M is the message transmitted and the eavesdropper receives Z^n for a codeword of length n channel uses, then $I(M;Z^n)/n \rightarrow 0$, as $n \rightarrow \infty$. This notion of secrecy is not stringent enough in various cases [9]. In [11], Maurer proposed a notion of *strong secrecy*: $I(W;Z^n) \rightarrow 0$ as $n \rightarrow \infty$. For a point-to-point channel, he showed that it can be achieved without any change in secrecy capacity. Since then, other methods have been proposed for achieving strong secrecy [12], [13] and [14]. The methods shown in [12] and [14] have been used to obtain strong secrecy for a MAC-WT in [15] and [16], respectively.

In all these works one may notice that security is achieved at the cost of transmission rate. For a single user AWGN wiretap channel, if C_b is the capacity of the legitimate receiver (Bob) and C_e is the capacity of Eve's channel, then the secrecy capacity of this channel is $C_s = (C_b - C_e)^+$, where $(x)^+ = \max(0,x)$ [17]. In recent years, some work has been devoted to mitigating the secrecy rate loss. A feedback channel is used in [18] and [19] to enhance the secrecy rate, and under certain conditions the authors prove that the secrecy capacity can approach the main channel capacity. In [20], the authors assume that the transmitter (Alice) and Bob have access to a secret key, and then they propose a coding scheme which utilizes that key to enhance the secrecy rate. A secure multiplex scheme has been proposed in [21] which achieves Shannon channel capacity for a point-to-point wiretap channel. In this model, multiple messages are transmitted. The authors show that the mutual information of the currently transmitted message with respect to (w.r.t.) all the information received by Eve goes to zero as the codeword length $n \rightarrow \infty$. In [22], Shah *et al.* propose a simple coding scheme, without any feedback channels or access to a specific key, and enhance the secrecy capacity of a wiretap channel to the Shannon capacity of the main channel. In this work, only the message currently being transmitted is secure with respect to all information possessed by Eve.

In [23], the coding scheme of [22] was extended to a multiple access wiretap channel and it was shown that the Ahleswede-Liao region of the MAC can be achieved as the secrecy rate region, while keeping the currently transmitted message secure with respect to all information of Eve. In this paper, we extend the coding/decoding schemes of [22] and [23] to a multiple access wiretap channel and prove that we can achieve the Ahleswede-Liao region [24], [25] of the MAC as the secrecy-rate region, while keeping all recent messages secure with respect to the information possessed by Eve until the present. Finally, we achieve the same for a fading MAC-WT.

The remaining part of the paper is organized as follows. In Section 2, we define the channel model and recall some previous results which will be used in this paper. We extend the coding/decoding scheme from [22] to two user discrete memoryless MAC-WT (DM-MAC-WT) in Section 3, and prove the achievability of the Ahleswede-Liao region, under the security constraint that only the currently transmitted message is secured with respect to all data received by Eve. In Section 4, we consider a two-user DM-MAC-WT where each user, i.e. the receiver as well as Eve, has infinite length buffers to store previous messages. We propose a coding scheme to enhance the secrecy rate region to the Ahleswede-Liao region of the usual MAC, this time with the security constraint that *all recent* messages are secure with respect to all information possessed by Eve. In Section 5, we consider a two-user fading MAC-WT and extend the coding scheme from the previous sections to enhance the secrecy-rate region of the fading MAC-WT to the Ahleswede-Liao region of the MAC in the ergodic sense. Section 6 concludes the paper. The Appendix contains several lemmas used in the proofs of the main theorems.

In this paper, random variables will be denoted by capital letters X, Y, Z, vectors will be denoted with upper-bar letters, e.g. $\overline{X} = (X_1, \ldots, X_n)$, and scalar constants will be denoted by lower case letters a, b, etc.

2. Multiple Access Wiretap Channel

A discrete memoryless multiple access channel with a wiretapper and two users is considered (Fig. 1). The channel is represented by a transition probability matrix $p(y,z|x_1,x_2)$, where $x_i \in \mathscr{X}_i$ is the channel input from user i, $i = 1,2$, $y \in \mathscr{Y}$ is the channel output to Bob and $z \in \mathscr{Z}$ is the channel output to Eve. Sets $\mathscr{X}_1, \mathscr{X}_2, \mathscr{Y}, \mathscr{Z}$ are finite. The two transmitting users want to securely and reliably send messages $M^{(1)}$ and $M^{(2)}$ to Bob (legitimate receiver), while ensuring that eavesdropper (Eve) cannot decode the messages.

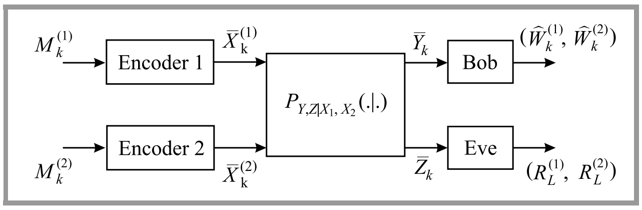


Fig. 1. Discrete memoryless multiple access wiretap channel.

Definition 2.1. For a MAC-WT, a $(2^{nR_1^{(s)}}, 2^{nR_2^{(s)}}, n)$ codebook comprises of (1) two sets of messages $\mathscr{M}^{(i)}$, $i = 1,2$, where cardinality of each message set is 2^{nR_i}, (2) uniformly distributed messages $M^{(1)}$ and $M^{(2)}$ over corresponding message sets (messages are assumed to be independent of each other), (3) two stochastic encoders:

$$f_i : \mathscr{M}^{(i)} \to \mathscr{X}_i^n, \quad i = 1,2, \tag{1}$$

and (4) a decoder at Bob:

$$g : \mathscr{Y}^n \to \mathscr{M}^{(1)} \times \mathscr{M}^{(2)}. \tag{2}$$

The decoded (estimated) messages are denoted by $(\hat{M}^{(1)}, \hat{M}^{(2)})$.

The average error probability at the receiver (Bob) is:

$$P_e^{(n)} \triangleq P\left\{\left(\hat{M}^{(1)}, \hat{M}^{(2)}\right) \neq \left(M^{(1)}, M^{(2)}\right)\right\}, \tag{3}$$

and leakage rate at Eve is:

$$R_L^{(n)} = \frac{1}{n} I(M^{(1)}, M^{(2)}; Z^n). \tag{4}$$

In [6], the authors have defined two types of security requirements depending upon the mutual trust of the transmitting users. If each user is conservative, such that when the other user is transmitting, then it may compromise with Eve and provide Eve with its codeword, then *individual leakage* constraints:

$$R_{L,1}^{(n)} = \frac{1}{n} I(M^{(1)};\ Z^n|\overline{X}^{(2)}), \tag{5}$$

$$R_{L,2}^{(n)} = \frac{1}{n} I(M^{(2)};\ Z^n|\overline{X}^{(1)}), \tag{6}$$

are relevant, where $\overline{X}^{(i)}$ denotes the codeword for user i.

In a scenario where the users trust each other, *collective leakage*:

$$R_L^{(n)} = \frac{1}{n} I(M^{(1)}, M^{(2)}; Z^n). \tag{7}$$

is relevant. Since, $M^{(1)} \perp M^{(2)}$ and, hence, also $\overline{X}^{(1)} \perp \overline{X}^{(2)}$, where $X \perp Y$ denotes that random variable X is independent of Y:

$$\begin{aligned} nR_L^{(n)} &= I(M^{(1)}, M^{(2)};\ Z^n) \\ &= I(M^{(1)};Z^n) + I(M^{(2)};\ Z^n|M^{(1)}) \\ &= H(M^{(1)}) - H(M^{(1)}|Z^n) + H(M^{(2)}) \\ &- H(M^{(2)}|Z^n, M^{(1)}) \\ &\leq H(M^{(1)}|X_2^n) - H(M^{(1)}|Z^n, X_2^n) \\ &+ H(M^{(2)}|X_1^n) - H(M^{(2)}|Z^n, X_1^n) \\ &= I(M^{(1)};\ Z^n|X_2^n) + I(M^{(2)};\ Z^n|X_1^n) \\ &= nR_{L,1}^{(n)} + nR_{L,2}^{(n)} \end{aligned} \tag{8}$$

and hence, if individual leakage rates are small, then so is the collective leakage rate. In this paper, we consider the secrecy notion (7).

Definition 2.2. The secrecy-rates $(R_1^{(s)}, R_2^{(s)})$ are achievable if there exists a sequence of codes $(2^{nR_1^{(s)}}, 2^{nR_2^{(s)}}, n)$ with $P_e^{(n)} \to 0$ as $n \to \infty$ and

$$\limsup_{n\to\infty} R_{L,i}^{(n)} = 0, \quad \text{for } i = 1,2. \tag{9}$$

By taking closure of convex-hull of the achievable secrecy-rate tuple $(R_1^{(s)}, R_2^{(s)})$, we obtain secrecy-capacity region for MAC-WT.

In [6], the authors propose a superposition coding-based scheme to obtain the following secrecy-rate region.

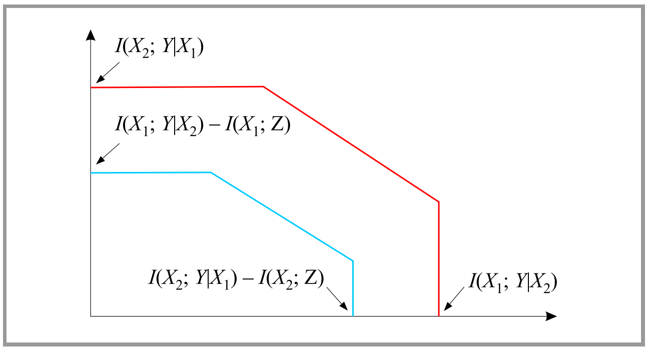


Fig. 2. Capacity region and secrecy rate region of MAC.

Theorem 2.1. Rates $(R_1^{(s)}, R_2^{(s)})$ are achievable with a $\limsup_{n\to\infty} R_{L,i}^{(n)} = 0$, $i = 1,2$, if there exist independent random variables (X_1, X_2) as channel inputs satisfying:

$$\begin{aligned} R_1^{(s)} &< I(X_1;Y|X_2) - I(X_1;Z), \\ R_2^{(s)} &< I(X_2;Y|X_1) - I(X_2;Z), \\ R_1^{(s)} + R_2^{(s)} &< I(X_1,X_2;Y) - I(X_1;Z) - I(X_2;Z), \end{aligned} \tag{10}$$

where Y and Z are the corresponding symbols received by Bob and Eve. □

The capacity region for a multiple access wiretap channel with the secrecy constraint is not known. Without the secrecy constraint, the capacity region for a multiple access channel is obtained by taking a convex closure of the regions in Theorem 2.1, excluding the terms $I(X_i;Z)$, $i = 1,2$ on the RHS of (10) (Fig. 2) [24]. In the following section, we show that we can achieve the Ahleswede-Liao capacity region of a MAC even when some modified metric of security metrics is satisfied. First, we restate the result of [24] and [25] which defines the capacity region for MAC without a security constraint, which we call as the Ahleswede-Liao capacity region.

Theorem 2.2. The capacity region of MAC is given by convex hull of rate pairs (R_1, R_2) satisfying:

$$\begin{aligned} R_1 &< I(X_1;Y|X_2)\ , \\ R_2 &< I(X_2;Y|X_1)\ , \\ R_1 + R_2 &< I(X_1,X_2;Y)\ . \end{aligned} \tag{11}$$

3. Mitigating Rate Loss in MAC-WT to Achieve Ahleswede-Liao Capacity

In this section, the coding scheme proposed in [22] for a single-user point-to-point wiretap channel is extended to the multiple transmitters case in order to enhance the achievable secrecy rates of discrete time MAC-WT. As in [22], we assume that the system is time slotted (i.e. each user transmits one message in one time slot), where each slot consists of n channel uses. In slot 1, the first message is encoded via point-to-point wiretap coding, as in [1]. In slot 2, the message transmitted in slot 1 is used as a secret key along with the usual wiretap code, and then the two messages are transmitted in that slot (the number of channels uses remains the same). Hence, the secrecy-rate is twice as high as in slot 1. We use the same coding scheme in the respective time slots *mutatis mutandis*, until the secrecy rate becomes saturated with the Shannon capacity of the main channel. After this time slot, the previously transmitted secure message is used as a key and no wiretap coding is used. We show that the proposed scheme ensures that the message currently being transmitted is secure with respect to all of Eve's outputs, i.e. if message M_k is securely transmitted in slot k then:

$$\frac{1}{n} I(M_k; \overline{Z}_1, \ldots, \overline{Z}_k) \to 0, \tag{12}$$

as the length of codeword $n \to \infty$, where $\overline{Z}_i$ is the information received by Eve in slot i.

Next, we not only extend this coding/decoding scheme to a multiple access wiretap channel, but also modify the scheme, so that it can be used to improve its secrecy criterion (12) and can be used for fading multiple access wiretap channels as well. The following secrecy criterion is used. In slot k, if user i transmits message $\overline{M}_k^{(i)}$, we need:

$$I(\overline{M}_l^{(1)}, \overline{M}_l^{(2)};\ \overline{Z}_1, \ldots, \overline{Z}_k) \leq n\varepsilon, \text{ for } l = 1, \ldots, k\ . \tag{13}$$

for any given $\varepsilon > 0$. This will be strengthened so that it satisfies strong secrecy requirement also, $I(\overline{M}_l^{(1)}, \overline{M}_l^{(2)};$

$\overline{Z}_1,\dots,\overline{Z}_k) \to 0$ as $n \to \infty$ at the end of the section. To achieve this goal, we modify the message sets and encoders/decoders with respect to Section 2 in the following manner.

Each slot has n channel uses and is divided into two parts. The first part has n_1 channel uses and the second n_2, $n_1 + n_2 = n$. The message sets are $\mathscr{M}^{(i)} = \{1,\dots,2^{n_1 R_i^s}\}$ for users $i = 1,2$, where (R_1^s, R_2^s) satisfy (10) for some (X_1, X_2). Here, there are two parts of the each encoder:

$$f_1^s : \mathscr{M}^{(1)} \to \mathscr{X}_1^{n_1}, \quad f_1^d : \mathscr{M}^{(1)} \times \mathscr{K}_1 \to \mathscr{X}_1^{n_2}\ , \tag{14}$$

$$f_2^s : \mathscr{M}^{(2)} \to \mathscr{X}_2^{n_1}\ , \quad f_2^d : \mathscr{M}^{(2)} \times \mathscr{K}_2 \to \mathscr{X}_2^{n_2}\ , \tag{15}$$

where $X_i \in \mathscr{X}_i$, $i = 1,2$, and $\mathscr{K}_i$, is the set of secret keys generated for the respective user $i = 1,2$, f_i^s, $i = 1,2$ are the usual wiretap encoders corresponding to each transmitting user, as in [6], and $f_i^d, i = 1,2$ are the deterministic encoders (used for channel models without security constraint) corresponding to each user in the usual MAC. User i may transmit multiple messages from $\mathscr{M}^{(i)}$ in a slot. In the first part of each slot of n_1 length, one message from $\mathscr{M}^{(i)}$ may be transmitted using wiretap coding via f_i^s (denoted by $\overline{M}_{k,1}^{(i)}$ in slot k) and in the second part multiple messages from $\mathscr{M}^{(i)}$ may be transmitted (denoted by $\overline{M}_{k,2}^{(i)}$) using messages transmitted in previous slots as keys. The overall message transmitted in slot k by user i is $\overline{M}_k^{(i)} = (\overline{M}_{k,1}^{(i)}, \overline{M}_{k,2}^{(i)})$.

Theorem 3.1. The secrecy rate region achieved while satisfying (13) is the Ahleswede-Liao capacity region for MAC, i.e. it is the closure of convex hull of all rate pairs $(R_1^{(s)}, R_2^{(s)})$ satisfying:

$$\begin{aligned} R_1^{(s)} &< I(X_1;Y|X_2), \\ R_2^{(s)} &< I(X_2;Y|X_1), \\ R_1^{(s)} + R_2^{(s)} &< I(X_1,X_2;Y), \end{aligned} \tag{16}$$

for some independent random variables X_1, X_2.

Proof. We fix distributions p_{X_1}, p_{X_2}. Initially we take $n_1 = n_2 = n/2$. In the first slot, i-th user selects message $\mathscr{M}_1^{(i)} \in \mathscr{M}^{(i)}$ to be securely transmitted in the first part of the slot, while the second part is not used. Both users use the multiple access wiretap coding scheme of [6]. Hence, the achievable rate pair $(R_1^{(s)}, R_2^{(s)})$ satisfies (10) and $R_{L,i}^{(n)} \leq n_1 \varepsilon$, $i = 1,2$. In slot 2, the two users select two messages each, $(\overline{M}_{2,1}^{(1)}, \overline{M}_{2,2}^{(1)})$ and $(\overline{M}_{2,1}^{(2)}, \overline{M}_{2,2}^{(2)})$ to be transmitted. Both transmitting users use the multiple access wiretap coding scheme (as in [6]) for the first part of the message, i.e. $(\overline{M}_{2,1}^{(1)}, \overline{M}_{2,1}^{(2)})$, and for the second part, transmitter i first takes XOR of $\overline{M}_{2,2}^{(i)}$ with the previous message, i.e. $\overline{M}_{2,2}^{(i)} \oplus \overline{M}_1^{(i)}$. This message (i.e. XOR of the second part and the previous message) is transmitted over the multiple access wiretap channel using a usual MAC coding scheme, i.e. without security [24], [25]. Therefore, the achievable secrecy rate in both sub-slots satisfies (10) for both the transmitting users. This achievable secrecy rate is also the overall rate of slot two.

In the third slot, the rate satisfies (10) in the first part (via wiretap coding). Since in the second part we XOR with message $\overline{M}_2^{(i)}$ and are able to send two messages, hence the rate of (10), assuming $2(R_1^{(s)}, R_2^{(s)})$ via (10), is within the range of (16).

Define:

$$\lambda_1 \triangleq \left\lceil \frac{I(X_1;Y|X_2)}{I(X_1;Y|X_2) - I(X_1;Z)} \right\rceil , \tag{17}$$

where $\lceil x \rceil$ is the ceiling of x, i.e. the smallest integer greater than or equal to x. In slot $\lambda_1 + 1$ the rate of user 1 in the second part of the slot satisfies:

$$\begin{aligned} R_1^{(s)} &\leq \min\left(\lambda_1 \left(I(X_1;Y|X_2) - I(X_1;Z)\right), I(X_1;Y|X_2)\right) \\ &= I(X_1;Y|X_2). \end{aligned} \tag{18}$$

Similarly, we define λ_2 as:

$$\lambda_2 \triangleq \left\lceil \frac{I(X_2;Y|X_1)}{I(X_2;Y|X_1) - I(X_2;Z)} \right\rceil . \tag{19}$$

In slot $\lambda_2 + 1$, the rate $R_2^{(s)}$ satisfies:

$$R_2^{(s)} \leq I(X_2;Y|X_1). \tag{20}$$

In slot $\lambda = \max\{\lambda_1, \lambda_2\} + 1$, the sum-rate will satisfy:

$$\begin{aligned} R_1^{(s)} + R_2^{(s)} \leq \min \Bigg\{ \lambda \left[I(X_1,X_2;Y) - \sum_{i=1}^{2} I(X_i;Z) \right], \\ I(X_1,X_2;Y) \Bigg\}. \end{aligned} \tag{21}$$

After some particular slot, say, λ^* which is greater than λ, the achievable secrecy sum-rate will get saturated by the Shannon sum-rate (i.e. sum-capacity of the usual MAC), i.e. $I(X_1,X_2;Y)$, and, hence, thereafter the rate pair $(R_1^{(s)}, R_2^{(s)}) \triangleq (R_1^{(s)*}, R_2^{(s)*})$ in the second part of the slot will be at a boundary point of (16) and the overall rate for the whole slot is the average of the rates in the first mini-slot and the second mini-slot.

In k-th slot, $(k > \lambda^*)$ to securely transmit a pair of messages $(\overline{M}_k^{(1)}, \overline{M}_k^{(2)})$, where $\overline{M}_k^{(i)} = (\overline{M}_{k,1}^{(i)}, \overline{M}_{k,2}^{(i)})$, $i = 1,2$, we

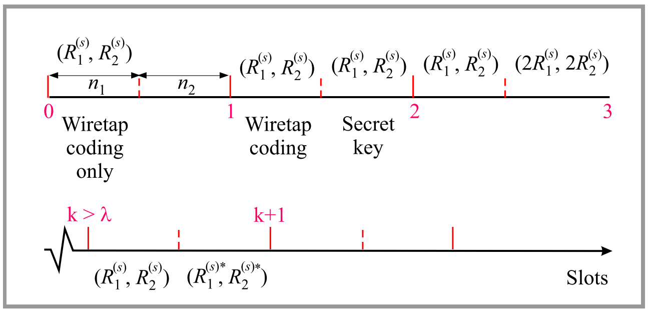


Fig. 3. Coding scheme to achieve Ahleswede-Liao region in MAC.

use the usual wiretap coding for $(\overline{M}_{k,1}^{(1)}, \overline{M}_{k,1}^{(2)})$ and for the second part of the message, we XOR it with the previously transmitted message, i.e. $\overline{M}_{k,2}^{(i)} \oplus \overline{M}_{k-1,2}^{(i)}$, $i = 1,2$, and transmit the overall codeword over the MAC-WT (Fig. 3). We let $n_2 = ln_1$ such that the overall rate of a slot is close to that in (16). Hence, by taking a sufficiently large l, one can achieve rates arbitrarily close to the boundary of (16). For the above-mentioned coding scheme, $P_e^n \to 0$. A convex combination of the achievable rates in (16) can be achieved by time sharing. Now, we show that our coding/decoding scheme also satisfies (13).

Leakage rate analysis. Before we compute the leakage rate, we set up the notation which will be used in the subsequent part of the proof. For transmitting user i, we represent the codeword sent in slot k by $\overline{X}_k^{(i)}$. Similarly, $\overline{X}_{k,1}^{(i)}$ and $\overline{X}_{k,2}^{(i)}$ will represent n_1-length and n_2-length codewords of i-th user i in slot number k. We define a notation here, when $i = 1$ then $\bar{i} = 2$ and when $i = 2$ then $\bar{i} = 1$. In k-th slot, the noisy version of the codeword received by Eve is $\overline{Z}_k \equiv (\overline{Z}_{k,1}, \overline{Z}_{k,2})$, where $\overline{Z}_{k,1}$ is the sequence corresponding to the wiretap coding part and $\overline{Z}_{k,2}$ is corresponding to the XOR part in which the previous message is used as a key.

Since wiretap coding of [6] is employed in slot 1, the leakage rate will satisfy:

$$I(\overline{M}_1^{(1)}; \overline{Z}_1 | \overline{X}_1^{(2)}) \leq n_1\varepsilon, \quad I(\overline{M}_1^{(2)}; \overline{Z}_1 | \overline{X}_1^{(1)}) \leq n_1\varepsilon\,. \tag{22}$$

For user 1 in slot 2, we show:

$$\begin{aligned} &I(\overline{M}_1^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \leq n_1\varepsilon\,, \\ &I(\overline{M}_2^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(1)}) \leq n_1\varepsilon\,. \end{aligned} \tag{23}$$

A similar calculation can be made for user 2.

First, we note that:

$$\begin{aligned} &I(\overline{M}_1^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \\ &= I(\overline{M}_1^{(1)}; \overline{Z}_1) + I(\overline{M}_1^{(1)}; \overline{Z}_2 | \overline{Z}_1, \overline{X}_2^{(2)}) \\ &\overset{(a)}{\leq} n_1\varepsilon + H(\overline{M}_1^{(1)} | \overline{Z}_1, \overline{X}_2^{(2)}) - H(\overline{M}_1^{(1)} | \overline{Z}_1, \overline{X}_2^{(2)}, \overline{Z}_2) \\ &\overset{(b)}{=} n_1\varepsilon + H(\overline{M}_1^{(1)} | \overline{Z}_1) - H(\overline{M}_1^{(1)} | \overline{Z}_1) = n_1\varepsilon\,. \end{aligned} \tag{24}$$

where (a) follows from the usual wiretap coding and (b) follows from the fact that $\overline{X}_2^{(2)} \perp (\overline{M}_1^{(1)}, \overline{Z}_1)$, and $(\overline{X}_2^{(2)}, \overline{Z}_2) \perp (\overline{M}_1^{(1)}, \overline{Z}_1)$.
Next, we consider:

$$\begin{aligned} &I(\overline{M}_2^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \\ &= I(\overline{M}_{2,1}^{(1)}, \overline{M}_{2,2}^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \\ &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) + I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}) \\ &\triangleq I_1 + I_2\,. \end{aligned} \tag{25}$$

We get the upper bounds on I_1 and I_2. The first term:

$$\begin{aligned} I_1 &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \\ &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_1, \overline{Z}_{2,1}, \overline{Z}_{2,2} | \overline{X}_2^{(2)}) \\ &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_1 | \overline{X}_2^{(2)}) + I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,1} | \overline{X}_2^{(2)}, \overline{Z}_1) \\ &+ I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,2} | \overline{X}_2^{(2)}, \overline{Z}_1, \overline{Z}_{2,1}) \\ &\overset{(a)}{=} 0 + I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,1} | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1) \\ &+ I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,2} | \overline{X}_2^{(2)}, \overline{Z}_1, \overline{Z}_{2,1}) \\ &\triangleq I_{11} + I_{12}\,, \end{aligned} \tag{26}$$

where (a) follows because $\overline{Z}_1$ is independent of $(\overline{M}_{21}^{(1)}, \overline{X}_2^{(2)})$. Furthermore:

$$\begin{aligned} I_{11} &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,1} | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1) \\ &= H(\overline{M}_{2,1}^{(1)}; | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1) \\ &- H(\overline{M}_{2,1}^{(1)}; | \overline{Z}_{2,1}, \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1) \\ &\overset{(a)}{=} H(\overline{M}_{2,1}^{(1)}; | \overline{X}_{2,1}^{(2)}) - H(\overline{M}_{2,1}^{(1)}; | \overline{Z}_{2,1}, \overline{X}_{2,1}^{(2)}) \\ &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,1}, | \overline{X}_{2,1}^{(2)}) \overset{(b)}{\leq} n_1\varepsilon\,, \end{aligned} \tag{27}$$

where (a) follows, since $(\overline{X}_{2,2}^{(2)}, \overline{Z}_1) \perp (\overline{M}_{2,1}^{(1)}, \overline{Z}_{2,1}, \overline{X}_{2,1}^{(2)})$ and (b) follows because the first part of the message is encoded via the usual wiretap coding scheme for the multiple access wiretap channel. Also:

$$\begin{aligned} I_{12} &= I(\overline{M}_{2,1}^{(1)}; \overline{Z}_{2,2} | \overline{X}_2^{(2)}, \overline{Z}_1, \overline{Z}_{2,1}) \\ &= H(\overline{M}_{2,1}^{(1)}; | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1, \overline{Z}_{2,1}) \\ &- H(\overline{M}_{2,1}^{(1)} | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_1, \overline{Z}_{2,1}, \overline{Z}_{2,2}) \\ &\overset{(a)}{=} H(\overline{M}_{2,1}^{(1)}; | \overline{X}_{2,1}^{(2)}, \overline{Z}_{2,1}) - H(\overline{M}_{2,1}^{(1)}; | \overline{X}_{2,1}^{(2)}, \overline{Z}_{2,1}) = 0\,, \end{aligned}$$

where (a) follows, since $(\overline{X}_{2,2}^{(2)}, \overline{Z}_1, \overline{Z}_{2,2}) \perp (\overline{M}_{2,1}^{(1)}, \overline{X}_{2,1}^{(2)}, \overline{Z}_{2,1})$. From Eqs. (25), (26) and (27), we have $I_1 = I_{11} + I_{12} \leq n_1\varepsilon$. Next, we consider:

$$\begin{aligned} I_2 &= I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}) \\ &= I(\overline{M}_{2,2}^{(1)}; \overline{Z}_2 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}) \\ &+ I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}, \overline{Z}_2)\,. \end{aligned} \tag{28}$$

We have:

$$\begin{aligned} &I(\overline{M}_{2,2}^{(1)}; \overline{Z}_2 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}) \\ &= I(\overline{M}_{2,2}^{(1)}; \overline{Z}_{2,1} | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}) \\ &+ I(\overline{M}_{2,2}^{(1)}; \overline{Z}_{2,2} | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}, \overline{Z}_{2,1}) \\ &\overset{(a_1)}{=} 0 + I(\overline{M}_{2,2}^{(1)}; \overline{Z}_{2,2} | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}, \overline{Z}_{2,1}) \\ &\overset{(a_2)}{=} I(\overline{M}_{2,2}^{(1)}; \overline{Z}_{2,2} | \overline{X}_{2,2}^{(2)}) \overset{(a_3)}{=} 0, \end{aligned}$$

and (a_1) follows, since $\overline{M}_{2,2}^{(1)} \perp (\overline{Z}_{2,1}, \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)})$; (a_2) holds because $(\overline{X}_{2,1}^{(2)}, \overline{M}_{2,1}^{(1)}) \perp (\overline{M}_{2,2}^{(1)}, \overline{Z}_{2,2}, \overline{X}_{2,2}^{(2)})$; and (a_3) is true, since $\overline{M}_{2,2}^{(1)} \perp (\overline{X}_{2,2}^{(2)}, \overline{Z}_{2,2})$.
In addition:

$$\begin{aligned}
& I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1 | \overline{X}_2^{(2)}, \overline{M}_{2,1}^{(1)}, \overline{Z}_2) \\
& = I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1 | \overline{X}_{2,1}^{(2)}, \overline{X}_{2,2}^{(2)}, \overline{M}_{2,1}^{(1)}, \overline{Z}_{2,1}, \overline{Z}_{2,2}) \\
& \overset{(b_1)}{=} I(\overline{M}_{2,2}^{(1)}; \overline{Z}_1 | \overline{X}_{2,2}^{(2)}, \overline{Z}_{2,2}) \overset{(b_2)}{=} 0 \,,
\end{aligned}$$

where (b_1) follows, since $(\overline{M}_{2,1}^{(1)}, \overline{Z}_{2,1}, \overline{X}_{2,1}^{(2)}) \perp (\overline{Z}_{2,2}, \overline{X}_{2,2}^{(2)}, \overline{M}_{2,2}^{(1)}, \overline{Z}_1)$ and (b_2) follows because $\overline{Z}_1 \perp (\overline{M}_{2,2}^{(1)}, \overline{X}_{2,2}^{(2)}, \overline{Z}_{2,2})$. Hence, from (28) we have $I_2 = 0$.
From (25), we have:

$$I(\overline{M}_2^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) \le n_1 \varepsilon \,. \tag{29}$$

Similarly, one can show that:

$$I(\overline{M}_2^{(2)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(1)}) \le n_1 \varepsilon \,. \tag{30}$$

Therefore, from (8):

$$\begin{aligned}
& I(\overline{M}_2^{(1)}, \overline{M}_2^{(2)}; \overline{Z}_1, \overline{Z}_2) \\
& \le I(\overline{M}_2^{(1)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(2)}) + I(\overline{M}_2^{(2)}; \overline{Z}_1, \overline{Z}_2 | \overline{X}_2^{(1)}) \,.
\end{aligned}$$

To prove that (13) holds for any slot, we use the principle of mathematical induction in the lemma below. For a proof, please see [23].

Lemma 3.2. Let (13) hold for k, then it also holds for $k+1$. □

Remark 1 (extension to strong secrecy notion). We have used the notion of *weak secrecy* in the above proof, i.e. if message W is transmitted via wiretap coding and Eve receives sequence Z^n, then $I(W; Z^n) \le n_1 \varepsilon$. The criteria of *strong secrecy* provide not only for the information leakage rate, but also require that the absolute information vanishes, i.e. $I(W; Z^n) \le \varepsilon$. In a single-user point-to-point wiretap channel, if the weak secrecy notion is replaced by the strong secrecy notion, the secrecy capacity of the channel does not change [26]. A similar result has been proved for a MAC-WT in [16], using the channel resolvability-based coding scheme. If we use, in the coding scheme proposed in this paper (Theorem 2), a coding scheme based on the resolvability technique in slot 1, and in other slots use both coding schemes together (i.e. resolvability-based coding in the first part of the slot) and the previously transmitted message (which is now secure in the strong sense with respect to Eve) as a secret key in the second part of the slot, we can achieve the same secrecy-rate region, i.e. the capacity region of the usual multiple access channel without Eve, satisfying the following leakage rate:

$$\limsup_{n \to \infty} I(\overline{M}_k^{(1)}, \overline{M}_k^{(2)}; \overline{Z}_1, \overline{Z}_2, \ldots, \overline{Z}_k) = 0 \,, \tag{31}$$

as $n \to \infty$, because in the RHS of (13), we can get ε instead of $2n_1\varepsilon$.

4. Discrete Memoryless MAC-WT with Buffer

In this section we improve the result from Theorem 3.1 by obtaining rates (16) while enhancing the secrecy requirement from (13) to:

$$\begin{aligned}
I(\overline{M}_k^{(1)}, \overline{M}_{k-1}^{(1)}, \ldots, \overline{M}_{k-N_1}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k | \overline{X}_k^{(2)}) &\le n_1 \varepsilon \,, \\
I(\overline{M}_k^{(2)}, \overline{M}_{k-1}^{(2)}, \ldots, \overline{M}_{k-N_1}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k | \overline{X}_k^{(1)}) &\le n_1 \varepsilon \,, \\
I(\overline{M}_k^{(1)}, \overline{M}_k^{(2)}, \ldots, \overline{M}_{k-N_1}^{(1)}, \overline{M}_{k-N_1}^{(2)}; \overline{Z}_1, \ldots, \overline{Z}_k) &\le 2n_1 \varepsilon \,,
\end{aligned} \tag{32}$$

where N_1 can be arbitrarily large. This will satisfy the requirements of any practical system. Therefore, we use a key buffer at each of the users and instead of using the messages transmitted in slot $k-1$ as the key in slot k, we use the messages transmitted in slots before $k - N_1 - 1$.
Let each user have an infinite key buffer to store the key bits. The message $\overline{M}_k^{(i)}$ after transmission in slot k from user i is stored in its key buffer at the end of the slot. However, now in slot $k+1$ we use the *oldest* bits stored in its key buffer as a key in the second part of its slot. Once certain bits from the key buffer have been used as a key, these are discarded from the key buffer.
Let $B_k^{(i)}$ be the number of key bits in the key buffer of the i-th user at the beginning of the k-th slot. Then, out of this, for $k \ge \lambda^*$, the number of key bits used in a slot by user 1 is $C_1 n_2$, since these are used only in the second part of the slot where $C_1 \le I(X_1; Y | X_2)$, while the total number of secret bits transmitted in the slot is $C_1 n_2 + R_s^{(1)} n_1$. These transmitted bits also get stored in its key buffer at time $k+1$. Similarly, the same holds for user 2. Thus, $B_k^{(i)} \to \infty$ as $k \to \infty$ for $i = 1, 2$.
After some time (say N_2 slots) has elapsed since using the oldest bits in the key buffer, for $k \ge N_2$, we will be using the secret key bits only from messages $(\overline{M}_1^{(i)}, \overline{M}_2^{(i)}, \ldots, \overline{M}_{k-N_1-1}^{(i)})$ for securing messages $(\overline{M}_k^{(i)}, \overline{M}_{k-1}^{(i)}, \ldots, \overline{M}_{k-N_1}^{(i)})$, for user $i = 1, 2$, respectively.
The following proof works for $N_1 > 0$. Theorem 2.1 for $N_1 = 0$.

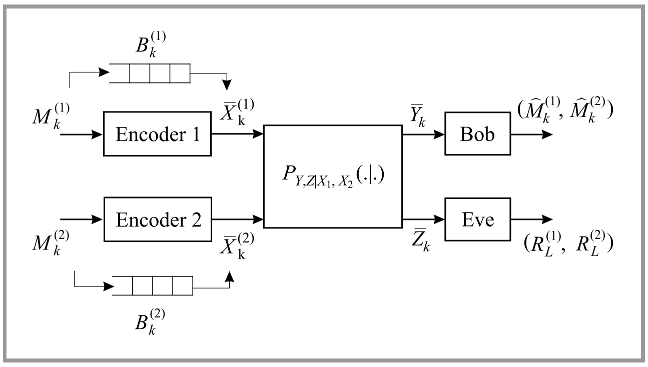


Fig. 4. Discrete memoryless multiple access wiretap channel with secret key buffers.

Theorem 4.1. The secrecy-rate region with the leakage rate constraints (32) of a DM-MAC-WT is equal to the usual Ahleswede-Liao region (16) of MAC.

Proof. With the proposed modification of the coding-decoding scheme presented in Section 3, in any slot k, the legitimate receiver is able to decode the message pair $(\overline{M}_k^{(1)}, \overline{M}_k^{(2)})$ with the error probability of $P_e^{(n)} \to 0$ as $n \to \infty$. Also (13) along with $R_{L,i}^{(n)} \leq n_1\varepsilon_1, i = 1,2$ continue to be satisfied, where $\varepsilon_1 > 0$ will be fixed later on.

Now we consider the leakage rate. We have:

$$\begin{aligned}
&I(\overline{M}_k^{(1)}, \overline{M}_{k-1}^{(1)}, \ldots, \overline{M}_{k-N_1}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k|\overline{X}_k^{(2)}) \\
&\quad = I(\overline{M}_{k,1}^{(1)}, \overline{M}_{k-1,1}^{(1)}, \ldots, \overline{M}_{k-N_1,1}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k|\overline{X}_k^{(2)}) \\
&\quad + I(\overline{M}_{k,2}^{(1)}, \overline{M}_{k-1,2}^{(1)}, \ldots, \overline{M}_{k-N_1,2}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k \\
&\qquad |\overline{X}_k^{(2)}, \overline{M}_{k,1}^{(1)}, \ldots, \overline{M}_{k-N_1,1}^{(1)}) .
\end{aligned}$$

From Lemma 7.1 and Lemma 7.2 in the Appendix:

$$I(\overline{M}_{k,1}^{(1)}, \overline{M}_{k-1,1}^{(1)}, \ldots, \overline{M}_{k-N_1,1}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k|\overline{X}_k^{(2)}) \leq n_1\varepsilon \quad (33)$$

and

$$I\Big(\overline{M}_{k,2}^{(1)}, \overline{M}_{k-1,2}^{(1)}, \ldots, \overline{M}_{k-N_1,2}^{(1)}; \overline{Z}_1, \ldots, \overline{Z}_k|\overline{X}_k^{(2)}, \overline{M}_{k,1}^{(1)}, \ldots, \overline{M}_{k-N_1,1}^{(1)}\Big) \leq 6n_1\varepsilon \, . \quad (34)$$

Thus, taking $\varepsilon = \varepsilon/7$, we obtain the first inequality in (32). Similarly, we can show the second inequality.

To prove the third inequality, we define $\widetilde{M}^{(1)} \triangleq (\overline{M}_k^{(1)}, \overline{M}_{k-1}^{(1)}, \ldots, \overline{M}_{k-N_1}^{(1)})$, $\widetilde{M}^{(2)} \triangleq (\overline{M}_k^{(2)}, \overline{M}_{k-1}^{(2)}, \ldots, \overline{M}_{k-N_1}^{(2)})$ and $\widetilde{Z} \triangleq (\overline{Z}_1, \ldots, \overline{Z}_k)$, and we have:

$$\begin{aligned}
&I(\widetilde{M}^{(1)}, \widetilde{M}^{(2)}; \widetilde{Z}) \\
&\quad = I(\widetilde{M}^{(1)}; \widetilde{Z}) + I(\widetilde{M}^{(2)}; \widetilde{Z}|\widetilde{M}^{(1)}) \\
&\quad = H(\widetilde{M}^{(1)}) - H(\widetilde{M}^{(1)}|\widetilde{Z}) + H(\widetilde{M}^{(2)}) - H(\widetilde{M}^{(2)}|\widetilde{Z}, \widetilde{M}^{(1)}) \\
&\quad \overset{(a)}{\leq} H(\widetilde{M}^{(1)}|\overline{X}_k^{(2)}) - H(\widetilde{M}^{(1)}|\widetilde{Z}, \overline{X}_k^{(2)}) + H(\widetilde{M}^{(2)}|\overline{X}_k^{(1)}) \\
&\qquad - H(\widetilde{M}^{(2)}|\widetilde{Z}, \overline{X}_k^{(1)}) \\
&\quad = I(\widetilde{M}^{(1)}; \widetilde{Z}|\overline{X}_k^{(2)}) + I(\widetilde{M}^{(2)}; \widetilde{Z}|\overline{X}_k^{(1)}), \qquad (35)
\end{aligned}$$

where (a) follows because: conditioning decreases the entropy, all transmitted messages are independent of each other and the codeword is a one-to-one function of the message to be transmitted. Hence, from (33) and (34):

$$I\Big(\overline{M}_k^{(1)}, \overline{M}_k^{(2)}, \ldots, \overline{M}_{k-N_1}^{(1)}, \overline{M}_{k-N_1}^{(2)}; \overline{Z}_1, \ldots, Z_k^n\Big) \leq n_1\varepsilon \, . \quad (36)$$

5. Fading MAC-WT

In this section, we consider a two-user discrete time additive white Gaussian fading channel. If X_1, X_2 are the channel inputs, then Bob receives:

$$Y = \widetilde{H}_1 X_2 + \widetilde{H}_2 X_2 + N_1 \quad (37)$$

and Eve receives:

$$Z = \widetilde{G}_1 X_1 + \widetilde{G}_2 X_2 + N_2, \quad (38)$$

where $\widetilde{H}_i$ is the channel gain to Bob, $\widetilde{G}_i$ is the channel gain to Eve, and N_i has Gaussian distribution with a mean 0 and variance σ_i^2, $i = 1,2$. We assume that the random variables $\widetilde{H}_1, \widetilde{H}_2, \widetilde{G}_1, \widetilde{G}_2, N_1, N_2$ are independent of each other. The channel is experiencing slow fading, i.e. the channel gains remain the same during the transmission of the entire codeword. Let $H_i = |\widetilde{H}_i|^2$ and $G_i = |\widetilde{G}_i|^2$, $i = 1,2$. The average power constraint for user i is $\overline{P}_i$.

We define some notation for convenience. For $H = (H_1, H_2)$, $G = (G_1, G_2)$:

$$\begin{aligned}
C_1(P_1(H,G)) &\triangleq \frac{1}{2}\log\left(1 + \frac{H_1 P_1(H,G)}{\sigma_1^2}\right), \\
C_2(P_2(H,G)) &\triangleq \frac{1}{2}\log\left(1 + \frac{H_2 P_1(H,G)}{\sigma_1^2}\right), \\
C_1^e(P_1(H,G)) &\triangleq \frac{1}{2}\log\left(1 + \frac{G_1 P_1(H,G)}{\sigma_2^2 + G_2 P_2(H,G)}\right), \\
C_2^e(P_2(H,G)) &\triangleq \frac{1}{2}\log\left(1 + \frac{G_2 P_2(H,G)}{\sigma_2^2 + G_1 P_1(H,G)}\right), \\
C(P_1(H,G), P_2(H,G)) &\triangleq \frac{1}{2}\log\left(1 + \frac{H_1 P_1(H,G) + H_2 P_2(H,G)}{\sigma_1^2}\right). \quad (39)
\end{aligned}$$

The achievable secrecy rate region for this channel is:

$$\mathscr{R}_g^s(\overline{P}) = \left\{ \begin{array}{c} (R_1^{(s)}, R_2^{(s)}) : \\ R_1^{(s)} \leq \mathbb{E}_{H,G}\left[(C_1(P_1) - C_1^e(P_1))^+\right] \\ R_2^{(s)} \leq \mathbb{E}_{H,G}\left[(C_2(P_2) - C_2^e(P_2))^+\right] \\ R_1^{(s)} + R_2^{(s)} \leq \mathbb{E}_{H,G}\left[\left(C(P_1, P_2) - \sum_{i=1}^2 C_i^e(P_i)\right)^+\right] \end{array} \right\} \quad (40)$$

where $\overline{P} = (\overline{P}_1, \overline{P}_2)$. To achieve these rates (with $P_i(H,G) \equiv \overline{P}_i$), the transmitters need not know the channel states, but Bob's receiver needs to know all H_i, G_i. We assume this in this section.

If the channel states (H,G) are known to each of the users, as well as at the receiver of Bob, then we can improve over the rate region in (40) by making the transmit power as functions of (H,G):

$$\mathscr{P} : H \times G \to \mathbb{R}_+^2 \, , \quad (41)$$

where $\mathscr{P} = (P_1, P_2)$. Now we note the rate region as $\mathscr{C}_f^s(\mathscr{P})$. Therefore, the secrecy capacity region of MAC-WT ($\mathscr{C}_f^s(\mathscr{P})$) is not known, but $\mathscr{R}_f^s(\mathscr{P}) \subseteq \mathscr{C}_f^s(\mathscr{P})$ [27].

Now, we apply the coding scheme of Section 3 to the two-user fading MAC-WT in order to enlarge the secrecy rate region to the usual capacity region of the fading channel. The message pair $(\overline{M}_k^{(1)}, \overline{M}_k^{(2)})$ is to be transmitted confidentially by the two users over the fading MAC in slot k, and will be stored in their respective secret key buffers at the end of the k-th slot. Let $B_k^{(1)}, B_k^{(2)}$ be the number of bits in the key buffers of users 1 and 2, respectively, at the beginning of the slot k.

Let $\overline{R}_k^{(i)}$ bits be taken from the key buffer of user i to act as a secret key for the transmission of message $\overline{M}_k^{(i)}$. The two users satisfy the long-term average power constraint:

$$\limsup_{k\to\infty} \frac{1}{k}\sum_{m=1}^{k} \mathsf{E}\left[P_i(H_k, G_k)\right] \leq \overline{P}_i, \quad i = 1, 2, \tag{42}$$

where $H_k,\ G_k$ are the channel gains in slot k and $P_i(H_k,\ G_k)$ is the average power used by user i in slot k. We need to compute $P_i(H,\ G)$ and $\overline{R}_k^{(i)},\ i = 1, 2$, such that the resulting average rate region $(\overline{r}^{(1)}, \overline{r}^{(2)})$ is maximized, where:

$$\overline{r}^{(i)} = \limsup_{k\to\infty} \frac{1}{k}\sum_{l=1}^{k} r_l^{(i)}, \tag{43}$$

$r_k^{(i)}$ is the transmission rate of user i in slot k, subject to the long-term respective power constraints (42). The secrecy-rate region is computed when:

$$\Pr\left(\{H_{1k} > G_{1k}\} \cup \{H_{2k} > G_{2k}\}\right) > 0\ , \tag{44}$$

where $\Pr(A)$ represents the probability of event A. Otherwise, the secrecy-rate region is zero. Actually, we state the following theorem for $\Pr(H_{ik} > G_{ik}) > 0,\ i = 1, 2$. If it is not true for any one i, then the secrecy rate for that user is zero. For both transmitting users, at the end of slot k, $\hat{r}_k^{(i)} = n(l+1)r_k^{(i)}$ bits are stored in the secret key buffer for future use as a key, where $n_2 = ln_1$. Hence, $B_k^{(i)}$ evolves as:

$$B_{k+1}^{(i)} = B_k^{(i)} + \hat{r}_k^{(i)} - \overline{R}_k^{(i)}\ , \tag{45}$$

where $\hat{r}_k^{(i)} \geq \overline{R}_k^{(i)}$ and $\hat{r}_k^{(i)} > \overline{R}_k^{(i)}$ with positive probability $Pr(H_{ik} > G_{ik})$. Therefore, $B_k^{(i)} \to \infty$ a.s. for $i = 1, 2$.

Theorem 5.1. If $\Pr(H_{ik} > G_{ik}) > 0,\ i = 1, 2$, and all the channel gains are available at all the transmitters, then the following long-term average rates that maintain the leakage rates (32), are achievable:

$$\begin{aligned} R_1^{(s)} &\leq \frac{1}{2}\mathsf{E}_{H,G}\left[C_1\left(P_1(H)\right)\right]\ , \\ R_2^{(s)} &\leq \frac{1}{2}\mathsf{E}_{H,G}\left[C_2\left(P_2(H)\right)\right]\ , \\ R_1^{(s)} + R_2^{(s)} &\leq \frac{1}{2}\mathsf{E}_{H,G}\left[C\left(P_1(H), P_2(H)\right)\right]\ . \end{aligned} \tag{46}$$

where P is any policy that satisfies the average power constraint. If Bob is the only party knowing all channel states (not the transmitters), then $(R_1^{(s)}, R_2^{(s)})$ satisfies (46) with $P_i(H,G) \equiv \overline{P}_i,\ i = 1, 2$.

Achievability scheme outline. We use the coding-decoding scheme proposed in Section 3 with appropriate changes to account for the fading process. Assuming that $B_0^{(i)} = 0,\ i = 1, 2$, user i transmits the first time when $H_{ik} > G_{ik}$. Then, it uses the usual MAC wiretap coding as proposed in [6] in all its $l+1$ mini-slots.

In the next slot (say k-th), user i uses the first mini-slot for wiretap coding (if $H_{ik} > G_{ik}$ for user i) and the rest of the m mini-slots for transmission via the secret key (if $H_{ik} < G_{ik}$ the first mini-slot is not used). It uses $\overline{R}_k^{(i)} = \min\left[B_k^{(i)}, lC_i(P_i(H,G)n_1)\right]$ key bits which are removed from the key buffer at the end of the slot. The total number of bits transmitted by user i in slot k is:

$$\hat{r}_k^{(i)} = \overline{R}_k^{(i)} + n_1\left[C_i(P_1(H_k, G_k)\right] - C_i^e\left[P_i(H_k, G_k)\right]^+\ . \tag{47}$$

These bits are stored in the key buffer at the end of the slot. Thus, $\hat{r}_k^{(i)} \geq \overline{R}_k^{(i)}$ and since $\Pr(H_{ik} > G_{ik}) > 0,\ i = 1, 2$, $\Pr(\hat{r}_k^{(i)} > \overline{R}_k^{(i)}) > 0$. Finally, $B_k^{(i)} \to \infty$ a.s. for $i = 1, 2$.

Also, as before, we can show that after some slot $k \geq N_2$, with an arbitrarily large probability, the messages transmitted in slots $k,\ k-1, \ldots,\ k-N_1$ will use the messages transmitted before $k - N_1 - 1$, and the rate used in the first mini-slot will satisfy (40), but the rate used in the second mini-slot will satisfy (46). The overall rate of the slot can be made as close to (46) as we wish, by taking a large value of l. Thus, the rest of the proof demonstrating $P_e^n \to 0$ and that (32) is satisfied follows from Theorem 3.1.

All the above results extend in *strong* secrecy sense, as in Section 3, by using the *resolvability*-based coding scheme of [16] instead of the usual wiretap coding for MAC-WT of [6].

6. Conclusions

In this paper, we obtain the secrecy-rate region for a time-slotted MAC-WT. By using the previously transmitted message as a secret key in the next slot, we show that we can mitigate the rate loss and achieve the secrecy-rate region equal to the Ahleswede-Liao region of a multiple access channel (without wiretapper), if we consider the secrecy rate of the individual messages. We then extend the results to a scenario in which an arbitrarily large number of recently transmitted multiple messages is now secure with respect to the information of Eve, by using the secret key buffer for both transmitters. Finally, we further extend our coding scheme to a fading Gaussian channel and show that the usual Ahleswede-Liao region can be obtained while retaining the secrecy of the multiple messages.

Appendix

DM-MAC-WT with Secret Key Buffer

Lemma 7.1. The following inequality is satisfied

$$I(\overline{M}^{(1)}_{k,1},\overline{M}^{(1)}_{k-1,1},\dots,\overline{M}^{(1)}_{k-N_1,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k)\le n_1\varepsilon. \quad (48)$$

Proof. We have:

$$\begin{aligned}
&I(\overline{M}^{(1)}_{k,1},\overline{M}^{(1)}_{k-1,1},\dots,\overline{M}^{(1)}_{k-N_1,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k)\\
&=I(\overline{M}^{(1)}_{k,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k)\\
&+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})\\
&+\dots+I(\overline{M}^{(1)}_{k-N_1,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{M}^{(1)}_{k-1,1},\dots,\\
&\overline{M}^{(1)}_{k-N_1+1,1})\triangleq I_1+I_2+\dots+I_{N_1}\ . \qquad (49)
\end{aligned}$$

Now let us evaluate each term. Denoting the two parts of $\overline{Z}_k$ by $\overline{Z}_{k,1},\overline{Z}_{k,2}$, and choosing the wiretap coding with leakage rate $\le n_1\varepsilon_1$, where $\varepsilon_1=\varepsilon/N_1$:

$$\begin{aligned}
I_1&=I(\overline{M}^{(1)}_{k,1};\overline{Z}_{1,1},\overline{Z}_{1,2},\dots,\overline{Z}_{k,1},\overline{Z}_{k,2}|\overline{X}^{(2)}_k)\\
&=I(\overline{M}^{(1)}_{k,1};\overline{Z}_{k,1}|\overline{X}^{(2)}_k)+I(\overline{M}^{(1)}_{k,1};\overline{Z}_1,\dots,\overline{Z}_{k-1},\overline{Z}_{k,2}|\overline{X}^{(2)}_k)\\
&\overset{(a)}{\le} n_1\varepsilon_1+I(\overline{M}^{(1)}_{k,1};\overline{Z}_1,\dots,\overline{Z}_{k-1},\overline{Z}_{k,2}|\overline{X}^{(2)}_k,)\\
&=n_1\varepsilon_1+H(\overline{M}^{(1)}_{k,1}|\overline{X}^{(2)}_k)-H(\overline{M}^{(1)}_{k,1}|\overline{X}^{(2)}_k,\overline{Z}_1,\dots,\overline{Z}_{k-1},\overline{Z}_{k,2})\\
&\overset{(b)}{=} n_1\varepsilon_1+H(\overline{M}^{(1)}_{k,1}|\overline{X}^{(2)}_k)-H(\overline{M}^{(1)}_{k,1}|\overline{X}^{(2)}_k)=n_1\varepsilon_1\ , \qquad (50)
\end{aligned}$$

where (a) follows from wiretap coding and (b) follows, since $(\overline{Z}_1,\dots,\ \overline{Z}_{k-1},\ \overline{Z}_{k,2})\perp(W^{(1)}_{k,1},\overline{X}^{(2)}_k)$.
Next consider I_2. We have:

$$\begin{aligned}
I_2&=I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_1,\dots,\overline{Z}_{k-1,1},\overline{Z}_{k-1,2},\overline{Z}_k|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})\\
&=I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})+I(\overline{M}^{(1)}_{k-1,1};\\
&\quad(\overline{Z}_1,\dots,\overline{Z}_k)\backslash\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&=H(\overline{M}^{(1)}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})-H(\overline{M}^{(1)}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1})\\
&+I(\overline{M}^{(1)}_{k-1,1};(\overline{Z}_1,\dots,\overline{Z}_k)\backslash\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1}) \quad (51)\\
&\overset{(a)}{=} H(\overline{M}^{(1)}_{k-1,1})-H(\overline{M}^{(1)}_{k-1,1}|\overline{Z}_{k-1,1})\\
&+I(\overline{M}^{(1)}_{k-1,1};(\overline{Z}_1,\dots,\overline{Z}_k)\backslash\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&=I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k-1,1})I(\overline{M}^{(1)}_{k-1,1};(\overline{Z}_1,\dots,\overline{Z}_k)\backslash\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\\
&\quad\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&\overset{(b)}{\le} n_1\varepsilon_1+I(\overline{M}^{(1)}_{k-1,1};(\overline{Z}_1,\dots,\overline{Z}_k)\backslash\overline{Z}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&=n_1\varepsilon_1+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_1,\dots,\overline{Z}_{k-1,2},\overline{Z}_k|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&=n_1\varepsilon_1+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_1,\dots,\overline{Z}_{k-2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1})\\
&+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_k,\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1}\overline{Z}_1,\dots,\overline{Z}_{k-2})\\
&\overset{(c)}{=} n_1\varepsilon_1+0+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_k,\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1}\overline{Z}_1,\dots,\\
&\quad\overline{Z}_{k-2})\\
&=n_1\varepsilon_1+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1}\overline{Z}_1,\dots,\overline{Z}_{k-2})\\
&+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k,2},\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\\
&\quad\overline{Z}_{k,1})\\
&=n_1\varepsilon_1+H(\overline{M}^{(1)}_{k-1,1};|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1}\overline{Z}_1,\dots,\overline{Z}_{k-2})\\
&-H(\overline{M}^{(1)}_{k-1,1};|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1}\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1})\\
&+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k,2},\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\\
&\quad\overline{Z}_{k,1})\\
&\overset{(d)}{=} n_1\varepsilon_1+H(\overline{M}^{(1)}_{k-1,1};|\overline{Z}_{k-1,1})-H(\overline{M}^{(1)}_{k-1,1};|\overline{Z}_{k-1,1})\\
&+I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k,2},\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\overline{Z}_1,\dots,\\
&\quad\overline{Z}_{k-2},\overline{Z}_{k,1})\ , \qquad (52)
\end{aligned}$$

where (a) follows since $\overline{M}^{(1)}_{k-1,1}\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})$ and $(\overline{M}^{(1)}_{k-1,1},\overline{Z}_{k-1})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})$, (b) follows from wiretap coding, (c) follows since $(\overline{M}^{(1)}_{k-1,1},\overline{Z}_{k-1})\perp(\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})$, $(\overline{Z}_1,\dots,\overline{Z}_{k-2})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})$ and $(\overline{Z}_1,\dots,\overline{Z}_{k-1})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1})$ and (d) follows since $(\overline{M}^{(1)}_{k-1,1},\overline{Z}_{k-1,1})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_1,\dots,\overline{Z}_{k-2})$.
But:

$$\begin{aligned}
&I(\overline{M}^{(1)}_{k-1,1};\overline{Z}_{k,2},\overline{Z}_{k-1,2}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\\
&\quad\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1})\\
&=H(\overline{M}^{(1)}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1})\\
&-H(\overline{M}^{(1)}_{k-1,1}|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_{k-1,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1},\\
&\quad\overline{Z}_{k,2},\overline{Z}_{k-1,2})\\
&\overset{(a)}{=} H(\overline{M}^{(1)}_{k-1,1}|\overline{Z}_{k-1,1})-H(\overline{M}^{(1)}_{k-1,1}|\overline{Z}_{k-1,1})\\
&=0\ , \qquad (53)
\end{aligned}$$

where (a) follows, since $(\overline{M}^{(1)}_{k-1,1},\overline{Z}_{k-1,1})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1})$ and $(\overline{M}^{(1)}_{k-1,1},\overline{Z}_{k-1,1})\perp(\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\overline{Z}_1,\dots,\overline{Z}_{k-2},\overline{Z}_{k,1},\overline{Z}_{k,2},\overline{Z}_{k-1,2})$.
Hence we have:

$$I_2\le n_1\varepsilon_1\ . \qquad (54)$$

One can similarly prove that $I_i\le n_1\varepsilon_1$ for $i=3,4,\ \dots,\ N_1$. Therefore:

$$\begin{aligned}
&I(\overline{M}^{(1)}_{k,1},\overline{M}^{(1)}_{k-1,1},\dots,\overline{M}^{(1)}_{k-N_1,1};\overline{Z}_1,\dots,\overline{Z}_k|\overline{X}^{(2)}_k)\\
&\le N_1n\varepsilon_1=n_1\varepsilon\ . \qquad (55)
\end{aligned}$$

□

Lemma 7.2. The following inequality is satisfied

$$\begin{aligned}
&I(\overline{M}^{(1)}_{k,2},\overline{M}^{(1)}_{k-1,2},\dots,\overline{M}^{(1)}_{k-N_1,2};\overline{Z}_1,\dots,\overline{Z}_k\\
&|\overline{X}^{(2)}_k,\overline{M}^{(1)}_{k,1},\dots,\overline{M}^{(1)}_{k-N_1,1})\le 6n_1\varepsilon\ . \qquad (56)
\end{aligned}$$

Proof.

$$\begin{aligned}
&I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_1, \ldots, \overline{Z}_k | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1})\\
&= I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1} | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1})\\
&+ I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1 2}; \overline{Z}_{k-N_1}, \ldots, \overline{Z}_k | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1})\\
&\overset{(a)}{=} 0 + I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_{k-N_1}, \ldots, \overline{Z}_k | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1})\\
&= I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_{k-N_1,1}, \ldots, \overline{Z}_{k,1} | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1})\\
&+ I(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,1}; \overline{Z}_{k-N_1,2}, \ldots, \overline{Z}_{k,2} | \overline{X}^{(2)}_k,\\
&\quad \overline{M}^{(1)}_{k,1}, \overline{M}^{(1)}_{k-1,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1},\\
&\quad \overline{Z}_{k,1}, \overline{Z}_{k-1,1}, \ldots, \overline{Z}_{k-N_1,1})\\
&\overset{(b)}{=} 0 + I(\overline{M}^{(1)}_{k,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_{k-N_1,2}, \ldots, \overline{Z}_{k,2} | \overline{M}^{(1)}_{k,1}, \ldots,\\
&\quad \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_1, \ldots, \overline{Z}_{k-N_1}, \overline{Z}_{k-N_1,1}, \ldots, \overline{Z}_{k-1}, \overline{X}^{(2)}_k)\\
&\overset{(c)}{=} I(\overline{M}^{(1)}_{k,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}; \overline{Z}_{k-N_1,2}, \ldots, \overline{Z}_{k,2} | \overline{Z}_1, \ldots,\\
&\quad \overline{Z}_{k-N_1}, \overline{X}^{(2)}_k)\\
&\triangleq I(\hat{M}^{(1)}_2; \hat{Z}_2 | \hat{Z}_1, \overline{X}^{(2)}_k),
\end{aligned}$$

where (a) follows, since $(\overline{M}^{(1)}_{k,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2}) \perp (\overline{Z}_1, \ldots, \overline{Z}_{k-N_1-1}, \overline{M}^{(1)}_{k,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{X}^{(2)}_k)$, (b) follows, since $(\overline{M}^{(1)}_{k,2}, \overline{M}^{(1)}_{k-1,2}, \ldots, \overline{M}^{(1)}_{k-N_1,2})$ is independent of the other random variables (RVs) in the first expression, (c) follows, since $(\overline{M}^{(1)}_{k,1}, \ldots, \overline{M}^{(1)}_{k-N_1,1}, \overline{Z}_{k-N_1,1}, \ldots, \overline{Z}_{k-1,1})$ is independent of all other RVs in the expression, and in the last inequality we denote the respective random sequences with their respective widehat symbols.

Now we observe that:

$$\begin{aligned}
&I(\hat{M}^{(1)}_2; \hat{Z}_1, \hat{Z}_2 | \overline{X}^{(2)}_k)\\
&\quad = I(\hat{M}^{(1)}_2; \hat{Z}_1 | \overline{X}^{(2)}_k) + I(\hat{M}^{(1)}_2; \hat{Z}_2 | \hat{Z}_1, \overline{X}^{(2)}_k)\\
&\quad \overset{(a)}{=} 0 + I(\hat{M}^{(1)}_2; \hat{Z}_2 | \hat{Z}_1, \overline{X}^{(2)}_k)\\
&\quad \le I(\hat{M}^{(1)}_2; \hat{Z}_1, \hat{Z}_2 | \overline{X}^{(2)}_k)\\
&\quad = I(\hat{M}^{(1)}_2; \hat{Z}_2 | \overline{X}^{(2)}_k) + I(\hat{M}^{(1)}_2; \hat{Z}_1 | \hat{Z}_2, \overline{X}^{(2)}_k)\\
&\quad \overset{(b)}{=} 0 + I(\hat{M}^{(1)}_2; \hat{Z}_1 | \hat{Z}_2, \overline{X}^{(2)}_k)\ ,
\end{aligned} \tag{57}$$

where (a) follows, since $\hat{M}^{(1)}_2 \perp (\hat{Z}_1,\ \overline{X}^{(2)}_k)$, and (b) follows, since $\hat{M}^{(1)}_2 \perp (\hat{Z}_2,\ \overline{X}^{(2)}_k)$.

We will also use the following notation: $\hat{M}^{(1)}_1 \triangleq (\overline{M}^{(1)}_{k,1}, \ldots, \overline{M}_{k-N_1,1})$, A_i are the indices of messages transmitted in slots $1, \ldots, k-N_1-1$ that are used as secret keys by user i for transmitting messages in slots $k-N_1, \ldots, k$, $\overline{M}^{(i)}_{A_i} = \left(\overline{M}^{(i)}_k, k \in A_i\right)$, $\overline{M}^{(i)}_{A_i^c} = \left(\overline{M}^{(i)}_k, k \in \{1, \ldots, k-N_1-1\}\right)$, similarly we define $\overline{Z}_{A_i}, \overline{Z}_{A_i^c}$. Then we have:

$$\begin{aligned}
&I(\hat{M}^{(1)}_2; \hat{Z}_1 | \hat{Z}_2,\ \overline{X}^{(2)}_k)\\
&\quad \le I(\hat{M}^{(1)}_2,\ \overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2}; \hat{Z}_1,\ |\hat{Z}_2,\ \overline{X}^{(2)}_k)\\
&\quad = I(\overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2}; \hat{Z}_1,\ |\overline{X}^{(2)}_k,\ \hat{Z}_2)\\
&\quad + I(\hat{M}^{(1)}_2; \hat{Z}_1 | \overline{X}^{(2)}_k,\ \hat{Z}_2,\ \overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2})\\
&\quad \overset{(a)}{\le} I(\overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2}; \hat{Z}_1) + I(\hat{M}^{(1)}_2; \hat{Z}_1 | \overline{X}^{(2)}_k,\ \hat{Z}_2,\ \overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2})\\
&\quad \overset{(b)}{=} I(\overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2}; \hat{Z}_1) + 0\\
&\quad = I(\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(1)}_{A_1,2},\ \overline{M}^{(2)}_{A_2,1},\ \overline{M}^{(2)}_{A_2,2}; \hat{Z}_1)\\
&\quad = I(\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2,1}; \hat{Z}_1)\\
&\quad + I(\overline{M}^{(1)}_{A_1,2},\ \overline{M}^{(2)}_{A_2,2}; \hat{Z}_1 | \overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2,1})\\
&\quad \overset{(c)}{=} I(\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2,1}; \hat{Z}_1) + 0\\
&\quad = I(\overline{M}^{(1)}_{A_1,1}; \hat{Z}_1) + I(\overline{M}^{(2)}_{A_2,1}; \hat{Z}_1 | \overline{M}^{(1)}_{A_1,1})\\
&\quad \le I(\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_1,1}; \hat{Z}_1) + I(\overline{M}^{(2)}_{A_2,1}; \hat{Z}_1 | \overline{M}^{(1)}_{A_1,1})\\
&\quad \overset{(d)}{\le} 2n_1\varepsilon + I(\overline{M}^{(2)}_{A_2,1}; \hat{Z}_1 | \overline{M}^{(1)}_{A_1,1})\\
&\quad \overset{(e)}{=} 2n_1\varepsilon + I(\overline{M}^{(2)}_{A_2,1}; \overline{Z}_{A_2},\ \overline{Z}_{A_2^c} | \overline{M}^{(1)}_{A_1,1})\\
&\quad = 2n_1\varepsilon + I(\overline{M}^{(2)}_{A_2,1}; \overline{Z}_{A_2} | \overline{M}^{(1)}_{A_1,1})\\
&\quad + I(\overline{M}^{(2)}_{A_2,1}; \overline{Z}_{A_2^c} | \overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2})\\
&\quad \triangleq 2n_1\varepsilon + I_1 + I_2\ ,
\end{aligned} \tag{58}$$

where:

- (a) follows, because $\hat{Z}_1 \leftrightarrow (\overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2}) \leftrightarrow (\hat{Z}_2,\ \overline{X}^{(2)}_k)$,
- (b) follows, since $\hat{M}^{(1)}_2 \leftrightarrow (\overline{M}^{(1)}_{A_1},\ \overline{M}^{(2)}_{A_2},\ \hat{Z}_2,\ \overline{X}^{(2)}_k) \leftrightarrow \hat{Z}_1$,
- (c) follows, since $(\overline{M}^{(1)}_{A_1,2},\ \overline{M}^{(2)}_{A_2,2}) \perp (\hat{Z}_1,\ \overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2,1})$,
- (d), (j) and (m) follows by wiretap coding,
- (e) follows, since $\hat{Z}_1 = (\overline{Z}_1,\ \ldots,\ \overline{Z}_{k-N_1}) = (\overline{Z}_{A_2},\ \overline{Z}_{A_2^c})$.

Now, we evaluate I_2:

$$\begin{aligned} I_2 &= I(\overline{M}^{(2)}_{A_2,1};\overline{Z}_{A_2^c}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2}) \\ &= H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2}) \\ &- H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2},\ \overline{Z}_{A_2^c}) \\ &\overset{(a)}{=} H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2,1},\ \overline{Z}_{A_2,2}) \\ &- H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1\cap A_2,1},\ \overline{Z}_{A_2,1}) \\ &\overset{(b)}{=} H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1\cap A_2,1},\ \overline{Z}_{A_2,1}) \\ &- H(\overline{M}^{(2)}_{A_2,1}|\overline{M}^{(1)}_{A_1\cap A_2,1},\ \overline{Z}_{A_2,1}) = 0\ , \end{aligned} \tag{59}$$

where (a) and (b) follow because $\overline{M}^{(1)}_{A_1,1}$ and $\overline{M}^{(1)}_{A_1,1}$ are used as keys only in slots $k-N_1, \ldots, k$.

Next, we evaluate I_1:

$$\begin{aligned} I_1 &= I(\overline{M}^{(2)}_{A_2,1};\ \overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1,1};\ \overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,\ 1}) \\ &\triangleq I_3 + I_4\ . \end{aligned} \tag{60}$$

Now:

$$\begin{aligned} I_3 &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2\cap A_1},\overline{Z}_{A_2\cap A_1^c}|\overline{M}^{(1)}_{A_1,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2\cap A_1^c}|\overline{M}^{(1)}_{A_1,1}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\ \overline{Z}_{A_2\cap A_1}|\overline{M}^{(1)}_{A_1,1},\overline{Z}_{A_2\cap A_1^c}) \\ &\triangleq I_{31} + I_{32}\ . \end{aligned} \tag{61}$$

Consider:

$$\begin{aligned} I_{31} &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,\ 1};\overline{Z}_{A_2\cap A_1^c}|\overline{M}^{(1)}_{A_1,\ 1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,\ 1};\overline{Z}_{A_2\cap A_1^c,\ 1},\ \overline{Z}_{A_2\cap A_1^c,2}|\overline{M}^{(1)}_{A_1,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1^c,2}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c,1}) \\ &\overset{(a)}{\leq} I(\overline{M}^{(1)}_{A_2\cap A_1^c,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1^c,1}) + 0 \\ &\overset{(b)}{\leq} 2n_1\varepsilon, \end{aligned} \tag{62}$$

where (a) follows, since $\overline{Z}_{A_2\cap A_1^c,2} \perp (\overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c,1})$, (b) follows from wiretap coding and that $\overline{M}^{(1)}_{A_1,1} \perp (\overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1^c,1})$. Next consider the second term of (61). We get:

$$\begin{aligned} I_{32} &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1,1},\ \overline{Z}_{A_2\cap A_1,2}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1,2}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c},\ \overline{Z}_{A_2\cap A_1,1}) \\ &\overset{(a)}{=} I(\overline{M}^{(2)}_{A_2\cap A_1^c,1};\overline{Z}_{A_2\cap A_1,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c}) + 0 \\ &= H(\overline{M}^{(2)}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c}) \\ &- H(\overline{M}^{(2)}_{A_2\cap A_1^c,1};|\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c},\ \overline{Z}_{A_2\cap A_1,1}) \\ &\overset{(b)}{=} H(\overline{M}^{(2)}_{A_2\cap A_1^c,1}|\overline{Z}_{A_2\cap A_1^c}) - H(\overline{M}^{(2)}_{A_2\cap A_1^c,1}|\overline{Z}_{A_2\cap A_1^c}) \\ &= 0\ , \end{aligned} \tag{63}$$

where (a) follows, since $\overline{Z}_{A_2\cap A_1,2} \perp (\overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1^c},\ \overline{Z}_{A_2\cap A_1,1})$, (b) follows, since $\overline{M}^{(1)}_{A_1,1} \perp (\overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1^c})$ and $(\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1,1}) \perp (\overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1^c})$.

Finally, we consider:

$$\begin{aligned} I_4 &= I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2,1},\ \overline{Z}_{A_2,2}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2,2}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2,1}) \\ &\overset{(a)}{=} I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) + 0 \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2\cap A_1,1},\ \overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &= I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2\cap A_1,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &+ I(\overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1,1}) \\ &\leq I(\overline{M}^{(2)}_{A_2\cap A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1,1};\overline{Z}_{A_2\cap A_1,1}|\overline{M}^{(1)}_{A_1,1}) \\ &+ H(\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1,1}) \\ &- H(\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(1)}_{A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1},\ \overline{Z}_{A_2\cap A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1,1}) \\ &\overset{(b)}{\leq} 2n_1\varepsilon + H(\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &- H(\overline{Z}_{A_2\cap A_1^c,1}|\overline{M}^{(2)}_{A_2\cap A_1^c,1}) \\ &= 2n_1\varepsilon, \end{aligned} \tag{64}$$

where (a) follows, since $\overline{Z}_{A_2,2}$ is independent of the rest of the terms in the expression, (b) follows, because $(\overline{Z}_{A_2\cap A_1^c,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \perp (\overline{M}^{(1)}_{A_1,1},\overline{Z}_{A_2\cap A_1,1})$ and $(\overline{Z}_{A_2\cap A_1^c,1},\ \overline{M}^{(2)}_{A_2\cap A_1^c,1}) \perp (\overline{M}^{(1)}_{A_1,1},\ \overline{Z}_{A_2\cap A_1,1},\ \overline{M}^{(2)}_{A_2\cap A_1,1})$.

Hence, we have from (60) that $I \leq 6n_1\varepsilon$. Thus, we get:

$$I(\hat{M}_2^{(1)};\hat{Z}_2|\hat{Z}_1,\ \overline{X}_k^{(2)}) \leq 6n_1\varepsilon\ , \tag{65}$$

and the lemma is established. □

References

[1] A. D. Wyner, “The wire-tap channel”, *Bell Syst. Techn. J.*, vol. 54, no. 8, pp. 1355–1387, 1975 (DOI: 10.1002/j.1538-7305.1975.tb02040.x).

[2] P. K. Gopala, L. Lai, and H. El Gamal, “On the secrecy capacity of fading channels”, *IEEE Trans. on Inform. Theory*, vol. 54, no. 10, pp. 4687–4698, 2008 (DOI: 10.1109/TIT.2008.928990).

[3] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, “Wireless information-theoretic security”, *IEEE Trans. on Inform. Theory*, vol. 54, no. 6, pp. 2515–2534, 2008 (DOI: 10.1109/TIT.2008.921908).

[4] O. Gungor, J. Tan, C. E. Koksal, H. El-Gamal, and N. B. Shroff, “Secrecy outage capacity of fading channels”, *IEEE Trans. on Inform. Theory*, vol. 59, no. 9, pp. 5379–5397, 2013 (DOI: 10.1109/TIT.2013.2265691).

[5] Y. Liang and H. V. Poor, “Multiple-access channels with confidential messages”, *IEEE Trans. on Inform. Theory*, vol. 54, no. 3, pp. 976–1002, 2008 (DOI: 10.1109/TIT.2007.915978).

[6] E. Tekin and A. Yener, “The Gaussian multiple access wire-tap channel”, *IEEE Trans. on Inform. Theory*, vol. 54, no. 12, pp. 5747–5755, 2008 (DOI: 10.1109/TIT.2008.2006422).

[7] S. M. Shah, V. Kumar, and V. Sharma, “Achievable secrecy sum-rate in a fading MAC-WT with power control and without CSI of eavesdropper”, in *Proc. of Int. Conf. on Sig. Process. and Commun. SPCOM 2012*, Bangalore, India, 2012 (DOI: 10.1109/SPCOM.2012.6290033).

[8] Y. Liang *et al.*, “Information theoretic security”, *Foundations and Trends in Commun. and Inform. Theory*, vol. 5, no. 4–5, pp. 355–580, 2009 (DOI: 10.1561/0100000036).

[9] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011 (ISBN: 9780511977985).

[10] R. Liu and W. Trappe (Ed.), *Securing Wireless Communications at the Physical Layer*. Boston, MA: Springer, 2010 (ISBN: 9781441913852).

[11] U. M. Maurer, “Secret key agreement by public discussion from common information”, *IEEE Trans. on Inform. Theory*, vol. 39, no. 3, pp. 733–742, 1993 (DOI: 10.1109/18.256484).

[12] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel”, *IEEE Trans. on Inform. Theory*, vol. 51, no. 1, pp. 44–55, 2005 (DOI: 10.1109/TIT.2004.839515).

[13] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011 (ISBN: 9780511921889).

[14] M. Bloch and N. Laneman, “Strong secrecy from channel resolvability”, *IEEE Trans. on Inform. Theory*, vol. 51, no. 1, pp. 44–55, 2011 (DOI: 10.1109/TIT.2013.2283722).

[15] M. Wiese and H. Boche, “Strong secrecy for multiple access channels”, in *Information Theory, Combinatorics, and Search Theory*, H. Aydinian, F. Cicalese, and C. Deppe, Eds. *LNCS*, vol. 7777, pp. 71–122. Berlin, Heidelberg: Springer, 2013 (DOI: 10.1007/978-3-642-36899-8_4).

[16] M. H. Yassaee and M. R. Aref, “Multiple access wiretap channels with strong secrecy”, in *Proc. Inform. Theory Worksh. ITW 2010*. Dublin, Ireland, 2010 (DOI: 10.1109/CIG.2010.5592953).

[17] S. Leung-Yan-Cheong and M. E. Hellman, “The Gaussian wire-tap channel”, *IEEE Trans. on Inform. Theory*, vol. 24, no. 4, pp. 451–456, 1978 (DOI: 10.1109/TIT.1978.1055917).

[18] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y.-H. Kim, “Wiretap channel with secure rate-limited feedback”, *IEEE Trans. on Inform. Theory*, vol. 55, no. 12, pp. 5353–5361, 2009 (DOI: 10.1109/TIT.2009.2032814).

[19] L. Lai, H. El Gamal, and H. V. Poor, “The wiretap channel with feedback: Encryption over the channel”, *IEEE Trans. on Inform. Theory*, vol. 54, no. 11, pp. 5059–5067, 2008 (DOI: 10.1109/TIT.2008.929914).

[20] W. Kang and N. Liu, “Wiretap channel with shared key”, in *Proc. 2010 IEEE Informa. Theory Worksh.*, Dublin, Ireland, 2010 (DOI: 10.1109/CIG.2010.5592665).

[21] D. Kobayashi, H. Yamamoto, and T. Ogawa, “Secure multiplex coding attaining channel capacity in wiretap channels”, *IEEE Trans. on Inform. Theory*, vol. 59, no. 12, pp. 8131–8143 (DOI: 10.1109/TIT.2013.2282673).

[22] S. M. Shah, S. Parameswaran, and V. Sharma, “Previous messages provide the key to achieve Shannon capacity in a wiretap channel”, in *Proc. IEEE Int. Conf. on Commun. Workshops ICC 2013*, Budapest, Hungary, 2013, pp. 697–701 (DOI: 10.1109/ICCW.2013.6649323).

[23] S. M. Shah and V. Sharma, “Achieving Shannon capacity region as secrecy rate region in a multiple access wiretap channel”, in *Proc. IEEE Wireless Commun. and Network. Conf. WCNC 2015*, New Orleans, LA, USA, 2015 (DOI: 10.1109/WCNC.2015.7127565).

[24] R. Ahlswede, “Multi-way communication channels”, in *Second International Symposium on Information Theory: Tsahkadsor, Armenia, U. S. S. R., September 2-8, 1971*, F. Csáki, B. N. Petrov, Eds. Budapest: Akademiai Kiado, 1973 [Online]. Available: https://pub.uni-bielefeld.de/download/1780371/2312888/Ahlswede_12.pdf

[25] H. H.-J. Liao, “Multiple access channels”, DTIC Document, Defense Tech. Inform. Center, Fort Belvoir, VA, Tech. Rep., 1972 [Online]. Available: http://www.dtic.mil/docs/citations/AD0753127

[26] U. Maurer and S. Wolf, “Information-theoretic key agreement: From weak to strong secrecy for free”, in *Advances in Cryptology – EUROCRYPT 2000. International Conference on the Theory and Application of Cryptographic Techniques Bruges, Belgium, 2000, Proceedings*, B. Preneel, Ed. LNCS, vol. 1807, pp. 351–368. Springer, 2000 (DOI: 10.1007/3-540-45539-6_24).

[27] E. Tekin and A. Yener, “Secrecy sum-rates for the multiple-access wire-tap channel with ergodic block fading”, in *Proc. of 45th Ann. Allerton Conf. on Commun., Control and Comput.*, Monticello, IL, USA, 2007, vol. 2, pp. 856–863 (ISBN: 9781605600864).

Shahid Mehraj Shah received his B.Tech. in ECE from the National Institute of Technology, Srinagar, and Ph.D. from the Indian Institute of Science Bangalore, in 2008 and 2017, respectively. Since 2018, he has been working as an Assistant Professor at the Department of ECE at NIT Srinagar, where he leads the communication control and learning lab. His research interests include information theory, wireless communication, cyber-physical systems, machine learning.

https://orcid.org/0000-0002-8583-7904
E-mail: shahid.nit@gmail.com
Communication Control & Learning Lab
Department of Electronics and Communication Engineering
National Institute of Technology
Srinagar, Jammu and Kashmir, India

Paper

Telecom Operator's Approach to QoE

Karol Kowalik[1], Paweł Andruloniw[1,2], Bartosz Partyka[1,2], and Piotr Zwierzykowski[2]

[1] *Fiberhost S.A., Poznań, Poland*
[2] *Faculty of Computing and Telecommunications, Poznań University of Technology, Poznań, Poland*

https://doi.org/10.26636/jtit.2022.159522

Abstract—Telecommunication networks are ever more frequently relying on artificial intelligence and machine learning techniques to detect specific use patterns or potential errors and to take automated decisions when these are encountered. This concept requires that methods be employed to measure the level of quality of a given telecommunication service, i.e. to verify quality of service (QoS) metrics. In a broader context, methods assessing the entire user experience (quality of experience – QoE) are required as well. In this article, various approaches to assessing QoS, QoE and the related metrics are presented, with a view to implement these at an FTTH network operator in Poland. Since this article presents the architecture of the system used to analyze QoE performance based on a number of QoS metrics collected by the operator, we also provide a comprehensive introduction to the QoS and QoE metrics used herein.

Keywords—quality of experience, quality of service.

1. Introduction

In the telecommunications industry, quality of service is measured based on three parameters: throughput, delay, and jitter. This simplified approach is often further restricted to bandwidth only by crowdsourcing websites which measure the network's speed (e.g. speedtest.net or fireprobe.net). Statistics collected by such websites are often used to create Internet speed rankings listing specific countries or operators [1].

Throughput is relatively easy to measure and is often well understood by inexperienced users. On the other hand, many people are still unaware of what maximum download speed of their Internet connection is. For them, comfortable access to the services they use is far more important. Parameters assessing the quality of a given service from the technical point of view are known as quality of service (QoS) metrics. QoS is also relied upon to describe a set of methods used to define how the network prioritizes the resources available.

In contrast, quality of experience (QoE) is a measure that aims to reflect human perception of a given service that technical key performance indicators (KPIs) are unable to reflect. In some cases, the perception does not coincide with these KPIs [2].

In this article, we describe a different approach to QoS and QoE assessments, as well as the QoS and QoE metrics available to telecom operators. In addition, we present a practical implementation of QoE metrics by Fiberhost – an FTTH network operator from Poland that was spun off from INEA. Since this article presents the architecture of an QoE performance analysis system implemented by Fiberhost which relies, due to the specific needs, on numerous QoS metrics collected by the operator, we also provide a comprehensive introduction to the QoS and QoE metrics used.

2. Quality of Service Methods

In today's converged (fixed and mobile) IP networks, all traffic shares the same network resources. However, historically, voice and data services were rendered using separate networks. In order to enable converged IP networks to deal with different types of traffic, QoS mechanisms are implemented today, allowing different types of services to be distinguished, prioritized and forwarded according to the required characteristics of a given traffic category. In some extreme scenarios, routing decisions may also be based on QoS requirements. A group of services that share similar QoS requirements is often referred to as a class of service (CoS). A common approach in today's IP networks is to classify services into classes based on QoS requirements concerning throughput, delay, jitter and packet loss. The QoS DiffServ model [3] utilizes a 6-bit DS field in the IP header to define up to 64 different classes services. QoS methods relied upon by modern network devices define per hop behavior only. Consequently, they fail to monitor and control the services' overall performance. Therefore, we propose to deploy the QoS metrics described in the next section.

2.1. QoS Metrics

Many different metrics and KPIs may be distinguished. Their design may be based on two techniques:

- active probing – when probing packets are sent as part of the service in order to measure a given set of quality-of-service metrics (such as throughput, latency, jitter and packet loss). RFC 2544 [4] and ITU-T Y.1564 [5] documents provide examples of methodologies concerned with such active probing;

- passive probing – when network devices monitor the traffic flow in order to estimate quality-of-service metrics (such as throughput, latency, jitter and packet loss). Passive probing may be implemented by observing SNMP [6] counters specifying packet loss and throughput values, etc.

Various metrics and KPIs used for assessing QoS are available. In this paper, we present the most typical of them. throughput is the most frequently mentioned metric and it is often referred to as Internet speed.

Throughput a.k.a. Internet speed. While throughput is well understood by the telecommunications community, there are additional issues that network operators have to deal with. For example, in the European Union (as defined in [7]), each Internet service provider (ISP) should specify, in the contract offered to the end customer, not only the maximum Internet speed, but also they must precisely determine:

- how traffic management measures, throughput constraints or other quality-of-service parameters may affect the quality of Internet service,
- minimum, normally available, maximum and advertised download and upload speeds of Internet services for fixed networks, or estimated maximum and advertised download and upload speeds for mobile networks,
- on remedies available to the end user in the event of any continuous or recurring discrepancy between the actual performance of the Internet service in terms of speed or other QoS parameters and the performance indicated in the contract.

The speed of the Internet itself does not reflect QoS well, and speed is also not directly proportional to QoS.

Delay and jitter metric. Delay and jitter are mostly caused by packets queuing at every transmission hop from the source to the destination, but may also stem from physical distance, topology of the network, forwarding mechanisms used by the networking devices and the QoS control methods implemented. Generally, when packet queueing takes place in any network device, a bottleneck may be created at some point that will affect latency and jitter and, consequently, service response time. When latency is significant, the end-user may observe what is sometimes referred to as the "spinning wheel of death". Both delay and jitter may be measured using the methods described in RFC 2544 [4].

2.2. QoS Layers

An obvious but interesting observation concerning QoS metrics is that a specific value may differ depending on whether we focus on theoretical or measured values, and may depend significantly on the measurement methodology applied. Therefore, as a result of two European Projects titled "Mapping of fixed and mobile broadband services in Europe (SMART 2014/0016)" and "Study on Broadband and Infrastructure Mapping (SMART 2012/0022)" described in mapping project [8], three layers of QoS parameters are analyzed:

- QoS-1 – calculated availability of service, theoretical calculations the coverage offered by network operators,
- QoS-2 – measured provision of service, with the value measured using test equipment, without taking into account the end user's environment,
- QoS-3 – measured experience of service, where the value is determined via crowd sourcing tools, like online speed tests, and takes into consideration the end user's environment.

Usability and correlation of the metrics collected with respect to these three levels require that further studies be conducted. Unfortunately, data collected at these levels may be biased. For example, measurements sources during wireless test drives are often collected during business hours, when qualified technical teams are available, not but not during service peak hours. Also, speed tests performed by end users themselves are often conducted in response to service interruptions, and users treat those tests as a diagnostic tool. Therefore, such systematic errors must be taken into account when investigating the correlation between QoS-1, QoS-2, and QoS-3.

3. Quality of Experience

To an inexperienced user, simple QoS measurements may seem fairly unrelated to end user experience. It is not easy to estimate how specific values of throughput, delay, jitter and packet loss affect the quality of service. Therefore, the more complex quality-of-experience measure has been devised to reflect people's perception of a given service. Both ETSI and ITU-T proposed methods and recommendations for measuring the quality of service from the end-user's point of view. Many of other scoring methods rely on various methodologies that are difficult to compare. Their detailed descriptions are not publicly available in many instances and they are not prepared for performing QoE testing [2].

It is widely known that quality measurements might be subjective or objective in nature. ITU-T defined the mean opinion score (MOS) parameter that covers both of the aforementioned types of measurements. In the first type, the arithmetic mean of the values collected from observations is calculated. This type of assessment might be considered as difficult to conduct by telecommunications operators such as Fiberhost, due to the fact that such a task is time consuming and due to the scale factor. Hence, at Fiberhost, we use a subjective metric called the Net Promoter Score [20]. It is a sampled metric that is updated annually. In the implementation described in this article, we focused on QoS metrics we can monitor and aggregate on a daily basis. It is also worth to mentioning that subjective assessments may not be reliable due to human factors, i.e. the observer's condition or mood, as well as due to experiment environment-related factors, such as acoustic conditions, equipment used, etc. Objective assessments are based on predictive calculations that should reflect the subjective evaluation.

Objective models rely on external factors that are free from subjective judgement. In the process of creating objective models, the main issue is to find the relevant factors that cover specific network parameters or service stream packets. The main advantage is the repeatability of the evaluation process, its easy scalability, as well as real-time judgement. Nowadays, the objective methods might be realized with the use of machine learning models that may be more accurate than the standard approach due to the incremental learning process and due to their ability to analyze larger data sets.

3.1. QoE Assessment for Voice Services

The process of gathering data and assessing the quality of voice services might be based on intrusive or non-intrusive methods. The former are based on a probe that attempts to connect with the end user's terminal and collect specific information. The non-intrusive method requires a physical probe installed at the user's terminal, and is difficult to accomplish from the point of view of the telecom operator. In the case of a fixed network operator, voice quality might refer to VoIP services. Depending on whether the operator offers voice services in conjunction with other services, such as television or Internet access, quality assessment should be performed both for the VoIP service alone and for VoIP while taking advantage of other services [11].

3.2. QoE Assessment for Video Services

The quality of video services should be considered separately for IP television (IPTV) and for adaptive streaming. Video quality in IP television services is covered by ETSI standards, both for linear content and for video-on-demand services. The standards describe viable indicators enabling to perform an objective assessment based on models or evaluations performed by robots simulating end user behaviors. Service availability may be measured based on channel and service group availability, i.e. on the number of successful channel start-ups divided by the number of attempts.

Video quality evaluation relies on ITU-T recommendations and uses the MOS indicator with the ACR rating scale. It is worth mentioning that video quality measurements must be objective, i.e. computed by a known QoE algorithm. The main requirement is to use ”no reference” models, while the choice of the method is the operator's decision, as no specific algorithms or models are imposed by ETSI. “No reference” algorithms do not refer to the original signal or any part of that signal. The predictions created by the abovementioned models may be based on network parameters, IPTV stream data or infrastructure-related parameters. As in the case of video, audio quality must be analyzed with the use of a “no reference” model, and cannot be performed based on speech quality assessment models [12].

From the point of view of a telecommunications operator, over-the-top (OTT) services might be considered as two different entities. One of them includes typical OTT applications, i.e. YouTube, Netflix, HBO Max etc., while the other is an own television system that might be based on adaptive streaming with the use of the HTTP standard, just as it is the case in classic OTT applications, but is provided with the use of the operator's network. The abovementioned television architecture may rely on multicast technology not only in the core network, but also in the last mile network. OTT television systems are often referred to as IPTV 2.0.

The assessment of quality of adaptive streaming services is provided for in the ITU-T P.1204 recommendation [13] and in the ETSI TR 103 488 report [14]. Adaptive streaming systems may use different protocols, for example HTTP live streaming (HLS) or dynamic adaptive streaming over HTTP (DASH). Quality measurements may be performed for pure OTT applications and for television services based on adaptive streaming provided on a managed network. A comprehensive overview of developments related to visual quality monitoring is presented in [19].

3.3. Compound STQ Metric

In addition to single parameter metrics, such as throughput, delay and jitter, compound QoS metrics may be distinguished as well, including speech and multimedia transmission quality (STQ) [15]. STQ describes the best practices for benchmarking a mobile network and is also widely adapted and used by ISPs in wired networks. This document focuses mainly on the aspects of STQ that are significant for ISPs and allow the quality of the services provided to be evaluated. ETSI recommends mainly a clear interpretation of the benchmark results. All results must include information about:

- the scoring model used,
- basic KPIs values measured in the test,
- number of samples and/or number of tests,
- methodology used (including hardware configuration, connection sequences, test servers and test pages),
- data collection areas and packages (tariff plans) for data collection.

Tests should be performed on the same servers by all operators. The location of servers should not favor any service provider. The choice of the test page can have an impact on the test results. Therefore, to provide a representative comparison, it is necessary to prepare various and reliable pages for the test. In [15], ETSI suggests using a minimum of 6 tests, while 10 or more are recommended. Eliminating a single web page from the overall pool should not affect the diversity of those pages. The ranking should use pages that are popular among customers. Their minimum size should exceed the download success criterion. The principle of proportion of active pages should be applied – 4 common and 6 country dependent. It is not recommended to block ads and use websites operating within one CDN. Web pages that are mainly accessible through dedicated apps should not be used for testing as well. The pool used in the ranking should not contain subpages with discriminatory and indecent content or those prohibited by law.

STQ sets the scoring benchmark for mobile networks operating across large geographic regions. The overall result is calculated from the individual measurement results and aggregation is performed using a weighted factor. The authors distinguish four levels presented in Fig. 1. The highest index is important for the business assessment of the quality of the entire network of a given operator and for comparing it with other ISPs.

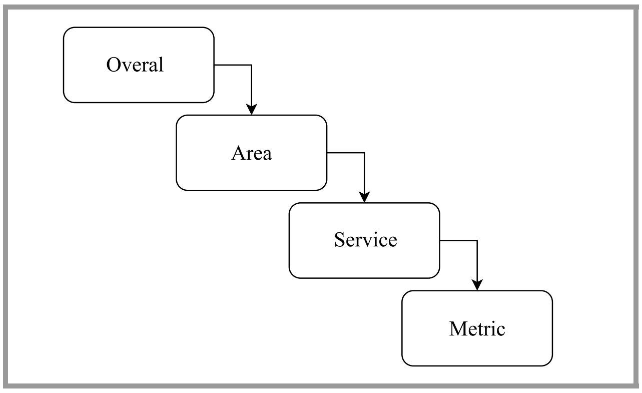


Fig. 1. Scoring layers.

The quality of a stationary cable network, unlike in the case of any mobile networks, does not depend on the geographical location of the customer's device. It is economically viable for an ISP to aim for a similar level of network saturation in all locations. Therefore, the "area" layer is not applicable in the weighting of stationary cable network's score. However, it may still be of interest for the operator, as a tool for comparing different parts of the network. In this context, a given "area" does not have to describe specific geographic coverage, but may also be an area in terms of the services provided based on the same technology or devices (e.g. OLT or aggregation router).

User profile is an important element of STQ, as it affects the "service" and "metric" layers. Different users have different service requirements and expectations. The recommendation is to rate the services based on the profile level associated with the highest requirements.

Research is being conducted at Fiberhost to define profiles and evaluate services in the context of different customer types. It has been assumed that profiles will be created on the basis of DNS queries collected over a period of time, though this is only one of the potential approaches.

The "service" level is the next aspect that may be assessed, as it covers the basic services provided. The division proposed in STQ is shown in Fig. 2.

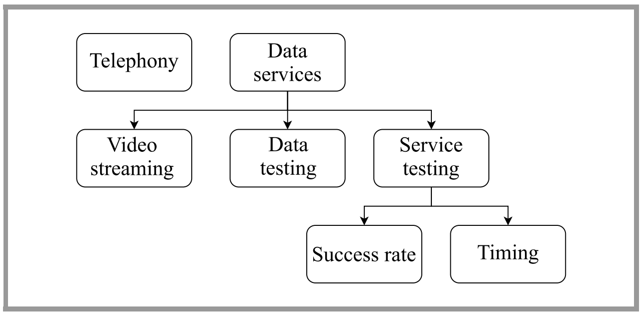


Fig. 2. Services testing.

At the lowest level, each service is assessed based on the metrics' basis obtained during the tests. The KPIs are mostly described in the document ETSI TS 102 250-2 [16] document. Table 1 lists the metrics for each of the services. Telephony is separated from data transmission services, due to the high impact of delays and errors on the assessment of its quality.

Each of the tests is multi-layered, with the overall test score at the top – a value calculated on the basis of the weighted results from the test scenarios for telephony and data services. The data service index consists of video, data, and service testing results. The total weight of the components at each level is always 100%.

STQ suggests the scoring and the weighting for each metric. Here, we are considering the general concept of QoS testing and parameter evaluation details are not covered here. Examples of weighting factors, limits and thresholds are provided in Annex A of ETSI TR 103 559 V1.1 [9].

3.4. Network Performance Score

Network performance score (NPS) is a proposal aiming to implement the good testing practices presented in STQ [18]. NPS, similarly to STQ, is designed primarily for mobile networks, but the concept is universal and is considered at Fiberhost. As it was the case STQ, NPS indicators are divided into 3 levels.

At the highest level, the coefficients from level 2 to one index describing the entire network are aggregated. However, when used in Fiberhost, instead of a weight based on geographic location, a weight was proposed based on the number of customers connected behind a given aggregation device (equivalent of a region in mobile networks).

The second level aggregates the service results by weighting. The indicator at this level shows the QoS in a given case, at a given time and location. In the case of an ISP, the location should be understood as the physical location of the infrastructure service.

Table 1
Services test metrics

Service	Metric	Description
Telephony	Telephony success ratio	Success rate of the call
	Setup time	Time from the initiation of the connection to the calling of the other party to the connection
	Listening quality	The value is calculated on a sample basis, using recommendations from ITU-T P.863 [17]
Video testing	Video streaming service success ratio	Success rate of the video stream received
	Setup time	Time from stream request to the displaying of the first picture
	Video quality	Video service metering has already been described in this document
Data testing	Success ratio	Success rate for HTTP uploads and downloads
	Throughput	Data rate or throughput for HTTP uploads and downloads
Services	Browsing	Success rate of all download and page open attempts, page response times to those operations
	Social media	The ratio of success user interacts with the media to all interaction and response times to those operations
	Messaging	The ratio of success sending message to sending message and delivery time

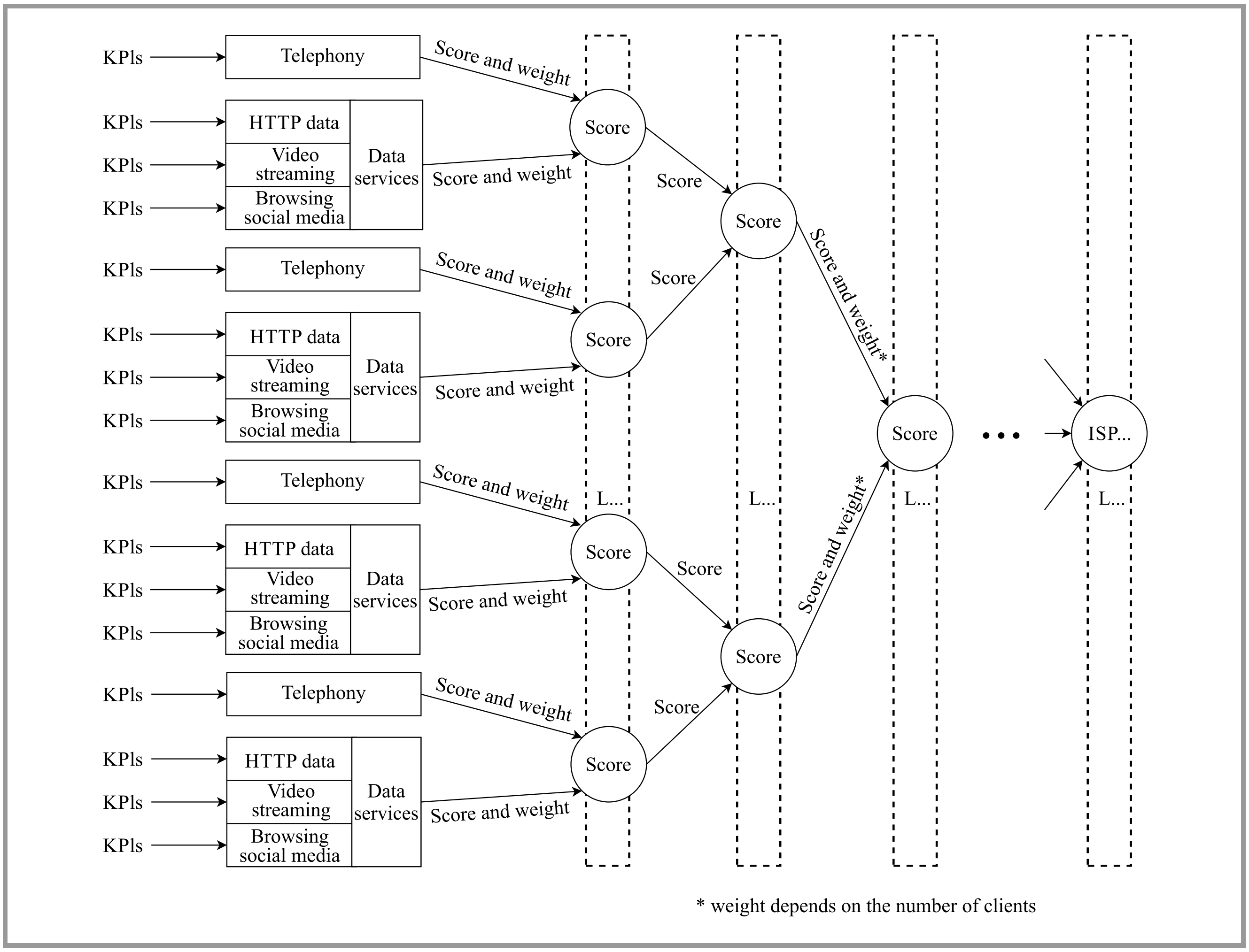


Fig. 3. Example implementation of DX based on NPS.

At the lowest (third) level, the recommendation is to divide QoE indicators into 3 groups:

- service availability – access and response time only,
- waiting for action – time between the request and commencing or ending the requested task,
- quality – media quality (video, voice, images).

Only those indicators that are decoded by human perception have an impact on the result. They all have to be normalized to the same scale and they must have a certain weight. STQ and NPS do not define the method of filtering data noise which can occur upon test device failure or a local power failure. Such filtering needs to be provided to show the overall condition of the services and not the extreme cases resulting from measurement errors and noise. There is also no definition of the scoring computation intervals at the aggregation levels. Excessively long sampling intervals will average out many short variations of the individual parameters and will reduce their impact on the overall result. Too frequent readings will introduce a lot of noise due to the number of measurements and potential local errors of the test equipment. Each ISP has to assess and define the above from their own perspective.

Figure 3 shows a sample NPS implementation customized to satisfy the needs of Fiberhost. Aggregation of the results by location (device) may occur on many levels, depending on the topology of the telecom operator's network. The division of the network into aggregation nodes for which the score is calculated may depend on a number of factors, such as physical and logical topology, geographic extension and the service provision model.

Both STQ and NPS fail to address issues related to the correctness of QoE representation by the KPIs. These issues were discussed in [17].

On October 4, 2021, a global failure of Facebook, the largest social networking site, occurred. As a result, telecom operators faced a new challenge in the form of an increase in the number of DNS queries to recursive servers. In the case of INEA, the number of requests doubled at 17:43 and during the peak hours it reached 250% of the normal query load, compared to the reference value recorded on Monday of the previous week. The problem was solved at 23:27. A conclusion was drawn from this failure, according to which service quality indicators used in telecommunications networks need to be verified and modified. In the recommendations, such as STQ (ETSI) and their proposed implementations, such as network performance score (NPS), the impact of the "quality" of social media services used by users is set at 15%. However, these types of situations prove a much greater impact on the overall indicator and the quality level perceived by the customer. The issue concerns an OTT service and the operator has little impact on its quality, but the failures have a large impact on the users themselves. ISPs do not have accurate data for such services and cannot optimize OTT services. Therefore, failures of this type are not included by INEA in its CX and DX logs. If we manage to develop, in the future, a form of cooperation providing better visibility of the quality of OTT services, we will be able to influence the level of that quality, and then quality scoring will be added to CX and DX.

The scoring should also distinguish between normal operation and a failure occurring outside the operator's network, such as the aforementioned unavailability of Facebook or problems at the interface with other operators or at traffic exchange points.

4. Implementation Example

In this section, we present the effort of Fiberhost – an FTTH operator from Poland – aiming at implementing a system for monitoring not only QoE related to the services offered, but all the interactions with the services and the operator as such. The aim was to compare the performance of different geographic areas of the operator's network, taking into account not only the average quality of experience in a given area, but also all interactions between the end user on the one hand, and the operator and the services purchased on the other. Thus, the definition used in this scenario was even of a broader nature that in the case of NPS or STQ, as described in previous sections. Because of such a broad scope of the study, Fiberhost realized that no third-party systems that are ready for use are available and decided to develop this system in-house. At Fiberhost, the system is referred to as CX. CX stands for Customer eXperience and is defined as:

$$CX = DX + AX.$$

Thus, CX is the sum of Digital eXperience (DX) which involves the measurements of complex QoE metrics similar to NPS, and Analog eXperience (AX), aiming to capture all human operator interactions, including purchasing the services, troubleshooting, payments, promotions, etc. AX is related to the operator's operational efficiency and covers the following: customer complaints made by phone or with the use of other forms of interaction, complaint handling time, installation time, number of customer complaints resolved during the first contact, etc. DX focuses on all digital services, taking into account such indicators as: availability of the services, scheduled works, maintenance, service failures, service degradation, CPE logout time, physical signal parameters, L2 (layer 2) or L3 (layer 3) transport performance metric, and network saturation events. Each metric contributing to DX is, just as it is the case with NPS or STQ, summed up and weighted in order to create a compound overall indicator.

Both CX components are structured in a similar way to NPS, so they form a weighted sum of all partial parameters and may be aggregated at different levels. Figure 3 shows details of the DX implementation. While AX is structured in the same way, it includes non-technical parameters related to purchasing the services, troubleshooting, payments and promotions, so we decided to exclude AX from the scope of this paper. The specific weighting of the various DX and AX parameters and, consequently, CX as well, are not disclosed by Fiberhost. However, all metrics are weighted and summed up to form a CX value that ranges from 0% to 100%.

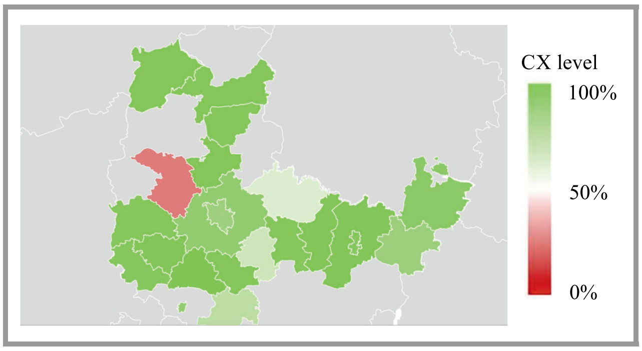


Fig. 4. Example of a CX aggregation calculated for a specific geographical area.

CX is calculated on a monthly and annual, per-customer basis and aggregated CX values may be used to visualize QoE measurements related to a given part of the network. An example of aggregation of CX values pertaining to a specific geographic area is shown in Fig. 4. The map shows the average value of CX for all customers using services provided by the operator in a given geographic area. The areas marked in dark green represent the highest CX rating close to 100%, and correspond to areas where most customers experience no technical issues (represented by the DX component in CX) or are affected by other non-technical issues (represented by the AX component in CX). The red areas represent the lowest CX score close to 0% and correspond to areas where the majority of customers have technical or other non-technical problems.

By presenting DX in the form of a map, we are able to have a good overview of the expected level of CX. However, such an approach fails to illustrate many technical and operational details. Therefore, at Fiberhost, the map is not the only analytical tool and is accompanied by reports that provide detailed information on all the components that make up a given CX value. These reports allow to conduct a detailed analysis in order to identify specific technical or operational aspects that require improvement.

CX is an attempt to reflect customer perceptions of service quality and customer experience in a broad context that includes all digital and analog interactions. CX has enabled Fiberhost to conduct in-depth technical and operational analyses. By providing such a complex metric, Fiberhost was able to determine what the average expected CX score was in all areas. this, in turn, served as a basis for identifying those areas where the CX score was too low and which required immediate attention to improve the maintenance parameters that affect the quality of experience or other non-technical parameters.

The implementation of CX allows Fiberhost to use ML methods for identifying more complex network dependencies and for optimizing service performance. However, ML-based optimization is planned as a future task. Current efforts are aimed at improving CX levels in all of the underperforming areas.

Although this paper does not provide any numerical results, it does provide an insight into the real live implementation of a QoE assessment system deployed by a telecom operator. Therefore, we hope that it will be a valuable guidance for other operators and companies willing to implement similar systems.

5. Conclusions

There are many QoS and QoE metrics available to telecom operators. Simple QoS metrics are easy to collect but do not reflect end user observations. Therefore, in order to compare different networks, by area or country, it is better to use QoE metrics that are meant to reflect human perception. However, any QoS and QoE metrics may be significantly influenced by the methodology, specific hardware, testing intervals, information source, etc. This error should be removed from the collected data to ensure that the QoE metric reflects the actual customer experience.

In this article, we analyze typical QoS and QoE metrics that are currently used by telecom operators. Additionally, we present a complex QoE metric known as CX (Customer eXperience), implemented by Fiberhost in order to get in-depth information about the complex experiences of its customers, including the quality of the service offered and all other interactions.

QoE metrics, such as STQ, NPS or CX, are very important for network operators planning to optimize their infrastructure by deploying machine learning mechanisms, since the first step preceding any optimization consists in understanding current performance of the network. By providing complex measurements of various parameters, QoE metrics are a source of reliable feedback that may be harnessed by artificial intelligence solutions in order to achieve what is often referred to as the intent network, i.e. a network whose operations are defined by the operator's ultimate intent, not the use of any technical terms.

Acknowledgements

The research was carried out as part of the Implementation Doctorate Program of the Ministry of Education and Science in 2021–2024 (Agreement no. DWD/4/24/2020 from 6.11.2020).

References

[1] I. Fogg, "Benchmarking the Global 5G Experience – September 2021", Opensignal, 02-03-2021 [Online]. Available: https://www.opensignal.com/2021/02/03/benchmarking-the-global-5g-experience

[2] J. Berger, "QoE in broadband telecommunication networks", Rohde and Schwarz, 11-2018 [Online]. Available: https://www.itu.int/en/ITU-T/Workshops-and-Seminars/qos/201811/Documents/S4P2-berger-slides.pdf

[3] D. L. Black *et al.*, "An Architecture for Differentiated Services", RFC2475, IETF, 1998 [Online]. Available: https://www.ietf.org/rfc/rfc2475.txt

[4] S. Bradner and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, IETF, RFC Editor, 1999 (DOI: 10.17487/RFC2544), March 1999 [Online]. Available: https://www.rfc-editor.org/info/rfc2544)

[5] Y.1564 Rec. "Ethernet service activation test methodology", ITU-T, 2016 [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.1564-201602-I!!PDF-E&type=items

[6] M. Fedor, M. L. Schoffstall, J. R. Davin, and J. D. Case, "Simple Network Management Protocol (SNMP)", RFC1157, IETF, RFC Editor, 1990 [Online]. Available: https://www.rfc-editor.org/rfc/pdfrfc/rfc1157.txt.pdf

[7] The European Parliament and the Council of the European Union, "Regulation (Eu) 2015/2120 of the European Parliament and of the Council", *Official Journal of the European Union*, 2015 [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/8fdf5d08-93fc-11e5-983e-01aa75ed71a1

[8] European Commission Directorate General for Communications Networks, Content & Technology (DG CNECT), "European broadband mapping project" [Online]. Available: https://www.broadband-mapping.eu/

[9] ITU-T P.800.1, "Methods for objective and subjective assessment of speech and video quality", ITU-T, 2016 [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.800.1-201607-I!!PDF-E&type=items

[10] ITU-T P.800, "Methods for objective and subjective assessment of quality", ITU-T, 1996 [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.800-199608-I!!PDF-E&type=items

[11] ETSI ES 202 765-2, "Speech and Multimedia Transmission Quality (STQ); QoS and network performance metrics and measurement methods", ETSI, 2014 [Online]. Available: https://www.etsi.org/deliver/etsi_es/202700_202799/20276502/01.02.01_60/es_20276502v010201p.pdf

[12] ETSI ES 202 765-4, "Speech and Multimedia Transmission Quality (STQ); Indicators for Supervision of Multiplay Services", ETSI, 2014 [Online]. Available: https://www.etsi.org/deliver/etsi_es/202700_202799/20276504/01.02.01_60/es_20276504v010201p.pdf

[13] ITU-T P.1204, "Video quality assessment of streaming services over reliable transport for resolutions up to 4K", ITU-T, 2020 [Online]. Available: https://www.itu.int/rec/T-REC-P.1204

[14] ETSI TR 103-488, "Speech and Multimedia Transmission Quality (STQ); Guidelines on OTT Video Streaming, Service Quality Evaluation Procedures", ETSI, 2019 [Online]. Available: https://www.etsi.org/deliver/etsi_tr/103400_103499/103488/01.01.01_60/tr_103488v010101p.pdf

[15] ETSI TR 103 559 V1.1, "Speech and multimedia Transmission Quality (STQ); Best practices for robust network QoS benchmark testing and scoring", ETSI, 2019 [Online]. Available: https://www.etsi.org/deliver/etsi_tr/103500_103599/103559/01.01.01_60/tr_103559v010101p.pdf

[16] ETSI, "Speech and multimedia Transmission Quality (STQ); QoS aspects for popular services in mobile networks; Part 2: Definition of Quality of Service parameters and their computation", 2015 [Online]. Available: https://www.etsi.org/deliver/etsi_ts/102200_102299/10225005/02.04.02_60/ts_10225005v020402p.pdf

[17] ITU-T P.863, "Perceptual objective listening quality prediction", 2018 [Online]. Available: https://www.itu.int/rec/T-REC-P.863

[18] W. Yoong, "Network Performance Score (NPS): A method for initiating network improvement with a single QoE centric score", Rohde & Schwarz, 2019 [Online]. Available: https://www.itu.int/en/ITU-T/Workshops-and-Seminars/qos/201903/Documents/Walter_Yoong_Presentation.pdf

[19] M. Leszczuk *et al.*, "Recent developments in visual quality monitoring by key performance indicators", *Multimed. Tools Appl.*, vol. 75, pp. 10745–10767, 2016 (DOI: 10.1007/s11042-014-2229-2).

[20] R. Bergevin, A. Kinder, W. Siegel, and B. Simpson, *Call Centers for Dummies*. Wiley, 2010 (ISBN: 9780470678404).

Karol Kowalik is a telecommunications engineer, holding a Ph.D. from Dublin City University in Ireland. He is a Technology Development Manager at INEA, where he participates in numerous projects covering a wide spectrum of technologies, such as GPON, Wi-Fi, WiMAX, DOCSIS, and MPLS. Currently, he is responsible for technical innovation and validation of new ideas. His research interests include AI/ML, networking, QoS, QoE, switching, routing, network management, as well as wireless and wired access.

E-mail: karol.kowalik@fiberhost.com
Fiberhost S.A.
Wysogotowo, Wierzbowa 84
62-081 Przeźmierowo, Poland

Paweł Andruloniw is graduated from the Electronics and Telecommunications Department of Poznań University of Technology. In 2018, he received his M.Sc. in Computer Networks and Internet Technologies. He is a Ph.D. student at the Poznań University of Technology, specializing in information and communication technologies, and an engineer at the Digital Television Department at Fiberhost S.A., where he conducts research on the usage of artificial intelligence in modern television services. His professional interests include programming automation scripts, big data analysis and continuous improvement of digital television services which have an effect on the quality of customer experience and on customer loyalty increase.

E-mail: pawel.andruloniw@fiberhost.com
Fiberhost S.A.
Wysogotowo, Wierzbowa 84
62-081 Przeźmierowo, Poland

Institute of Computer and Communication Networks
Faculty of Computing and Telecommunications
Poznań University of Technology
Polanka 3
60-965 Poznań, Poland

Bartosz Partyka graduated from Computer Science Department of Poznań University of Technology, majoring in Computer Networks. Bartosz Partyka has several years of experience in designing and implementing high-performance systems, distributed databases and data processing in ICT infrastructures. In 2009, he started to work at Poczta Polska in Poznań. Since 2016, he has been an engineer at the telecommunications company Fiberhost (formerly INEA S.A.) telecommunications company in Poznań. He was responsible for database optimization, data analysis, as well as for analyzing Core, GPON and HFC network parameters for the management. He participates in numerous involving the collection and analysis of network data, optimization of data collection mechanisms, improvement of the quality of services. In 2020, he commenced a Ph.D. studies at the Poznań University of Technology.

E-mail: bartosz.partyka@fiberhost.com
Fiberhost S.A.
Wysogotowo, Wierzbowa 84
62-081 Przeźmierowo, Poland

Institute of Computer and Communication Networks
Faculty of Computing and Telecommunications
Poznań University of Technology
Polanka 3
60-965 Poznań, Poland

Piotr Zwierzykowski received his M.Sc. degree in Telecommunications from Poznań University of Technology, Poland, in 1995, and then a Ph.D. degree (with honors) and D.Sc. degree in Telecommunications from PUT in 2002 and 2015, respectively. Since 1995, he has been working at Poznań University of Technology, Poland, first at the Institute of Electronics and Telecommunications at the Faculty of Electrical Engineering, and then, since 2005, at the Chair of Communications and Computer Networks at the Faculty of Electronics and Telecommunications and since 2019, at the Institute of Communications and Computer Networks at the Faculty of Computing and Telecommunications at Poznań University of Technology. Piotr Zwierzykowski is engaged in research and teaching activities in the field of analysis and modeling of multi-service switching systems and networks.

E-mail: piotr.zwierzykowski@put.poznan.pl
Institute of Computer and Communication Networks
Faculty of Computing and Telecommunications
Poznań University of Technology
Polanka 3
60-965 Poznań, Poland

Paper

Design of a Modified Interleaving Algorithm Based on Golden Section Theory Enhancing the Performance of Turbo Codes

Essedik Iftene[1,2], Adda Ali-Pacha[1], and Lahcène Hadj Abderrahmane[2]

[1] *Laboratoire de codage et de la sécurité de l'information, LACOSI, Département d'Electronique, Université des Sciences et de la Technologie d'Oran Mohamed-Boudiaf, USTO-MB, Oran, Algeria*
[2] *Algerian Space Agency (ASAL), Centre of Satellites Development, Oran, Algeria*

https://doi.org/10.26636/jtit.2022.158721

Abstract—This paper investigates the design of a modified matrix interleaving algorithm as a way to improve the performance of turbo codes. This proposed solution, known as the matrix-dithered golden (MDG) interleaver, utilizes the characteristics of a matrix interleaver combined with the golden section theory. The performance of the proposed interleaving method is compared with that of matrix (M), random (R), and dithered golden (DG) interleavers. The comparison is made in terms of bit error rate (BER), frame error rate (FER), computational complexity, and storage memory requirement. The turbo coded system is implemented and simulated using Matlab/Simulink software. Results of simulations performed both in the additive white Gaussian noise (AWGN) channel and the Rayleigh fading channel demonstrate the effectiveness of the proposed interleaver. The MDG interleaver is an effective replacement for random interleavers, as it improves BER and FER performance of the turbo code and is also capable of reducing the storage memory requirement without increasing the system's complexity.

Keywords*—*AWGN channel, golden section theory, interleaver, iterative decoding, Rayleigh fading channel, turbo code.

1. Introduction

Since their introduction in 1993 [1], turbo codes have received considerable attention and are currently the subject of extensive research [2], [3]. This is not only because of their powerful error correcting capability, but also because of their flexibility in terms of providing different block sizes and code rates [4]. A turbo code encoder consists of a parallel concatenation of two recursive systematic convolutional (RSC) encoders separated by an interleaver [5], [6]. The interleaver is a device that takes a given sequence of symbols at the input and produces identical symbols at the output, but in a different temporal order. The binary data sequence entering the turbo code's internal interleaver is denoted by d_N, where N is the length the data sequence. The binary data sequence at the output of the turbo code's internal interleaver is denoted by $d_{N-\Delta}$. The corresponding coded data is the binary output X_{3N}. A turbo code decoder employs two cascaded decoding blocks. An iterative

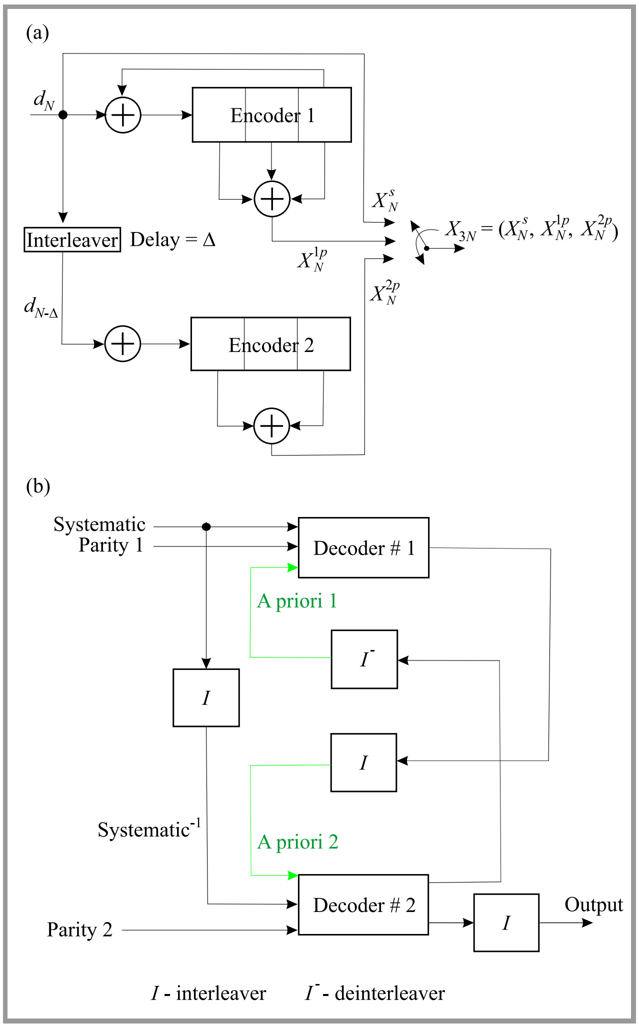


Fig. 1. (a) Turbo encoder structure, (b) turbo decoder structure.

scheme is used for decoding the turbo codes such that the overall performance can be improved [2], [3]. The structures of the turbo encoder and decoder are shown in Fig. 1.

It has been found that one way to improve the performance of a turbo coded system is to use a good interleaver structure [5], [7], [8]. Therefore, interleaver design is the subject of numerous research projects and a number of algorithms have been developed [9]–[20]. It is asserted for turbo codes that interleavers with some randomness tend to perform better than their fully structured counterparts, especially for large block sizes. However, a turbo coded system with a built-in random interleaver suffers from the problem of insufficient flexibility. A change in the length of interleaving requires another search of the interleaving pattern, which implies a more complex implementation. Furthermore, the generated interleavers, characterized by different lengths, should be stored separately in the memory [9], [10]. This causes a serious storage-related concern in a scenario in which many interleaving lengths need to be supported.

This paper introduces a modified architecture for a matrix interleaver, referred to as the matrix-dithered golden (MDG) interleaver which can resolve the problems of the existing random interleavers. The matrix-dithered golden interleaver aims to improve the BER and FER of turbo code, and to minimize the memory requirement by avoiding the need for generating and storing individual interleaving patterns for different interleaving lengths.

The rest of the paper is organized as follows. In Section 2, matrix, random and dithered golden interleavers are briefly explained. In Section 3, we present the designing method of the proposed interleaving algorithm. Simulation results and performance evaluation of these interleavers are provided and discussed in Section 4. The conclusion is presented in Section 5.

2. Interleavers

The interleaving process is a useful technique to enhance the error correcting capability of a turbo code [19]. Thanks to the interleaver, turbo codes can deal with burst errors by converting error patterns that contain long sequences of serial erroneous data into a more random error pattern, thus distributing errors among many code vectors [21]. Turbo codes work much better when errors in the received sequence are spread far apart [2]. An interleaver is used to randomize the error locations by taking a given sequence of symbols, and permutes their positions in a different temporal order [22], [23]. The inverse of the interleaving process is called deinterleaving and restores the interleaved sequence.

In general, we can classify interleavers into two broad categories [5], [7], [8]: random and deterministic interleavers. For deterministic interleavers, the position of every data bit is known according to an algorithm, while for random interleavers the position of each data bit is random. Some useful interleavers used in turbo code are discussed below.

The matrix interleaver [7], [11], [22] is one of the simplest types that is most commonly used in communication systems. This type of interleaver is easy to implement in practice and is characterized by a process in which data is permuted by being written row-wise and read column-wise. The matrix interleaver may have a good minimum distance, but the high multiplicity of low-weight code words makes this interleaver unsuitable [23], [24].

The interleaver with random properties is one of the essential building blocks of turbo codes [7], [18], [23]. Such an interleaver generates a random mapping between the input and output positions. Once the symbols are introduced into a random interleaver, the output symbols are chosen randomly, so that the same symbol that has already been selected is not repeated [21]. As the selection is random, it will be impossible to know the symbol positions at the interleaver output. Therefore, it would be necessary to maintain a correspondence table showing the dependence between the old and the new positions of the interleaved symbols, so that they can be deinterleaved [18], [24]. The random interleaver requires N indexes to be stored in order to implement an interleaver of a length of N. The fundamental concept of a random interleaver is simple, but its practical realization is more complex than that of a matrix interleaver [17], [23].

The golden section has applications in many mathematical problems [25]. It has been used for designing interleavers in turbo codes and is characterized by good proprieties [26]. Golden interleavers are based on sorting real-valued numbers derived from the golden section. Figure 2 illustrates the golden section principle.

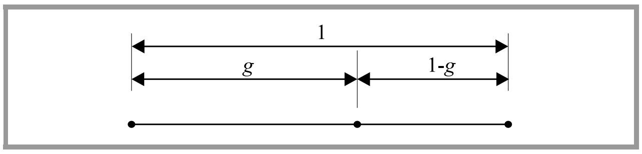


Fig. 2. Golden section principle.

For a given line segment of length 1, the problem is to divide it into a long segment of length g, and a shorter segment of length $1-g$, such that:

$$\frac{g}{1} = \frac{1-g}{g} \ .$$

Using this principle, the golden section value is calculated as $g \approx 0.618$ [8], [12], [22]. The first step in calculating the interleaver indexes is to compute the golden section value g. The second step is to compute the real increment value C, as:

$$C = \frac{N(g^m + j)}{r} \ , \tag{1}$$

where N is the length of the data sequence, m is a preselected non-zero positive integer preferably set to 1 or 2, r is a preselected non-zero integer defining a spacing between any pair of input elements that are to be maximally spread, and j is a preselected integer modulo r. In a typical implementation, j is set to 0, and r is set to 1 [8], [22], [26]. The third step is to generate a real-valued dithered golden vector v. The elements of v are calculated as:

$$v(n) = [s + n \times C + \mathbf{d}(n)] \mod N, \quad \text{for } n = 1 \text{ to } N \ . \tag{2}$$

In Eq. (2), s is any real starting value and d is a dither vector. The starting value s is usually set to 0, but other real values can be selected. The dither vector is uniformly distributed between 0 and $N \times D$, where D is the normalized width of the dither distribution $\mathbf{d}(n)$ and is set to 0.01, according to [8], [14], [26]. The next step is to sort the dithered golden vector v and find the index vector $\mathbf{Z}$ that defines this sort, i.e. to find a sort vector $\mathbf{Z}$ such that $a(n) = v[\mathbf{Z}(n)]$ for $n = 1$ to N, where $\mathbf{a} = \text{sort}(\mathbf{v})$. The dithered golden interleaver indexes are then given by $\alpha[\mathbf{Z}(n)] = n$, for $n = 1$ to N. In fact, vector $\mathbf{Z}$ is the inverse interleaver for α. The dithered golden interleaver requires the use of index memory for storing precomputed indexes. If the full indexes are stored, then the index memory can be excessive.

3. The Matrix-Dithered Golden Interleaving Algorithm

The process of designing a matrix-dithered golden (MDG) interleaver is performed according to the flowchart shown in Fig. 3. The interleaving design comprises four steps:

Step 1

Prepare the golden section model with its control parameters: D, s, m, r, and j. The values of the control parameters of the golden section model adopted in this paper are: $s = 0$, $m = 1$, $D = 0.1$, $j = 9$, and $r = 15$.

Using the defined golden section model, generate two dither vectors sequences $\mathbf{d}_{row}$ and $\mathbf{d}_{column}$ of real numbers with their length equal to the length of the largest frame N:

$$\mathbf{d}_{row}(n) = \{d_{r1}, d_{r2}, \dots, d_{rN}\}, \quad \mathbf{d}_{row}(n) \in [1, N \times D],\ n \in [1, N]\ , \tag{3}$$

$$\mathbf{d}_{column}(n) = \{d_{c1}, d_{c2}, \dots, d_{cN}\}, \quad \mathbf{d}_{column}(n) \in [1, N \times D],\ n \in [1, N]\ . \tag{4}$$

Step 2

Start with the conventional matrix interleaver M. Find an appropriate number of rows $\mathbf{N}_r$ and determine the number of columns N_c for a particular frame size N. The range of the input frame size is divided into two sub-blocks and each sub-block has a different row number and a different column number given by:

$$\begin{aligned} N_r &= \text{floor}(\sqrt{N}), \quad \text{if} \quad N < 512\ , \\ N_r &= \text{floor}(\sqrt[3]{N}), \quad \text{if} \quad N \geq 512\ , \\ N_c &= \text{ceil}\frac{N}{N_r}\ . \end{aligned} \tag{5}$$

Write, in a row-wise fashion, left to right, and starting with the top row, the input data $\mathbf{D}_{in}$ into a matrix $\mathbf{M}$ with N_r rows and N_c columns.

$$\mathbf{M}(i,:) = \mathbf{D}_{in}\left[(i-1) \times N_c + 1 : i \times N_c\right]\ , \quad \text{for } i = 1 \text{ to } N\ . \tag{6}$$

Write the vector $\boldsymbol{d_{row}}$ inside a matrix $\boldsymbol{M_{d_{row}}}$ having N_r rows and N_c columns to obtain N_r different $\boldsymbol{d_{row}}$ vectors, each with length N_c, and write the vector $\boldsymbol{d_{column}}$ inside a matrix $\boldsymbol{M_{d_{column}}}$, having N_r rows and N_c columns to obtain N_c different $\boldsymbol{d_{column}}$ vectors, each with length N_r.

Step 3

Using Eq. (7), generate the dithered golden matrix $\mathbf{v}_{row}$, and order each row according to its magnitude, to form the intra-row permutation matrix $\mathbf{Z}_{row}$. Indexes matrix $\mathbf{Z}_{row}$ and matrix $\mathbf{a}_{row}$, which is the sorted version of matrix $\mathbf{v}_{row}$, are related as Eq. (8):

$$\mathbf{v}_{row}(i,:) = \left[s + i \times C_{row} + \mathbf{M}_{\mathbf{d}_{row}}(i,:)\right] \bmod N_r, \quad \text{for } i = 1 \text{ to } N_r\ , \tag{7}$$

$$\mathbf{a}_{row}(i,:) = \mathbf{v}_{row}\left[\mathbf{Z}_{row}(i,:)\right], \quad \text{for } i = 1 \text{ to } N_r\ . \tag{8}$$

Perform the intra-row permutations of matrix $\mathbf{M}$, based on the constructed intra-row permutation pattern $\mathbf{Z}_{row}$.

$$\mathbf{M}_{row}(i,j) = \mathbf{M}\left[i, \mathbf{Z}_{row}(i,j)\right], \quad \text{for } i = 1 \text{ to } N_r,\ j = 1 \text{ to } N_c\ . \tag{9}$$

Step 4

Similarly, to the step 3 and using Eq. (10), generate another dithered golden matrix $\boldsymbol{v_{column}}$, and order each column of this matrix according to their magnitude, to form the intra-column permutation matrix $\mathbf{Z_{column}}$. Indexes matrix $\mathbf{Z_{column}}$ and the matrix $\boldsymbol{\alpha_{column}}$, which is the sorted version of the matrix $\boldsymbol{v_{column}}$, are related as Eq. (11):

$$\boldsymbol{v_{column}}(:,j) = \left[s + j \times C_{column} + \boldsymbol{M_{d_{column}}}(:,j)\right] \bmod N_c, \quad \text{for } j = 1 \text{ to } N_c\ , \tag{10}$$

$$\boldsymbol{\alpha_{column}}(:,j) = \boldsymbol{v_{column}}\left[\mathbf{Z_{column}}(:,j)\right], \quad \text{for } j = 1 \text{ to } N_C\ . \tag{11}$$

Perform the intra-column permutations of matrix $\mathbf{M}_{row}$ obtained in step 3, based on the constructed intra-column permutation pattern $\mathbf{Z}_{column}$:

$$\mathbf{M}_{column}(i,j) = \mathbf{M}_{row}\left[\mathbf{Z}_{column}(i,j), j\right], \quad \text{for } i = 1 \text{ to } N_r,\ j = 1 \text{ to } N_c\ . \tag{12}$$

Finally, the entire data block is read from the permuted matrix $\mathbf{M}_{column}$, column-wise, top to bottom, starting with the left column:

$$\mathbf{D}_{out}\left[(j-1) \times N_r + 1 : j \times N_r\right] = \mathbf{M}_{column}(:,j), \quad \text{for } j = 1 \text{ to } N_c\ . \tag{13}$$

4. Comparative Performance Analysis of Interleavers

To verify the effectiveness of the proposed interleaving approach, comparisons to matrix (M), random (R) and dithered golden (DG) interleavers have been made based on such parameters as complexity, BER, FER, and memory usage. These interleavers were introduced into an unpunctured turbo code at the rate of 1/3, in which two identi-

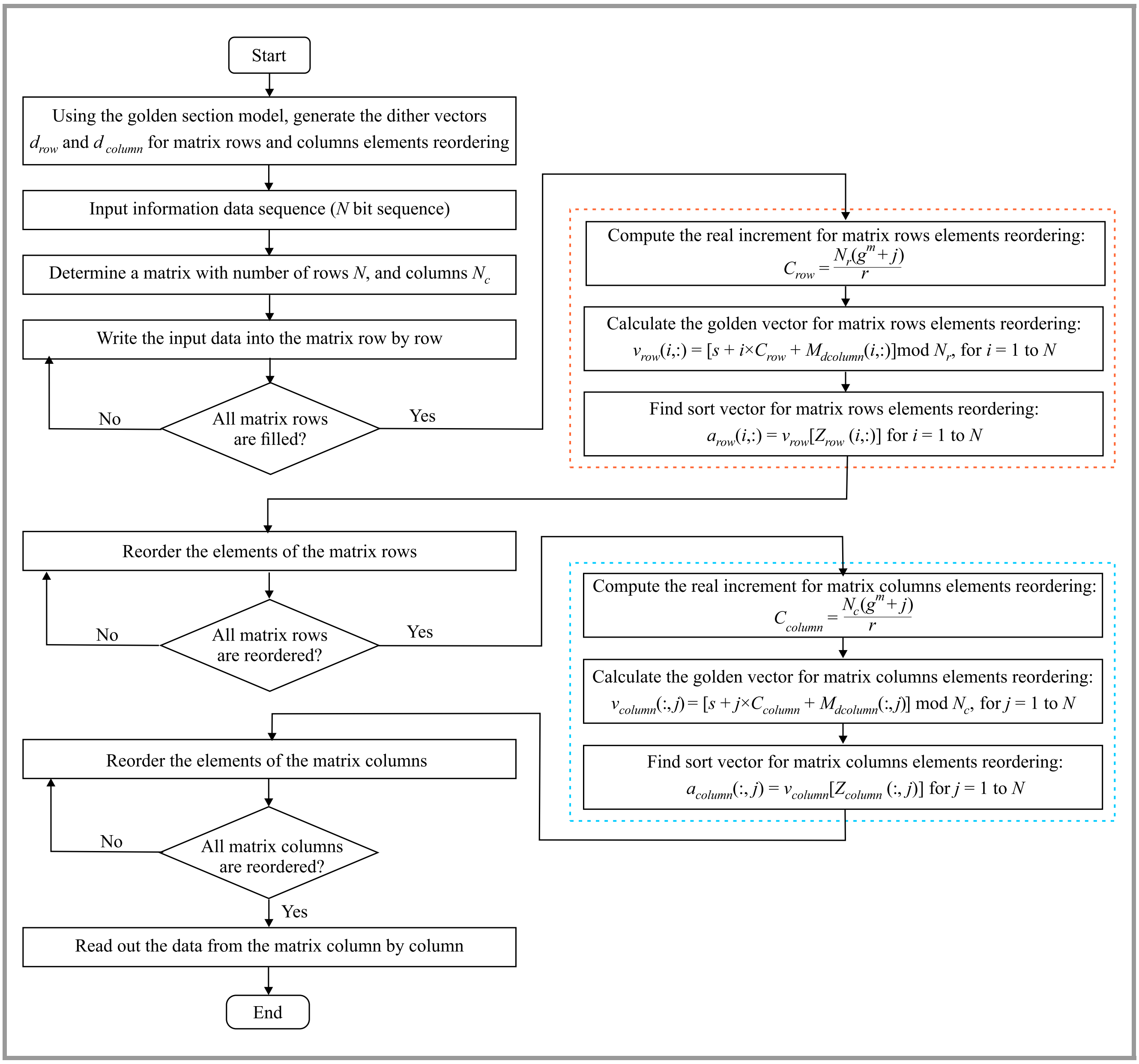

Start
Using the golden section model, generate the dither vectors d_{row} and d_{column} for matrix rows and columns elements reordering
Input information data sequence (N bit sequence)
Determine a matrix with number of rows N_r and columns N_c
Write the input data into the matrix row by row
No
All matrix rows are filled?
Yes
Compute the real increment for matrix rows elements reordering:
$C_{row} = \frac{N_r(g^m + j)}{r}$
Calculate the golden vector for matrix rows elements reordering:
$v_{row}(i,:) = [s + i \times C_{row} + M_{dcolumn}(i,:)] \bmod N_r$, for $i = 1$ to N
Find sort vector for matrix rows elements reordering:
$a_{row}(i,:) = v_{row}[Z_{row}(i,:)]$ for $i = 1$ to N
Reorder the elements of the matrix rows
No
All matrix rows are reordered?
Yes
Compute the real increment for matrix columns elements reordering:
$C_{column} = \frac{N_c(g^m + j)}{r}$
Calculate the golden vector for matrix columns elements reordering:
$v_{column}(:,j) = [s + j \times C_{column} + M_{dcolumn}(:,j)] \bmod N_c$, for $j = 1$ to N
Find sort vector for matrix columns elements reordering:
$a_{column}(:,j) = v_{column}[Z_{column}(:,j)]$ for $j = 1$ to N
Reorder the elements of the matrix columns
No
All matrix columns are reordered?
Yes
Read out the data from the matrix column by column
End

Fig. 3. Flowchart of the MDG interleaver algorithm.

cal recursive systematic convolutional encoders of generator polynomials $(7,5)_{oct}$, having the constraint length, $K = 3$, are connected in parallel [6], [7], [14], [26]. The turbo coded system is implemented and simulated using Matlab/Simulink software. All simulation results presented are based on iterative decoding using the maximum a posteriori (MAP) algorithm [2]. Turbo decoders suffer from high decoding latency due to the iterative decoding process. Latency can be lowered by reducing the number of required decoding iterations. Hence, the number of iterations in the decoder is selected to equal 7. Two noise models were considered: AWGN channel and Rayleigh fading channel. The data length is taken as 400 and 1024 bits. For each SNR value, the simulation stops after having counted at least 60 error frames. The trellis termination is applied to both RSC component encoders.

The normalized width of the dither distribution D and other parameters, such as m, r, and j is an important design parameter for generating the dithered golden vector. Hence, in this paper, we perform a search for the best values of these design parameters, by using the BER and FER as a measure of quality.

Figure 4 shows the influence of design parameters D, j, r, m, and the number of matrix row N_r on the performance of the matrix-dithered golden interleaver in a Rayleigh fading channel, for two interleaving sizes $N = 400$ and $N = 1024$.

Performance of the matrix-dithered golden interleaver depends on the choice of the design parameters. The result shows that for a small frame length of approximately 400 bits, the best results are obtained by selecting the number of rows of the interleaving matrix to be $N_r = \text{floor}(\sqrt{N})$. For large frame lengths, i.e. of 1024 bits, the

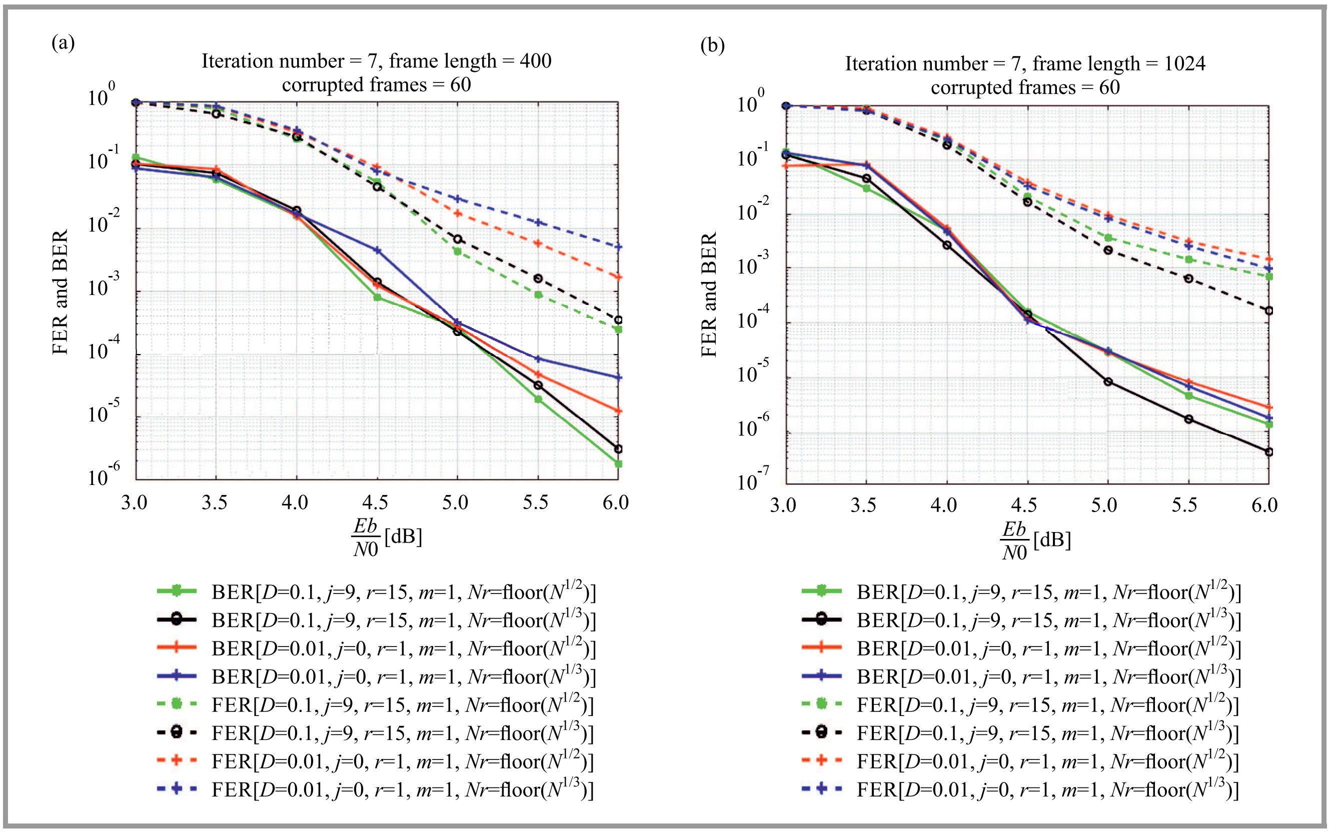


Fig. 4. Influence of the MDG interleaver design parameters on the BER and FER performance in the Rayleigh fading channel: (a) frame length $N = 400$ and (b) frame length $N = 1024$.

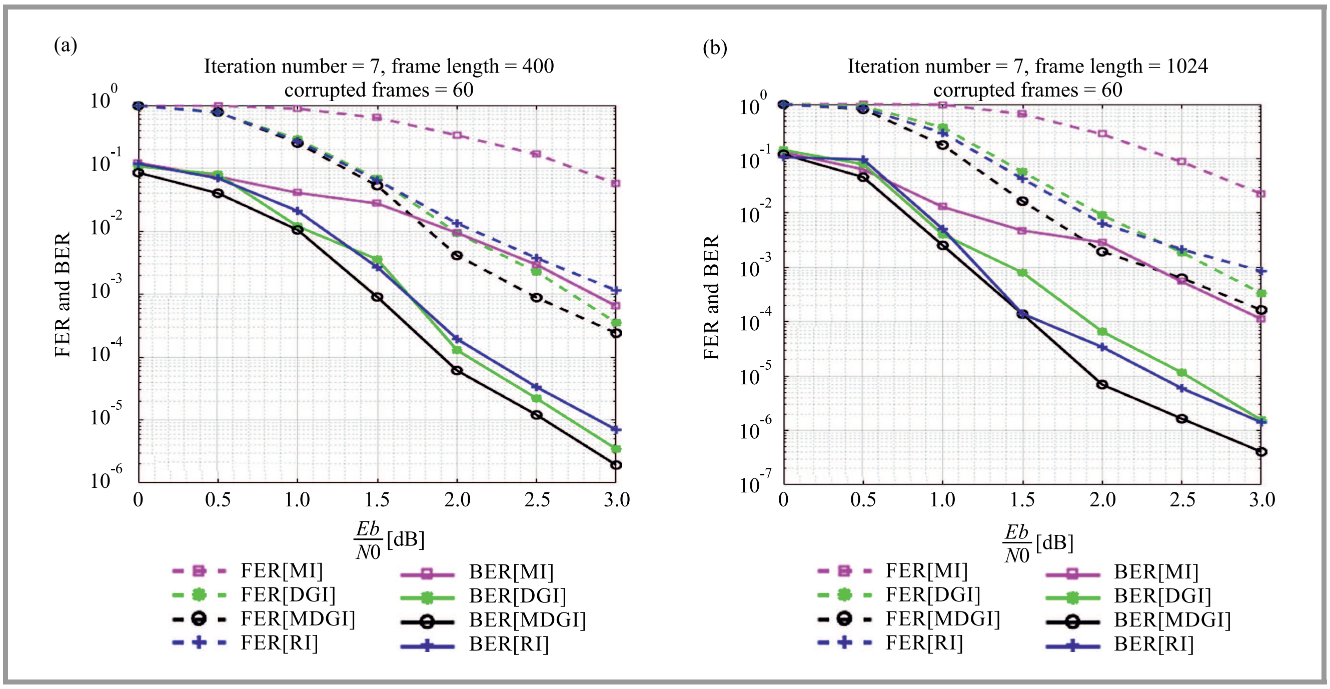


Fig. 5. BER and FER performance comparison between interleavers in the AWGN channel: (a) frame length $N = 400$ and (b) frame length $N = 1024$.

selection of the number of rows as $N_r = \text{floor}(\sqrt[3]{N})$, improves the interleaver's performance. The result also shows, that for any frame size, the best BER and FER performances is obtained for matrix-dithered golden interleaver with design parameters are set as the normalized width of the dither distribution $D = 0.1$, $j = 15$, $r = 9$, and $m = 1$.

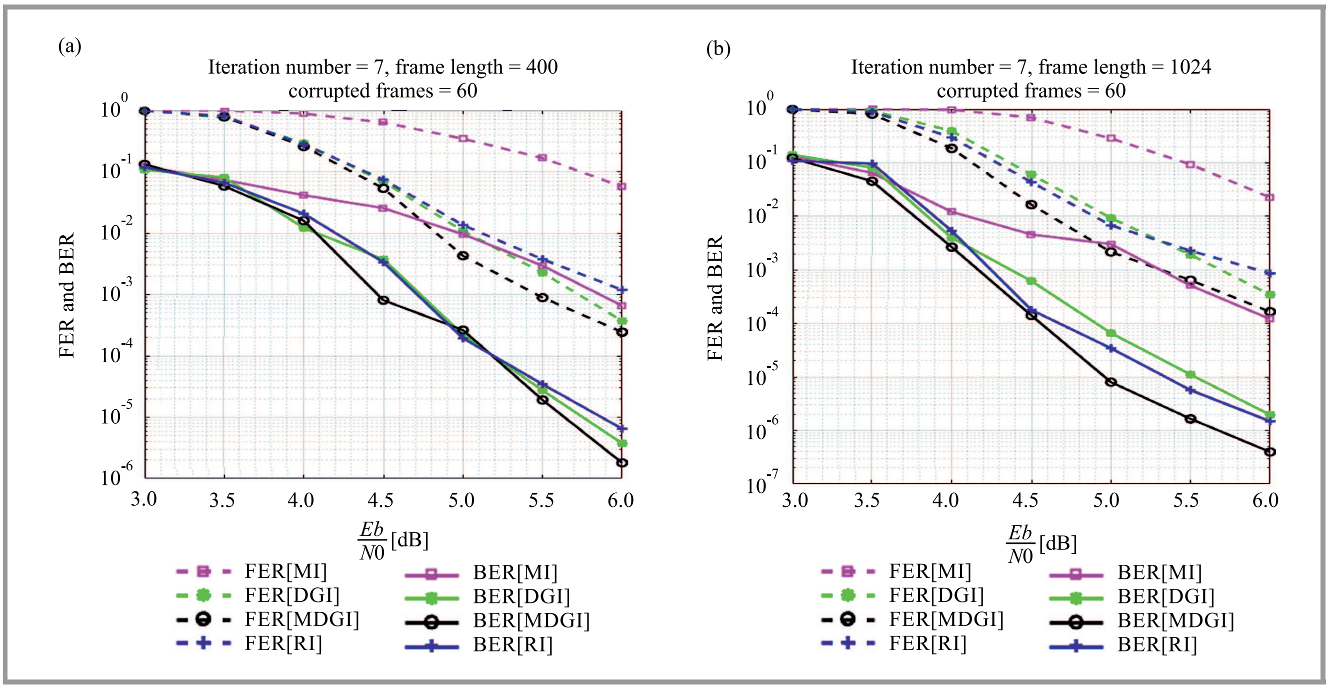


Fig. 6. BER and FER performance comparison between interleavers in the Rayleigh channel: (a) frame length $N = 400$ and (b) frame length $N = 1024$.

Table 1
Comparison of the computational complexity in terms of the number of cycles required to obtain the interleaving pattern

Interleaver length	Random interleaver	Dithered golden interleaver	Matrix interleaver	Matrix-dithered golden interleaver
$N = N_r \times N_c$	N	N	$2 \times (N_r \times N_c)$	$2 \times (N_r + N_c)$

Figures 5 and 6 show the comparisons of BER and FER performance in AWGN channel and Rayleigh fading channels, respectively. A turbo code with the interleaving length of $N = 400$ and 1024, and decoding iterations of 7 is tested in this comparison. The following algorithm design parameters are used: $D = 0.1$, $j = 9$, $r = 15$, $m = 1$, with the frame lengths equaling $N = 400$ and $N = 1024$.

Figures 5–6 show that almost the same turbo code behavior is recorded both in AWGN and Rayleigh fading channels. However, there is a fall in performance recorded in the Rayleigh fading channel, compared to AWGN. This loss equals approximately 3 dB. It is also observed that the performance of the turbo code improves significantly as the interleaving length increases, and that the MDG interleaver exhibits better BER and FER performance than other interleaver schemes for all SNRs. Considering that an interleaving length of 1024 bits and the AWGN channel are used, the results show that performance of the matrix-dithered golden interleaver is approximately 0.4 dB better than that of the random interleaver, and 0.6 dB better than that of the dithered golden interleaver at BER of 10^{-5}. A more significant gain is obtained relative to the performance of the matrix interleaver. The results show that, at an SNR value of 3 dB, the matrix interleaver has a BER of 10^{-4}, whereas the proposed interleaver has a BER of only $4 \times 10^{-7.}$

In Table 1, computational complexity of the different interleaving algorithms discussed in this work is presented. Here, complexity means the number of cycles required for the generation of interleaving patterns. The MDG interleaving scheme is extremely efficient in reducing computational complexity, compared to random, matrix, and dithered-golden interleaving schemes. By using the MDG interleaver, one may interleave a block of N_r rows and N_c columns in $2 \times (N_r + N_c)$ cycles, since only one cycle per row or column is needed. Performance is significantly improved compared to the traditional implementation which needs $2 \times (N_r \times N_c)$ cycles.

The memory requirement for different interleavers is shown in Table 2. The values are calculated based on the number of interleaving patterns to be stored, as a function of number of interleaving lengths n that need to be supported by the interleaver. The frame length is represented as N.

The results show that in the case of the random interleaver and the dithered golden interleaver, the memory size required for storing interleaving patterns depends on the number of interleaving lengths. Therefore, storage memory becomes large if multiple frame lengths have to be supported by the interleaving algorithm. However, the memory requirement of the matrix-dithered golden interleaver is independent of the number of interleaving lengths, as in this case, only one interleaving pattern, generated for the largest interleaving length, is to be stored instead of storing all interleaver patterns generated for different interleaving lengths. The slightly increased memory requirement of the

Table 2
Comparison of memory requirements of different interleaving algorithms

Interleaver	Memory requirement
Random interleaver	$n \times N$
Dithered golden interleaver	$n \times N$
Matrix interleaver	N
Matrix-dithered golden interleaver	$3 \times N$

MDG interleaver, compared with the matrix interleaver, is related to the calculation and storage of the dithered golden matrices for intra-row and intra-column permutations. Because the proposed approach offers better BER and FER performance than the matrix interleaver, such a slight additional memory requirement is acceptable.

5. Conclusion

Based on the golden section theory, a modified architecture for a matrix interleaving scheme referred to as the matrix-dithered golden (MDG) interleaver, has been suggested in this paper to further improve the performance of a turbo-coded system.

It was concluded that the proposed interleaving method improves BER and FER performance of the turbo codes. Compared with the random interleaver and the dithered golden interleaver, the MDG interleaver reduces computational complexity and storage memory requirements, as only one interleaving pattern needs to be generated and stored. The increased memory requirement of the MDG interleaver, compared with the matrix interleaver, is related to the calculation and storage of the dithered golden matrices for intra-row and intra-column permutations. Because the proposed approach is characterized by lower computational complexity, as well as by better BER and FER performance compared with the matrix interleaver, the slight additional memory requirement is acceptable.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1", in *Proc. of IEEE Int. Conf. on Commun. ICC'93*, Geneva, Switzerland, vol. 2, pp. 1064–1070, 1993 (DOI: 10.1109/ICC.1993.397441).

[2] D. Sah, "Iterative decoding of turbo codes", *J. of Adv. College of Engin. and Manag.*, vol. 3, pp. 15–30, 2017 (DOI: 10.3126/jacem.v3i0.18810).

[3] S. Jasim and A. Abbas, "Performance of turbo code with different parameters", *J. of Univer. of Babylon*, vol. 25, no. 5, pp. 1684–1692, 2017 [Online]. Available: https://iasj.net/iasj/pdf/42c4640e163c6f20

[4] R. M. Deshmukh and S. A. Ladhake, "Analysis of various puncturing patterns and code rates: Turbo code", *Int. J. Electron. Engin. Res.*, vol. 1, no. 2, pp. 79–88, 2009 [Online]. Available: https://www.idc-online.com/technical_references/pdfs/data_communications/Analysis%20of%20Various%20Puncturing%20Patterns.pdf

[5] Y. J. Harbi, "Effect of the interleaver types on the performance of the parallel concatenation convolutional codes", *Int. J. of Elec. and Comp. Sci. IJECSIJENS*, vol. 12, no. 3, pp. 25–31, 2012 [Online]. Available: http://ijens.org/Vol_12_I_03/128703-5858-IJECS-IJENS.pdf

[6] M. K. Gupta and V. Sharma, "To improve bit error rate of turbo coded OFDM transmission over noisy channel", *J. of Theoret. and Appl. Inform. Technol.*, vol. 8, no. 2, pp. 162–168, 2009 [Online]. Available: http://www.jatit.org/volumes/research-papers/Vol8No2/9Vol8No2.pdf

[7] M. Synthia and M. S. Ali, "Performance study of turbo code with interleaver design", *Int. J. of Scient. & Engin. Res.*, vol. 2, no. 7, pp. 1–5, 2011 [Online]. Available: https://www.ijser.org/onlineResearchPaperViewer.aspx?Performance_Study_of_Turbo_Code_with_Interleaver_Design.pdf

[8] A. S. Hadi, "Performance of turbo-codes with some proposed interleaver schemes", *Al-Khwarizmi Engin. J.*, vol. 2, no. 2, pp. 1–4, 2006 [Online]. Available: https://iasj.net/iasj/download/ee3c4595b8142501

[9] M. Salim, R. P. Yadav, K. Narwal, and A. Sharma, "A new block S-random interleaver for shorter length frames for turbo codes", *Bull. of Elec. Engin. and Inform.*, vol. 2, no. 4, pp. 293–298, 2013 (DOI: 10.11591/eei.v2i4.196).

[10] M. A. Khan, V. Jeoti, and R. S. Manzoor, "Performance evaluation of seed based random (SBR) interleaver in Rayleigh fading channel", in *Proc. of IEEE 4th Int. Conf. on Intell. and Adv. Syst. ICIAS 2012*, Kuala Lumpur, Malaysia, pp. 311–313, 2012 (DOI: 10.1109/ICIAS.2012.6306208).

[11] A. E. Hassan, M. Shokair, A. A. Elazm, D. Truhachev, and C. Schlegel, "Proposed deterministic interleavers for CCSDS turbo code standard", *J. of Theoret. & Appl. Inform. Technol.*, vol. 16, pp. 29–33, 2010 [Online]. Available: http://www.jatit.org/volumes/research-papers/Vol16No1/4Vol16No1.pdf

[12] X. Zou, M. Wang, and G. Feng, "A new interleaver design for iteratively decoded bit-interleaved coded modulation", *Int. J. of Soft Comput.*, vol. 3, pp. 338–343, 2008 [Online]. Available: http://docsdrive.com/pdfs/medwelljournals/ijscomp/2008/338-343.pdf

[13] O. Y. Takeshita and D. J. Costello, "New deterministic interleaver designs for turbo codes", *IEEE Trans. on Inform. Theory*, vol. 46, no. 6, pp. 1988–2006, 2006 (DOI: 10.1109/18.868474).

[14] L. H. Abderrahmane, M. Bacha, and A. Mebrek, "A new optimised interleaver structure for turbo coding", in *Proc. of 27th Canadian Conf. on Elec. and Comp. Engin. CCECE 2014*, Toronto, ON, Canada, 2014 (DOI: 10.1109/CCECE.2014.6900945).

[15] D. Wang and H. Kobayashi, "On design of interleavers with practical size for turbo codes", in *Proc. of IEEE Int. Conf. on Commun. ICC 2000. Global Convergence Through Commun.*, New Orleans, LA, USA, vol. 2, pp. 618–622, 2000 (DOI: 10.1109/ICC.2000.853570).

[16] H. R. Sadjadpour, N. J. Sloane, M. Salehi, and G. Nebe, "Interleaver design for turbo codes", *IEEE J. on Selec. Areas in Commun.*, vol. 19, no. 5, pp. 831–837, 2001 (DOI: 10.1109/49.924867).

[17] C. J. Corrada-Bravo and I. Rubio, "Deterministic interleavers for turbo codes with random-like performance and simple implementation", in *Proc. of 3rd Int. Symp. on Turbo Codes*, Brest, France, 2003 [Online]. Available: http://ccom.uprrp.edu/~labemmy/Wordpress/wp-content/uploads/2010/11/Deterministic-Interleavers-for-Turbo-Codes-with-Random-like-Performance-and-Simple-Implementation.pdf

[18] M. Kovaci, H. G. Balta, and M. M. Nafornita, "The performances of interleavers used in turbo codes", in *Proc. of Int. Symp. on Signals, Circuits and Syst. ISSCS 2005*, Iasi, Romania, vol. 1, pp. 363–366, 2005 (DOI: 10.1109/ISSCS.2005.1509929).

[19] C. C. Bravo and I. Rubio, "Algebraic construction of interleavers using permutation monomials", in *Proc. of IEEE Int. Conf. on Commun.*, Paris, France, vol. 2, pp. 911–915, 2004 (DOI: 10.1109/ICC.2004.1312634).

[20] D. S. Kim, H. Y. Oh, and H. Y. Song, "Collision-free interleaver composed of a latin square for parallel-architecture turbo codes", *IEEE Commun. Lett.*, vol. 12, no. 3, pp. 203–205, 2008 (DOI: 10.1109/LCOMM.2008.071423).

[21] G. A. A. Kareem, "Efficient selection of circular and random interleaver for turbo codes multiuser CDMA system with Hadamard spreading code in AWGN and Rayleigh fading channels", *J. of Engin. and Sustainable Develop.*, vol. 16, no. 2, pp. 117–132, 2012 [Online]. Available: https://www.iasj.net/iasj/pdf/d6335a53589140fe

[22] G. B. Purushottama and B. R. Sujatha, "Turbo codes with golden section interleaver", *Int. Res. J. of Engin. and Technol.*, vol. 2, no. 3, pp. 1069–1073, 2015 [Online]. Available: https://www.irjet.net/archives/V2/i3/Irjet-v2i3158.pdf

[23] P. D. Bahirgonde and S. K. Dixit, "BER analysis of turbo code interleaver", *Int. J. of Comp. Appl.*, vol. 126, no. 14, pp. 1–4, 2015 (DOI: 10.5120/ijca2015906278).

[24] B. Raje and K. Markam, "Review paper on study of various Interleavers and their significance", *Int. Res. J. of Engin. and Technol.*, vol. 5, no. 11, pp. 430–434, 2018 [Online]. Available: https://www.irjet.net/archives/V5/i10/IRJET-V5I1082.pdf

[25] S. Crozier, J. Lodge, P. Guinand, and A. Hunt, "Performance of turbo-codes with relative prime and golden interleaving strategies", in *Proc. of the 6th Int. Mobile Satellite Conf. IMSC'99*, Ottawa, Canada, 1999 pp. 268–275.

[26] L. H. Abderrahmane and S. Chellali, "Performance comparison between Gaussian interleaver, Rayleigh interleaver, and dithered golden interleaver", *Ann. of Telecommun.*, vol. 63, no. 7–8, pp. 449–452, 2008 (DOI: 10.1007/s12243-008-0037-2).

Essedik Iftene received his B.Sc. degree in Electronics from University of Sciences and Technologies, Algeria, in 2008, and M.Sc. in Space Communication Systems from High Institute of Aeronautics and Space, Toulouse, France, in 2010. He also earned an M.Sc. in Space Technologies and Applications from University of Sciences and Technologies, Algeria. Currently, he is a research assistant at the Department of Space Instrument Research at the Satellite Development Center (CDS), Algeria, and is pursuing a Ph.D. in Space Instrument Design at University of Sciences and Technologies, Algeria. His research interests include wireless networks, channel coding and modeling, and their application in space communication systems.

https://orcid.org/0000-0003-0995-5850
E-mail: essedik.iftene@univ-usto.dz
Laboratoire de codage et de la sécurité
de l'information (LACOSI)
Département d'Electronique
Université des Sciences et de la Technologie
d'Oran Mohamed-Boudiaf USTO-MB
BP 1505, El M'naouer
Bir El Djir 31000 Oran, Algeria

Algerian Space Agency (ASAL)
Centre of Satellites Development
Ibn Rochd USTO Oran BP 4065, Algeria

Adda Ali-Pacha majored in Telecommunications Engineering in January 1986. He received a B.Sc. in Mathematics in 1986 and an M.Sc. in Signal Processing in 1993. He also obtained a Ph.D. in Data Security in 2004. He worked for a telecommunications authority (PTT Oran), holding the position of the Head of Telephone Traffic for two years (1986-1988). He has been working for University of Sciences and Technology of Oran (USTO) Algeria since 1989, serving as an academic teacher/researcher in the Electronics Institute. His research focuses on telecommunication domains.

E-mail: adda.alipacha@univ-usto.dz
Laboratoire de codage et de la sécurité
de l'information (LACOSI)
Département d'Electronique
Université des Sciences et de la Technologie
d'Oran Mohamed-Boudiaf USTO-MB
BP 1505, El M'naouer
Bir El Djir 31000 Oran, Algeria

Lahcène Hadj Abderrahmane is a researcher at the Satellite Development Center. Currently, he works on RF systems and channel coding, focusing primarily on designing satellite communication systems. He is an RF team leader at the Department of Space Instrument Research and an Assistant Professor teaching telecommunications to Master's degree students.

E-mail: hadjabderrahmanel@yahoo.fr
Algerian Space Agency (ASAL)
Centre of Satellites Development
Ibn Rochd USTO Oran BP 4065, Algeria

Using of Golden Code Orthogonal Super-Symbol in Media-Based Modulation for Single-Input Multiple-Output Schemes

Ayodeji James Bamisaye and Tahmid Quazi

Department of Electrical, Electronic and Computer Engineering, University of Kwazulu-Natal, Durban, South Africa

https://doi.org/10.26636/jtit.2022.158921

Abstract—The media-based modulation (MBM) scheme is capable of providing high throughput, increasing spectrum efficiency, and enhancing bit error rate (BER) performance of communication systems. In this paper, an MBM employing radio frequency (RF) mirrors and golden code is investigated in a single-input multiple-output (GC-SIMO) application. The aim is to reduce complexity of the system, maximize linear relationships between RF mirrors and improve spectral efficiency of MBM to in order to obtain a high data rate with the use of less hardware. Orthogonal pairs of the super-symbol in the GC scheme's encoder are employed, transmitted via different RF mirrors at different time slots in order to achieve the full data rate and high diversity. In the results having BER of 10^{-5}, the GC-SIMO, MBM exhibits better performance than GD-SIMO, with the gain of approximately 7 dB and 6.5 dB SNR for 4 b/s/Hz and 6 b/s/Hz, respectively. The derived theoretical average error probability of the proposed scheme is validated with the use of the Monte Carlo simulation.

Keywords—golden code, media-based modulation, radio RF mirrors, SIMO.

1. Introduction

The continuous need for higher throughputs in wireless communications has led to increased popularity of multiple-input multiple-output (MIMO) systems which have shown great promise regarding high transmission capacity and improved link reliability and are considered to offer good prospects for modern wireless communications [1]–[6].

MIMO systems split signals into several separated bit streams for a high data rate, via simultaneous transmissions of information to multiple receivers. A typical example of a MIMO to consider is spatial modulation (SM) [7], a unique MIMO scheme which employs a number of transmit antennas for high data rate communication.

The basic idea of spatial modulation (SM) is to convey information using both amplitude/phase modulation (APM) and the transmit antenna index. For example, a conventional MIMO system with 4 transmit antennas and 4-QAM modulation yields a data rate of 2 b/s/Hz, while the same configuration in an SM system shows a data rate of 4 b/s/Hz. Similarly, the involvement of a single radio frequency (RF) chain in SM improves the eliminates the setbacks experienced in a conventional MIMO system, such as, for example, inter-antenna synchronization (IAS) and inter-channel interference (ICI) [7]–[11]. Improvements to SM schemes are currently receiving much attention in terms of research focusing on MIMO systems [12], [13].

Spatial diversity is a technique to improve link reliability through the channel's frequency, time, and space variations, with numerous copies of data received at the receiver side. An example to consider is space-time block code (STBC) transmit diversity scheme [14]–[16] which employs the precoding technique to send multiple packets of data from to the group of transmit antennas to optimize SNR and transmitting power with a suitable phase and amplitude. Two time slots are required to transmit two symbols – hence, the data rate remains unchanged [14], [15]. There is an improvement in the reliability of the link due to transmitting redundant data packets over an independent channel. Signal space diversity (SSD) [17], [18] is another example that should be mentioned here, as it achieves communication diversity by transmitting the in-phase and quadrature of the rotated multi-dimension signal to the receiver via an independent fading channel. It offers improved link reliability at no additional cost of hardware, bandwidth, and transmit power.

For the next generation of wireless communication systems, energy efficiency, spectrum usage and system complexity are essential for supporting demand related to multimedia services and applications. Such features bring the media-based modulation (MBM) [19]–[21] enhancing transmission data rate.

MBM uses numerous RF mirrors to design complicated fading symbols, even with a single transmit antenna, by positioning a number of RF mirrors near the transmit antenna that broadcasts a tone. The placement of RF mirrors near the transmit antenna is the same technique as the placement of scatterers near the transmitter in the propagation environment. The mirror activation pattern (MAP) can change the radiation properties of each of these scatterers, i.e. RF mirrors. The propagation environment adjacent to the transmitter varies from one MAP to the other. Thanks to minor perturbation in the propagation environment reinforced by many random reflections in a rich scattering environment, an independent channel feature is obtained. By serving as regulated scatterers, RF mirrors produce this type of radio interference, resulting in independent fading characteristics for different MAPs.

2. Related Work

Due to low hardware requirements and performance advantages of MBM, it has recently attracted considerable amounts of research attention [22]–[25] as a promising technique exhibiting greater benefits over the existing index modulation (IM) systems, such as frequency-domain IM (FD-IM) [26], [27], space domain IM (SD-IM), also referred to as spatial modulation (SM) [8], [28], and time-domain IM (TD-IM) [29] [30]. MBM offers better performance also when compared with conventional modulation schemes [14], [21], [23], [31]. Likewise, research focusing on MBM in MIMO and multiuser settings also showed improved performance of MBM [12], [24]. In addition, describing a scenario in which RF mirrors are employed, papers [24], [32]–[34] prove that MBM may further improve link reliability at a lower degree of hardware complexity, compared to other spatial multiplexing techniques. This is due to the linear correlation between spectral efficiency and the number of RF mirrors used, which is achieved by creating different channel fade realizations via the RF mirrors, known as mirror activation patterns (MAPs) [24]. In [33], a SIMO-aided MBM (SIMO-MBM) scheme is used, where the linear correlation between spectral efficiency and the number of RF mirrors m_{rf} reduces the system's complexity. Considering an example of an equivalent SM system, which would require $2^{m_{rf}}$ transmit antennas to achieve the same spectral efficiency as SIMO-MBM, the SIMO-MBM system is more efficient, in terms of data rate and hardware complexity, than SM. In MBM, RF mirrors may be positioned side-by-side, resulting in the received constellation size being independent of the transmit power, while in conventional MIMO schemes, transmit antennas are adequately separated to achieve independent fading. Therefore, a large increase in spectral efficiency is easily realizable in MBM schemes [35], [36].

In papers [37]–[41], golden code (GC) has been introduced as a scheme. It achieves a full rate and full diversity by employing a precoding technique, using different transmit antennas at various time slots based on the idea of space time label diversity. In [38], GC modulation was investigated, with SIMO systems maintaining the same bandwidth efficiency if a pair of super-symbols is transmitted, coupled with an extra diversity gain. Computation complexity (CC) of a GC-SIMO scheme presented in [37] is reduced by transmitting only the orthogonal pairs in the encoder, such that two symbols are transmitted in total. This still achieves an extra diversity gain when compared to the conventional SIMO system.

The MBM technique was also investigated in connection with GC modulation in [42]. The scheme employs four complex symbols to output the super-symbols, called the golden codewords, via the encoder. The super-symbols are transmitted via four independents transmit antennas in different time slots, such that four symbols are transmitted in total, achieving full rate and full diversity. Each pair of these super-symbols is transmitted via multi-active transmit antennas and in a different time slot. However, complexity of the system proposed in [42] is high, which limits its practical application.

When incorporating the MBM technique in the GC-SIMO scheme, only the super-symbol orthogonal pairs are employed, i.e. only two symbols are used in the encoder against four in [42], which allows to achieve a high data rate similar to that from [37], with reduced hardware complexity. This feature serves as our motivation to propose media-based golden codeword modulation for SIMO, referred to as GC-SIMO-MBM.

In this paper, we start by examining RF mirror-based MBM in a GC-SIMO scheme. Next, theoretical considerations of the proposed system are validated by means of Monte Carlo simulation.

Note. Scalar quantities are represented by regular letters, while vectors/matrices are indicated by bold/italic lowercase/uppercase symbols. $\|\cdot\|$ symbolizes the Frobenius norm, $Q(\cdot)$ denotes the Gaussian Q-function, $\underset{w}{\operatorname{argmin}}(\cdot)$ and $\underset{w}{\operatorname{argmax}}(\cdot)$ signify the minimum/maximum rate of an argument with reference to w, the binomial coefficient is represented by $\binom{\cdot}{\cdot}$, i is a complex number and the real component of the complex number is given by $R\{\cdot\}$, $|\cdot|$ signifies the Euclidean norm, $[\cdot]^T$ represent transpose and $\lfloor\cdot\rfloor$ indicates the closest integer to a lesser extent than the input argument.

3. Golden Code

The golden code offers the full rate and full diversity, employing the precoding technique. The golden code encoder employs 4 unique complex symbols to output 4 super-symbols which are transmitted via unique independent transmit antennas in two time slots. The golden codeword matrix is given as [37], [42]:

$$\boldsymbol{X} = \begin{bmatrix} \alpha(x_1 + x_2\theta)\frac{1}{\sqrt{5}} & \gamma\overline{\alpha}(x_3 + x_4\overline{\theta})\frac{1}{\sqrt{5}} \\ \alpha(x_3 + x_4\theta)\frac{1}{\sqrt{5}} & \overline{\alpha}(x_1 + x_2\overline{\theta})\frac{1}{\sqrt{5}} \end{bmatrix},$$

where $\theta = \frac{1+\frac{1}{\sqrt{5}}}{2}$, $\overline{\theta} = 1 - \theta$, $\alpha = 1 + j(1 - \overline{\theta})$ and $\gamma = j$.

Four super-symbols:

$$\alpha(x_1 + x_2\theta)\frac{1}{\sqrt{5}}, \quad \gamma\overline{\alpha}(x_3 + x_4\overline{\theta})\frac{1}{\sqrt{5}},$$
$$\alpha(x_3 + x_4\theta)\frac{1}{\sqrt{5}}, \quad \text{and} \quad (\overline{\alpha}x_1 + x_2\overline{\theta})\frac{1}{\sqrt{5}},$$

are generated and are referred to as the golden codeword. They comprise two pairs of super-symbols:

$$\left\{\alpha(x_1+x_2\theta)\frac{1}{\sqrt{5}}\ ,\ \ \overline{\alpha}(x_1+x_2\overline{\theta})\frac{1}{\sqrt{5}}\right\}$$

and

$$\left\{\alpha(x_3+x_4\theta)\frac{1}{\sqrt{5}}\ ,\ \ \gamma\overline{\alpha}(x_3+x_4\overline{\theta})\frac{1}{\sqrt{5}}\right\}\ .$$

The pair $\left\{\alpha(x_1+x_2\theta)\frac{1}{\sqrt{5}},\ \overline{\alpha}(x_1+x_2\overline{\theta})\frac{1}{\sqrt{5}}\right\}$ is employed for transmission in the proposed system.

4. Proposed GC-SIMO-MBM

The system model of the proposed GC-SIMO-MBM scheme is shown in Fig. 1. Spectral efficiency associated with this scheme is $m = \log_2(M) + m_{rf}$ [b/s/Hz], where M and m_{rf} represents the amplitude/phase modulation (APM) constellation size and the number of RF mirrors at the transmitting unit, respectively.

In the proposed GC-SIMO-MBM, input bit $\log_2(M)$ is fed into mappers Ω_1 and Ω_2 to map $\log_2(M)$ bits onto the constellation points from the signal set of $\alpha(x_1+x_2\theta)\frac{1}{\sqrt{5}}$ – and $\overline{\alpha}(x_1+x_2\overline{\theta})\frac{1}{\sqrt{5}}$ – in the Argand plane, which yields two super-symbols x_q^1 and x_q^2. In addition, m_{rf} bit chooses the RF mirror to be used for transmission. The number of available RF mirrors m_{rf}, yields the mirror activation pattern (MAP), such that $N_m = 2^{m_{rf}}$. For example, if $m_{rf} = 2$, then $N_m = 4$.

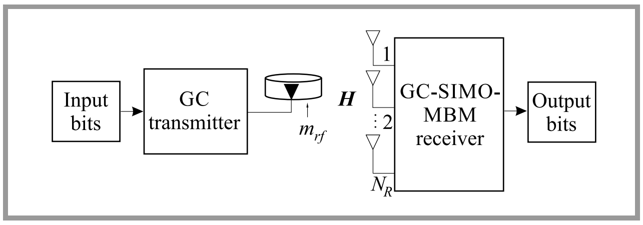


Fig. 1. System model of the proposed GC-SIMO-MBM.

The modulated symbol is conveyed across a channel $\boldsymbol{H}_i$ of magnitude $N_R \times N_m$ in the presence of additive white Gaussian noise (AWGN) $\boldsymbol{n}_i$ of magnitude $N_R \times 1$, e_{ℓ_i} is an $N_m \times 1$ vector. A Rayleigh frequency-flat fading channel is assumed. Therefore, the received signal vector $\boldsymbol{y}_i$ can be written as:

$$\boldsymbol{y}_i = \boldsymbol{H}_i x_q^i \boldsymbol{e}_{\ell_i} + \boldsymbol{n}_i\ , \tag{1}$$

where $i \in [1:2]$, the corresponding transmit antenna employed to transmit the modulated symbol is represented by ℓ_i, while $\boldsymbol{H}_i$ is the i-th column of the channel matrix, which is independent, and identically distributed (i.i.d.) complex Gaussian random variables are distributed as $CN(0,\ 1)$.

Using the maximum likelihood (ML) detector at the receiver, the received signal vector $\boldsymbol{y}_i$ is detected optimally, examining the total signal space of M^2 constellation points combined with all possible transmit antenna index. The ML detector can be defined as:

$$[\hat{\ell}_1, \ldots, \hat{\ell}_i, x_q^i] = \underset{\substack{\ell\in[1:i]\\ x\in\Omega}}{\operatorname{argmin}}\left(\left\|\boldsymbol{y}_i - \boldsymbol{H}_i x_q^i \boldsymbol{e}_{\ell_i}\right\|_F^2\right)\ . \tag{2}$$

5. Performance Analysis

In the performance evaluation of the proposed GC-SIMO-MBM, the BER metric is considered. Similarly to [31], ABEP is defined as:

$$P_e \le \frac{1}{2N_m M^2 m}\sum_{q=1}^{N_m M^2}\sum_{\hat{q}\neq q}^{N_m M^2} N(i,\hat{\imath})P(\boldsymbol{X}_q \to \boldsymbol{X}_{\hat{q}})\ , \tag{3}$$

where $P(\boldsymbol{X}_q \to \boldsymbol{X}_{\hat{q}})$ symbolizes the pairwise error probability (PEP) of $\boldsymbol{X}_{\hat{q}}$ detected at the receiver, given that $\boldsymbol{X}_q$ is transmitted, $\boldsymbol{X}_q = (x_q^1, x_q^2)$ and $\boldsymbol{X}_{\hat{q}} = (x_{\hat{q}}^1, x_{\hat{q}}^2)$, $N(i,\hat{\imath})$ stand for the bit error connected with the PEP event. Similarly to [35], the conditional PEP may be defined as:

$$\begin{aligned} P(\boldsymbol{X}_q\to\boldsymbol{X}_{\hat{q}}|\boldsymbol{H}_i) &= P\left(\left\|\boldsymbol{y}_i-\boldsymbol{H}_{\hat{\imath}}x_q^i\boldsymbol{e}_{\ell_i}\right\|_F^2 < \left\|\boldsymbol{y}_i-\boldsymbol{H}_i x_q^i\boldsymbol{e}_{\ell_i}\right\|_F^2|\boldsymbol{H}_i\right)\\ &= P\left(\sum_{i=1}^{2}\left\|\boldsymbol{H}_{\hat{\imath}}x_q^i\boldsymbol{e}_{\ell_i}+\boldsymbol{n}\right\|_F^2 < \sum_{i=1}^{2}\left\|\boldsymbol{n}\right\|_F^2\right)\\ &= Q\sum_{i=1}^{2}\alpha_i\ , \end{aligned} \tag{4}$$

where α_i is central chi-squared distribution with $2N_R$ degrees of freedom defined as:

$$\frac{\rho}{2}\left\|\boldsymbol{h}_i\right\|_F^2\left|d_x^i\right|^2 = \sum_{k=1}^{2N_R}\alpha_i^2$$

with $N(0,\ \sigma^2)$, $\sigma^2 = \frac{\rho}{4}\left|d_x^i\right|^2$.

The probability density function PDF of α_i^2, employing $f_{\alpha_i}(\alpha_i) = \frac{\alpha_i^{N_R-1}\mathrm{e}^{\frac{-\alpha_i}{2\sigma^2}}}{(2\sigma^2)^{N_R}(N_R-1)!}$ is similar to the PEP derivation of the GC-SIMO in [37], and is coupled with the trapezoidal approximation of the Q-function given in [17]. Therefore, the PEP for GC-SIMO-MBM can be defined as:

$$\frac{1}{4n}\left(\frac{1}{2}\prod_{i=1}^{2}\left(\frac{\rho}{4}\left|d_x^i\right|^2\right)^{-N_R} + \sum_{k=1}^{2}\left(\frac{\rho}{4}\frac{|d_x^i|^2}{u_k}\right)^{-N_R}\right)\ , \tag{5}$$

where ρ represents the signal-to-noise ratio (SNR), $n > 10$ for trapezoidal approximation convergence of the Q-function [17], $i \in [1:2]$, $k \in [1:2N_R]$, $u_k = \sin^2\left(\frac{k\pi}{2n}\right)$ and $\left|d_x^i\right|^2 = \left|x_q^i - x_{\hat{q}}^i\right|^2$.

6. Numerical Analysis and Discussion

The results of the simulation obtained for the proposed GC-SIMO-MBM scheme are presented in terms of average BER and SNR parameters. Likewise, the result of the evaluated theoretical ABEP is presented. In all cases, the ML detector is utilized.

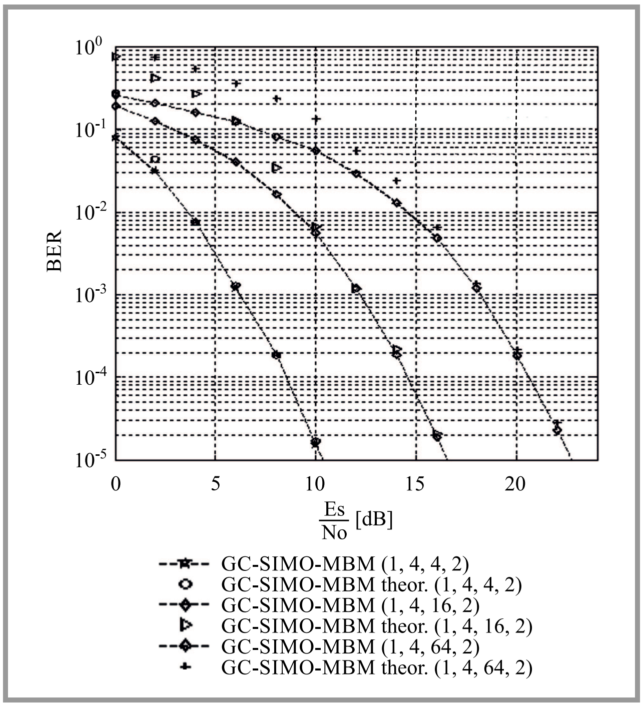


Fig. 2. Performance analysis validation for GC-SIMO-MBM for 4, and 8 b/s/Hz.

In Fig. 2, the GC-SIMO-MBM scheme is shown with a configuration setting of 1×4 4-QAM, 1×4 16-QAM and 1×4 64-QAM with 2 RF mirrors around each transmit antenna ($m_{rf} = 2$). This yields a spectral efficiency of 4, 6, and 8 b/s/Hz, respectively. The results of MC simulation obtained showed a close match with the average theoretical analysis in the high SNR region, validating the proposed scheme.

Figure 3 presents a comparison of performance between the GC modulation of SIMO-MBM [37] and the proposed GC-SIMO-MBM system with the same spectral efficiency of 4 and 6 b/s/Hz, respectively. The simulation results revealed that GC-SIMO-MBM outperforms its counterpart in 4 and 6 b/s/Hz.

One can see from the MC simulation results that using the MBM technique based on RF mirrors improves the system's error performance at a reduced hardware complexity. At a BER of 10^{-5}, GC-SIMO-MBM exhibits a significant performance gain of approximately 7 dB and 6.5 dB SNR for 4 and 6 b/s/Hz, respectively, compared to GC-SIMO from [37]. Similarly, GC-SIMO-MBM outperforms SIMO-MBM by 5 dB and 3.5 dB in 4 and 6 b/s/Hz, respectively.

7. Conclusion

In this paper media-based modulation was examined in a GC-SIMO scheme based on the RF mirrors to improve BER performance and enhance spectral efficiency. The topology based on the GC-SIMO-MBM technique offers

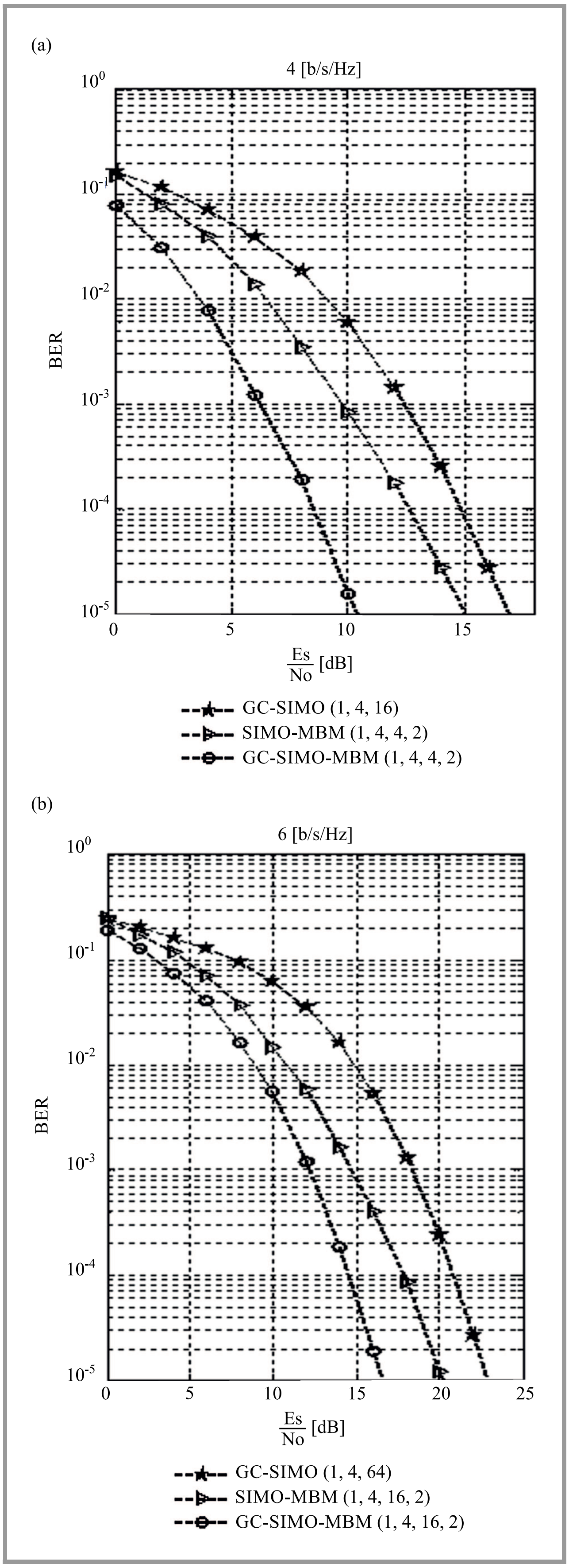


Fig. 3. BER performance comparison of GC-SIMO-MBM, GC-SIMO, and SIMO-MBM for 4 and 6 b/s/Hz.

better BER performance compared to SIMO-MBM and GC-SIMO of the same spectral efficiency. Results of the Monte Carlo simulation indicate that GC-SIMO-MBM shows a significant performance gain of approximately 7 dB and 6.5 dB SNR for 4 and 6 b/s/Hz, respectively, compared to GC-SIMO at a BER of 10^{-5}. The proposed GC-SIMO-MBM system is validated by theoretical and numerical results which show that the scheme is capable of significantly improving the system's hardware complexity, maximizing the linear relationship between RF mirrors and the spectral efficiency in MBM to accomplish a high data rate at a reduced hardware complexity by employing orthogonal pairs of the super-symbol in the GC scheme's encoder, transmitted via different RF mirrors at different time slots to achieve full rate and full diversity.

References

[1] A. Goldsmith, *Wireless Communications*, 1st ed. New York: Cambridge University Press, 2005 (ISBN: 9780511841224).

[2] S. Patel, T. Quazi, and H. Xu, "High-rate uncoded space-time labelling diversity with low-complexity detection", *Int. J. of Commun. Syst.*, vol. 33, no. 14, e4520, 2020 (DOI: 10.1002/dac.4520).

[3] T. Quazi and H. Xu, "SSD enhanced uncoded space-time labeling diversity", *Int. J. of Commun. Syst.*, vol. 31, no. 11, e3592, 2018 (DOI: doi.org/10.1002/dac.3592).

[4] A. Evans, T. Quazi, and H. Xu, "BER performance of a maximum ratio transmission enhanced hierarchical quadrature amplitude modulation (MRT-HQAM) system over Rayleigh fading channels", *IET Commun.*, vol. 10, no. 17, pp. 2473–2479, 2016 (DOI: 10.1049/iet-com.2016.0676).

[5] T. Quazi and H. Xu, "Simple UEP mechanism for multimedia traffic using the Alamouti structure with hierarchical modulation and signal space diversity", *IET Commun.*, vol. 8, no. 17, pp. 3128–3135, 2014 (DOI: 10.1049/iet-com.2014.0350).

[6] S. Solwa, M. K. Elmezughi, A. J. Bamisaye, D. Ayanda, A. Almaktoof, and M. T. E. Kahn, "Genetic algorithm-based uncoded M-ary phase shift keying space-time labelling diversity with three transmit antennas for future wireless networks", in *Proc. of the 9th Int. Conf. on Elec. and Electron. Engin. ICEEE 2022*, Alanya, Turkey, 2022, pp. 423-428 (DOI:10.1109/ICEEE55327.2022.9772544).

[7] N. R. Naidoo, H. Xu, and T. Quazi, "Spatial modulation: optimal detector asymptotic performance and multiple-stage detection", *IET Commun.*, vol. 5, pp. 1368–1376, 2019 (DOI: 10.1049/iet-com.2010.0667).

[8] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation", *IEEE Trans. on Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, 2008 (DOI: 10.1109/TVT.2007.912136).

[9] Z. Pan, J. Luo, J. Lei, L. Wen, and C. Tang, "Uplink spatial modulation SCMA system", *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 184–187, 2019 (DOI: 10.1109/LCOMM.2018.2882813).

[10] C. Zhong, X. Hu, X. Chen, D. W. K. Ng, and Z. Zhang, "Spatial modulation assisted multi-antenna non-orthogonal multiple access", *IEEE Wirel. Commun.*, vol. 25, no. 2, pp. 61–67, 2018 (DOI: 10.1109/MWC.2018.1700062).

[11] A. Khalid, T. Quazi, H. Xu, and S. Patel, "Performance analysis of M-APSK generalised spatial modulation with constellation reassignment", *Int. J. of Commun. Syst.*, vol. 33, no. 14, e4497, 2020 (DOI: 10.1002/dac.4497).

[12] A. Bhowal, R. Lalani, A. S. Sapre, and R. S. Kshetrimayum, "Advanced spatial modulation for efficient MIMO-based B2B communications in sporting activities", *IET Commun.*, vol. 13, no. 20, pp. 3529–3536, 2019 (DOI: 10.1049/iet-com.2019.0747).

[13] A. Bhowal and R. S. Kshetrimayum, "Advanced optical spatial modulation techniques for FSO communication", *IEEE Trans. on Commun.*, vol. 69, no. 2, pp. 1163–1174, 2021 (DOI: 10.1109/TCOMM.2020.3035400).

[14] T. Liu, "Analysis of the Alamouti STBC MIMO system with spatial division multiplexing over the Rayleigh fading channel", *IEEE Trans. on Wirel. Commun.*, vol. 14, no. 9, pp. 5156–5170, 2015 (DOI: 10.1109/TWC.2015.2433924).

[15] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Space-time block coded spatial modulation", *IEEE Trans. on Wirel. Commun.*, vol. 59, pp. 823–832, 2011 (DOI: 10.1109/TCOMM.2011.121410.100149).

[16] A. Saeed, H. Xu, and T. Quazi, "Alamouti space-time block coded hierarchical modulation with signal space diversity and MRC reception in Nakagami-m fading channel", *IET Commun.*, vol. 8, no. 4, pp. 516–524, 2014 (DOI: 10.1049/iet-com.2013.0519).

[17] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel", *IEEE Trans. on Inform. Theory*, vol. 44, no. 4, pp. 1453–1467, 1998 (DOI: 10.1109/18.681321).

[18] A. Saeed, T. Quazi, and H. Xu, "Hierarchical modulated QAM with signal space diversity and MRC reception in Nakagami-m fading channels", *IET Commun.*, vol. 7, no. 12, pp. 1296–1303, 2013 (DOI:10.1049/iet-com.2012.0750).

[19] A. K. Khandani, "Media-based modulation: Converting static Rayleigh fading to AWGN", in *Proc. IEEE Int. Symp. on Inform. Theory*, Honolulu, HI, USA, 2014, pp. 1549–1553 (DOI: 10.1109/ISIT.2014.6875093).

[20] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey", *IEEE Commun. Surv. and Tutor.*, vol. 21, no. 1, pp. 315–348, 2019 (DOI: 10.1109/COMST.2018.2858567).

[21] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors", *IEEE Trans. on Veh. Technol.*, vol. 66, no. 6, pp. 4967–4983, 2017 (DOI: 10.1109/TVT.2016.2620989).

[22] E. Seifi, M. Atamanesh, and A. K. Khandani, "Media-based MIMO: Outperforming known limits in wireless", in *Proc. IEEE Int. Conf. on Commun. ICC 2016*, Kuala Lumpur, Malaysia, 2016 (DOI: 10.1109/ICC.2016.7511273).

[23] E. Seifi, M. Atamanesh, and A. K Khandani, "Performance evaluation of media-based modulation in comparison with spatial modulation and legacy SISO/MIMO", in *Proc. 28th Biennial Symp. on Commun. BSC 2016*, Kelowna, British Columbia, 2016 [Online]. Available: http://cst.uwaterloo.ca/content/bsc-2016.pdf

[24] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors", *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4967–4983, 2017 (DOI: 10.1109/TVT.2016.2620989).

[25] Y. Naresh and A. Chockalingam, "A low-complexity maximum likelihood detector for differential media-based modulation", *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2158–2161, 2017 (DOI: 10.1109/LCOMM.2017.2687921).

[26] E. Basar, "Index modulation techniques for 5G wireless networks", *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, 2016 (DOI: 10.1109/MCOM.2016.7509396).

[27] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation", *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3892–3906, 2017 (DOI: 10.1109/TCOMM.2017.2710312).

[28] M. D. Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple antenna wireless systems: A survey", *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, 2011 (DOI: 10.1109/MCOM.2011.6094024).

[29] M. Nakao, T. Ishihara, and S. Sugiura, "Single-carrier frequency-domain equalization with index modulation", *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 298–301, 2017 (DOI: 10.1109/LCOMM.2016.2626447).

[30] M. Nakao, T. Ishihara, and S. Sugiura, "Dual-mode time-domain index modulation for Nyquist-criterion and faster-than-Nyquist single-carrier transmissions", *IEEE Access*, vol. 5, pp. 27659–27667, 2017 (DOI: 10.1109/ACCESS.2017.2768539).

[31] Y. Naresh and A. Chockalingam, "Performance analysis of media-based modulation with imperfect channel state information", *IEEE Trans. on Veh. Technol.*, vol. 67, no. 5, pp. 4192–4207, 2018 (DOI: 10.1109/TVT.2018.2791845).

[32] N. Pillay and H. Xu, "Quadrature spatial media-based modulation with RF mirrors", *IET Commun.*, vol. 11, pp. 2440–2448, 2017 (DOI: 10.1049/iet-com.2017.0269).

[33] R. Pillay, N. Pillay, and H. Xu, "A study of single-input multiple-output media-based modulation with RF mirrors", in *Proc. of the Southern Africa Telecommun. Netw. and Appl. Conf. SATNAC 2017*, Barcelona, Spain, 2017, pp. 20–25, 2017 [Online]. Available: https://www.satnac.org.za/assets/documents/proceedings/SATNAC_2017_Proceedings.pdf

[34] T. Mao, Q. Wang, M. Wen and Z. Wang, "Secure single-input-multiple-output media-based modulation", *IEEE Trans. on Veh. Technol.*, vol. 69, no. 4, pp. 4105–4117, 2020 (DOI: 10.1109/TVT.2020.2975303).

[35] A. J. Bamisaye and T. Quazi, "Two-way decode and forward quadrature media-based modulation for single-input multiple-output scheme", *Int. J. of Commun. Sys.*, e5186, 2022 (DOI: 10.1002/dac.5186).

[36] A. J. Bamisaye and T. Quazi, "Quadrature spatial modulation-aided single-input multiple output-media-based modulation", *Int. J. of Commun. Syst.*, vol. 34, no. 11, e4883, 2021 (DOI: 10.1002/dac.4883).

[37] H. Xu and N. Pillay, "Golden codeword based modulation schemes for single-input multiple-output systems", *Int. J. of Commun. Syst.*, e3963, 2019 (DOI: 10.1002/dac.3963).

[38] L. Luzzi, G. R. Othman, J. Belfiore, and E. Viterbo, "Golden space-time block-coded modulation", *IEEE Trans. on Inform. Theory*, 2021, vol. 55, no. 2, pp. 584–597, 2009 (DOI: 10.1109/TIT.2008.2009846).

[39] J. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2 × 2 full-rate space-time code with non-vanishing determinants", in *Proc. of Int. Symp. on Inform. Theory ISIT 2004*, Chicago, IL, USA, 2004 (DOI: 10.1109/ISIT.2004.1365347).

[40] H. Xu and N. Pillay, "The component-interleaved golden code and its low-complexity detection", *IEEE Access*, vol. 8, pp. 59550–59558, 2020 (DOI: 10.1109/ACCESS.2020.2982673).

[41] H. Xu and N. Pillay, "Multiple complex symbol golden code", *IEEE Access*, vol. 8, pp. 103576–103584, 2020 (DOI: 10.1109/ACCESS.2020.2997308).

[42] N. Pillay and H. Xu, "RF mirror media-based modulation for golden codes", *J. of Telecommun. Electron. and Comp. Engin.*, vol. 10,2021, no. 3, pp. 21–24, 2018 [Online]. Available: https://jtec.utem.edu.my/jtec/article/view/3376/3419

Ayodeji James Bamisaye received his M.Sc. in Electrical and Electronics Engineering (communication option) from Federal University of Technology, Akure, Nigeria, in 2010, for his research on evaluating multiusers' interference on radiolocation in CDMA cellular systems. In 2019, he commenced his Ph.D. research at the Department of Electrical, Electronics and Computer Engineering at the University of Kwazulu-Natal, Durban, South Africa. His Ph.D. research focused on quadrature spatial modulation (QSM) in an RF mirror-based MBM system and in cooperative networks. His current research interests are in digital communications, as well as wireless and mobile communications.

https://orcid.org/0000-0001-7329-3869
E-mail: ayobamisaye@gmail.com
219038995@stu.ukzn.ac.za
Department of Electrical, Electronic and Computer Engineering
University of Kwazulu-Natal
Durban, South Africa

Tahmid Quazi has been a wireless communications researcher since 2002. He obtained his M.Sc. in 2003 for his research and development work concerning mobile ad-hoc networks for tactical communications. He began his Ph.D. in the area of cross-layer design (CLD) in wireless communications in 2005, which he completed in 2009. More recently, his work has been focusing on a variety of diversity systems, including space, time, signal, labelling, and polarization diversity, with the general aim of improving error rate performance of wireless links. Tahmid is passionate about using technology for the betterment of mankind, and channels his research accordingly. In addition to his research in the field of advanced telecommunication systems, in order to help bridge the digital divide, he runs numerous R&D projects focusing on IoT, embedded systems and signal processing systems.

https://orcid.org/0000-0002-1288-4224
E-mail: quazit@ukzn.ac.za
Department of Electrical, Electronic and Computer Engineering
University of Kwazulu-Natal
Durban, South Africa

Paper

Shallow Layer Convolutional Features with Correlation Filters for UAV Object Tracking

Budi Syihabuddin, Suryo Adhi Wibowo, Agus D. Prasetyo, and Desti Madya Saputri

School of Electrical Engineering, Telkom University, Bandung, Indonesia

https://doi.org/10.26636/jtit.2022.150020

Abstract—In this paper, convolutional shallow features are proposed for unmanned aerial vehicle (UAV) tracking. These convolutional shallow features are generated by pre-trained convolutional neural networks (CNN) and are used to represent the target objects. Furthermore, to estimate the location of the target objects, an adaptive correlation filter based on the Fourier transform is used. This filter is multiplied with the convolutional shallow features by using pixel-wise multiplication in the Fourier domain. Then, the inverse of Fourier is performed to estimate the location of the target object, where its location is represented by the maximum value of the response map. Unfortunately, the target object always changes its appearance during tracking. Therefore, we proposed an updated model to address this issue. The proposed method is evaluated by using the UAV123_10fps benchmark dataset. Based on the comprehensive experimental results, the proposed method performs favorably against state-of-the-art tracking algorithms.

***Keywords*—CNN, convolutional features, correlation filter, object tracking, shallow layer, UAV tracking.**

1. Introduction

Unmanned aerial vehicles (UAV) with remote sensing capabilities are used in many modern applications, such as object tracking and object recognition [1], [2]. In object tracking, numerous problems are encountered, such as aspect ratio change, background clutter, camera motion, fast motion, full occlusion, illumination variations, low resolution, out-of-view situations, partial occlusion, scale variation, presence of similar objects, and viewpoint change [3]. Those problems mean that object tracking systems have to comply with a number of requirements. Such systems must be capable of defining the next state, if they were given an initial state, for instance the initial object location or the initial object size. Object recognition can be useful for surveillance and human-computer interactions, but requires that numerous problems and issues be solved.

At the beginning of 2000, many researchers proposed a generative approach, i.e. suggested that an adaptive color histogram be used to identify objects [4]. The adaptive color histogram can be represented as an object, and a particle filter is used to estimate the next state. A similar method was used by [5] and [6]. The method is simple and easy, but that is why it suffers from a specific disadvantage. It is hard to identify an object using an adaptive color histogram if the distractor is characterized by similar color features. The particle filter uses Bayesian distribution to achieve a high level of accuracy. Distributions with many particles make the system more complex.

To compensate for the disadvantages of the previous method, the researchers proposed a discriminative approach based on boosting the classifier. This method uses a background model as initial information to come up with a robust object tracking algorithm. Grabner *et al.* proposed online learning relying on the Adaboost classifier for object tracking [7]. This approach was developed in paper [8] by proposing semi-supervised online boosting for object tracking and multiple instance learning based on boosting the classifier proposed in [9]. Zhang *et al.* added weight calculation based on distance and updated the model to approve accuracy of the system. Kalal *et al.* proposed a discriminative approach for long-term tracking based on tracking learning detection (TLD).

The discriminative approach using a boosting classifier suffers from certain disadvantages, i.e. a limited area taken as a positive sample to be represented as the target object. If the object moves quickly or abruptly, the method will have difficulty detecting it. This method can achieve good performance even if the object's color characteristics are similar to those of the distractor. However, by using the integral image, this method will run into a problem if occlusion is encountered.

To solve the occlusion problem, some papers recommended object representation based on a sparse coefficient vector. This method uses a generative approach and is particle filter to estimate the tracked object's location. A sparse coefficient vector has been researched by [12]–[14], and Wibowo *et al.* proposed a sparse coefficient vector to minimize computation time [15]. Computational time still remains a problem to be solved besides the fast motion and background clutter. As this method is being developed, it can be replaced by a correlation filter.

A correlation filter estimates the tracked object's location using the Fourier transform. Bolme *et al.* proposed the minimum output sum of square error (MOSSE) approach that relies on a correlation filter and is capable of working adaptively [16]. This method can only work for simple linear classification problems. To boost the performance of the correlation filter, Henriques *et al.* used the ridge regression problem and the circulant matrix [17]. Other methods to increase the performance of the correlation filter include color histogram features, histogram of Gaussian (HOG) features and complementary learners [18]–[22]. For this paper, we proposed convolutional shallow features from a pre-trained CNN network to predict the movement of the object.

The remaining part of this paper is organized as follows. Section 2 discusses the correlation filter, and Section 3 presents the proposed method. Experimental results of the UAV123_10fps benchmark dataset and the result are presented and discussed in Section 4. Finally, Section 5 concludes this paper.

2. Correlation Filters

A correlation filter can be used to estimate the location of the targeted object. We work in the frequency domain to minimize computation time. Correlation filter h and the input that has been proceeded x (i.e. smoothen feature extraction) have to be transformed using the discrete Fourier transform (DFT). The output of h and x can be multiplied by each element to substitute the convolution process. Practically, DFT may be changed by means of the fast Fourier transform (FFT) to make the process more efficient. The maximum value of the inverse Fourier transform can be assumed as the location of the target object. Those processes can be described in the following manner:

$$\begin{aligned} x \otimes h &= F^{-1}(\hat{x} \odot \hat{h}^{*}) , \\ m &= F^{-1}(\hat{x} \odot \hat{h}^{*}) , \end{aligned} \tag{1}$$

where F^{-1} is the inverse Fourier transform, $\hat{x}$ is the smoothened feature extraction in the frequency domain, $\hat{h}$ is the correlation filter in the frequency domain, $*$ is complex conjugate, and $\odot$ is element-wise multiplication.

The correlation filter may use several data training approaches. In this case, it uses an image patch from the initial position from the first frame. To obtain the correlation filter, we can rely on minimization based on the following equation:

$$\min_{\bar{h}} \sum_{i} L\big(f(h,x_i),m_i\big) + \omega\,||h||^2 , \tag{2}$$

where $L(\cdot)$ is the loss function.

The $L(f(h,x_i),m_i) = ||h \cdot x_i - m_i||^2$ can be expressed as:

$$\min_{\bar{h}} \sum_{i} ||h \cdot x_i - m_i||^2 + \omega\,||h||_2^2 , \tag{3}$$

where h is the correlation filter, x is the smoothened feature extraction, m is the expected output, and ω is the control value to prevent overfitting.

Equation (3) can be simplified to:

$$h = (X^T X + \omega D)^{-1} X^T \mathbf{m} , \tag{4}$$

where D is the identity matrix, $\mathbf{m}$ is the labeled vector, and X is the training samples matrix. Since the computation takes place in the frequency domain, X^T has to be transformed into $X^H = (X^*)^T$, with $\mathbf{H}$ as the Hermitian matrix. To simplify the calculations for Eq. (4), we can use the circulant matrix as an approach. So, it can be expressed as:

$$\hat{h} = A \operatorname{diag}\left(\frac{\hat{x}}{\hat{x}^{*} \odot \hat{x} + \omega}\right) A^H \mathbf{m} , \tag{5}$$

with A being the DFT matrix. In the frequency domain, Eq. (5) will be:

$$\hat{h}_i = \frac{\hat{m}_i \odot \hat{x}_i^{*}}{\hat{x}_i \odot \hat{x}_i^{*} + \omega} . \tag{6}$$

3. Proposed Method

The CNN is a deep learning method that comprises convolutional layers, normalization layers and pooling layers. In the convolutional layers, it can be defined from the shallow layer to the deep layer. In this paper, we investigate the shallow layer from the convolutional layers to represent the target. We are using the CNN model proposed by Simonyan *et al.*, where the data set is trained by a big benchmark dataset [24]. We suggest checking [24] for architectural details and for the pre-trained CNN model. The proposed method is shown in Fig. 1.

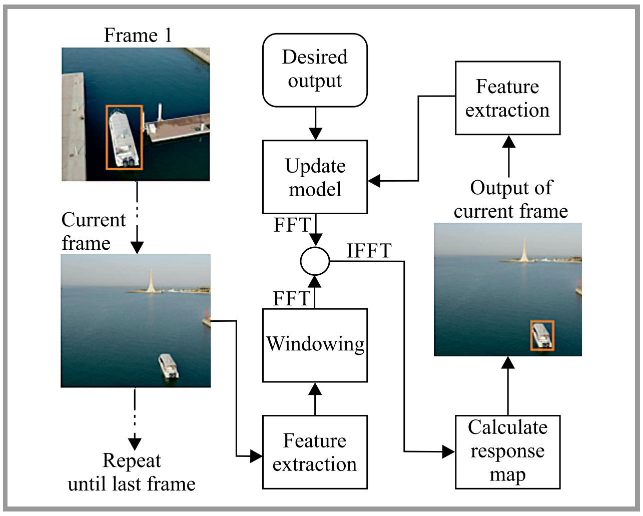


Fig. 1. Framework of the proposed method.

When we track an object from the same point of view but with the object moving, the appearance of the object will not be the same as it is in the initial state. It may be different in shape. To solve this problem, we must design a robust system for tracking objects. Updating the model is one of the potential methods. If we use the correlation filter, then the filters will be updated in every frame. Referring to Eq. (6), for updating the correlation filters we use:

$$h_t = \frac{\beta_t}{\gamma_t + \omega} \ , \tag{7}$$

with h_t, β_t, γ_t, and ω being the correlation filters, numerator, denominator, and value at frame t, respectively. Equation (7) is solved using a linear system $I \times I$. For the numerator, we can use the following formula:

$$\beta_t = \alpha_1 \beta_{t-1} + \alpha_2 \left(m \odot x_t^* \right) \ , \tag{8}$$

where α_1, α_2, and β_{t-1} are weigh factor 1, weigh factor 2, and numerator for the previous frame $t-1$, respectively. The denominator can be solved by:

$$\gamma_t = \alpha_1 \gamma_{t-1} + \alpha_2 \sum_i x_t \odot x_t^* \ , \tag{9}$$

where x_t and γ_{t-1} are feature extraction at t and denominator at the previous frame $t-1$, respectively.

4. Experimental Results and Discussion

An updated model is needed to overcome changes in the appearance of the target object. In this step, variables ω, α_1, and α_2 exist, having the values of 0.001, 0.02, and 0.08, respectively. To validate the proposed method, referred to as csfUAVt, our tracker is evaluated with the use of the UAV123_10fps benchmark dataset. This dataset consists of 72 videos that contain several challenging problems, such as aspect ratio change, background clutter, camera motion, fast motion, full occlusion, illumination variation, low resolution, out-of-view, partial occlusion, scale variation, presence of a similar object, and viewpoint change. The proposed method is evaluated quantitatively using success plots based on the overlap ratio and precision plots based on the center location error. This evaluation is carried out following the one-pass-evaluation (OPE) protocol described in [3]. The proposed method was implemented using Matlab.

In this quantitative evaluation, the proposed method is compared with nine state-of-the-art tracking methods, such as ASLA [25], CSK [27], KCF [17], DSST [19], IVT [23], MOSSE [16], DCF, Struck [26], and TLD [11]. The results of the evaluation for the success plots of OPE are presented in Fig. 2. In the case of aspect ratio change, our proposed method obtains the best performance with a success rate of 0.289 and an overlap threshold value of 0.5. Meanwhile, TLD, Struck, DSST, and ASLA trackers are ranked second, third, fourth, and fifth with success rates of 0.287, 0.232, 0.227, and 0.212, respectively. For the TLD tracker, the features used are points and motion prediction is aided using optical flow. Meanwhile, for DSST, the feature used is the histogram of Gaussian (HOG). Based on the results for the aspect ratio change, our proposed method, using convolutional shallow features and correlation filters, offers the best performance compared to nine other tracker algorithms.

In the case of background clutter, the proposed method ranks second with a success rate of 0.279, while the winning Struck tracker outperforms the proposed method with a success rate of 0.361 for an overlap threshold value of 0.5. DSST, KCF, and DCF are ranked third, fourth, and fifth with success rates of 0.240, 0.232, and 0.231, respectively. Furthermore, the Struck tracker itself is developed based on a kernelized structured output support vector machine (SVM). Based on the results of these experiments, the tracker achieves superior performance compared to nine other trackers for background clutter problems.

In the case of camera motion, the proposed method provides the best performance, with a success rate of 0.376. Meanwhile, Struck, TLD, DSST, and MOSSE trackers are ranked second, third, fourth, and fifth with success rates of 0.280, 0.278, 0.244, and 0.237, respectively. The MOSSE tracker itself is a tracking algorithm based on adaptive correlation filters using the HOG feature. The experimental results show that the proposed method achieves better performance than nine other trackers for camera motion problems. Furthermore, for fast motion problems, the first, second, third, fourth, and fifth places are occupied by the proposed method, DSST, TLD, DCF, and CSK, respectively, with their respective success rates equaling 0.257, 0.160, 0.152, 0.149, and 0.136. In the case of fast motion, the proposed method shows better performance than the remaining tracking algorithms, with a success rate difference of 0.097 compared to the second rank.

Figure 2 shows the full occlusion problem in object tracking, with the entire shape of the target object being obscured by the distractor and making it invisible. In this case, the proposed method ranks first, with a success rate of 0.239, winning by a difference of 0.083 compared with the Struck tracker, ranking second. Meanwhile, ranks three, four, and five are occupied by TLD, MOSSE, and CSK, with their success rates equaling 0.149, 0.138, and 0.123, respectively. In the case of illumination variation, the proposed method ranks first with a success rate of 0.252, while the second, third, fourth, and fifth ranks are occupied by Struck, DSST, DCF, and KCF trackers achieving the success rates of 0.215, 0.187, 0.174, and 0.170, respectively.

In the case of low resolution, the Struck approach ranks first, with a success rate of 0.296, while the proposed method is placed fourth, rank with a success rate of 0.232. The second, third, and fifth ranks are occupied by TLD, ASLA, and MOSSE trackers, respectively. ASLA is a tracking algorithm that utilizes the sparse coefficient vector feature. Furthermore, in the case of out-of-view scenario, the proposed method ranks first, with a success rate of 0.5, outperforming the second-ranking DCF approach by 0.173. Meanwhile, DSST, KCF, and CSK occupy the third, fourth, and fifth place, with success rates of 0.327, 0.327, and 0.314, respectively.

Furthermore, in the case of partial occlusion, scale variation, similar objects, and viewpoint change, the proposed method offers superb results. It ranks first, showing the success rates of 0.375, 0.334, 0.438, and 0.342. The success plot of OPE is summarized in Table 1. Based on the

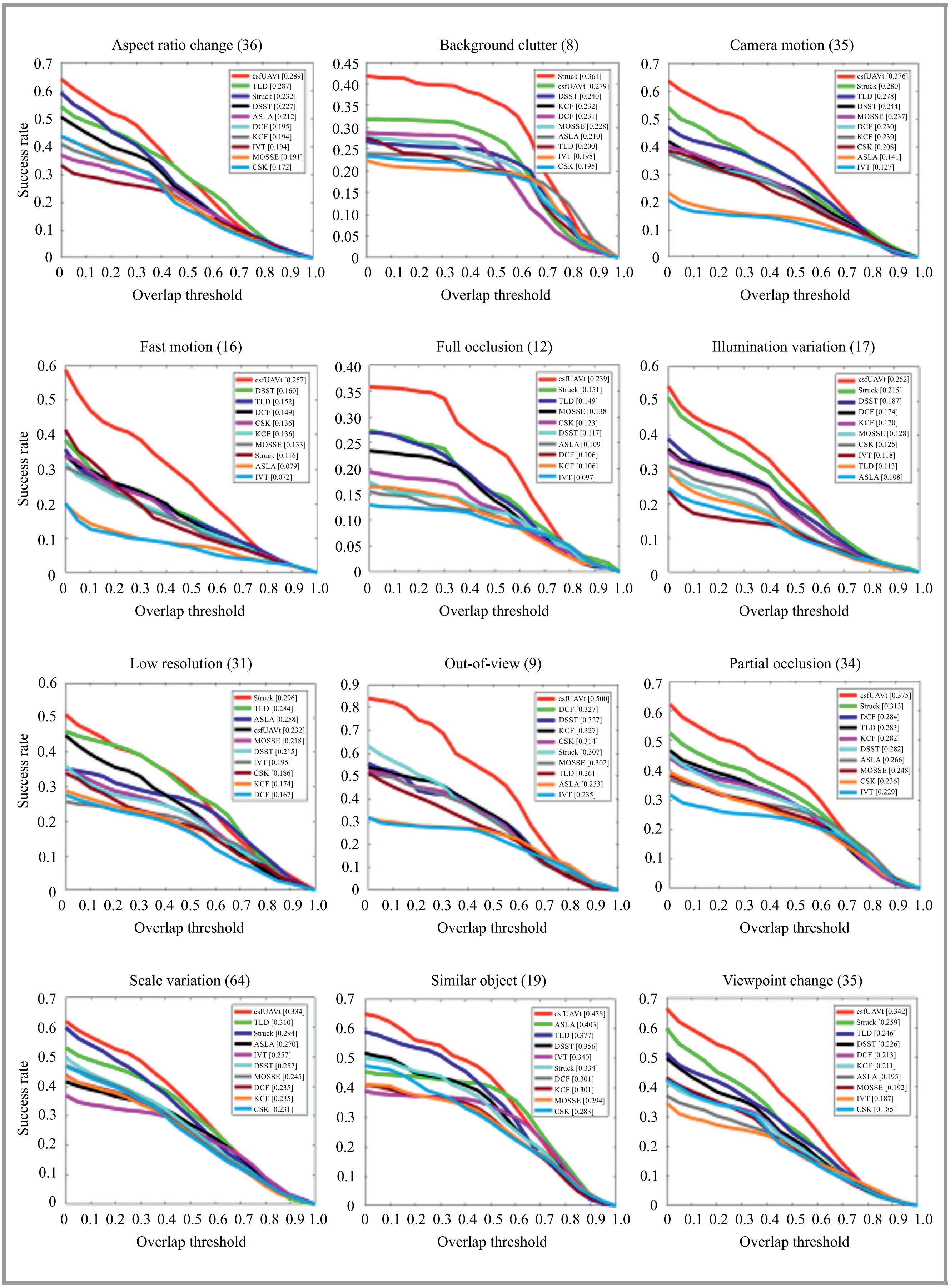


Fig. 2. Success plots of OPE for all types of problems encountered.

Table 1
Success plot of OPE

	Success plots of OPE for all types of problems encountered									
	csfUAVt	TLD	Struck	DSST	ASLA	DCF	KCF	IVT	MOSSE	CSK
Aspect ratio change	**0.289**	0.287	0.232	0.227	0.212	0.195	0.194	0.194	0.191	0.172
Background clutter	0.279	0.200	**0.361**	0.240	0.210	0.231	0.232	0.198	0.228	0.195
Camera motion	**0.376**	0.278	0.280	0.244	0.141	0.230	0.230	0.127	0.237	0.208
Fast motion	**0.257**	0.152	0.116	0.160	0.116	0.149	0.136	0.072	0.133	0.136
Full Occlusion	**0.239**	0.149	0.151	0.117	0.109	0.106	0.106	0.097	0.138	0.123
Illumination variation	**0.252**	0.113	0.215	0.187	0.108	0.174	0.170	0.118	0.128	0.125
Low resolution	0.232	0.284	**0.296**	0.215	0.258	0.167	0.174	0.195	0.218	0.186
Out-of-view	**0.500**	0.261	0.307	0.327	0.253	0.327	0.327	0.235	0.302	0.314
Partial occlusion	**0.375**	0.283	0.313	0.282	0.266	0.284	0.282	0.229	0.248	0.236
Scale variation	**0.334**	0.310	0.294	0.257	0.270	0.235	0.235	0.257	0.245	0.231
Similar object	**0.438**	0.377	0.334	0.356	0.403	0.301	0.301	0.340	0.294	0.283
Viewpoint change	**0.342**	0.246	0.259	0.226	0.195	0.213	0.211	0.187	0.192	0.185

experiment results, it is proved that the proposed method is more robust and useful than the nine other algorithms that are used as a benchmark for this specific UAV tracking application.

In addition the success rate, in this quantitative evaluation, the performance of our proposed method is also evaluated based on the precision plot parameters. The evaluation results for the precision plots of OPE are presented in Fig. 3. In the case of ratio change, the proposed method ranks first with a precision plot of 0.458 and a location error threshold of 20 pixels. Meanwhile, Struck, TLD, DSST, and DCF approaches are ranked second, third, fourth, and fifth, with success rates of 0.376, 0.376, 0.369, and 0.301, respectively. DCF is a tracking algorithm that relies on correlators and HOG as its features.

In the case of background clutter, the proposed method ranks second with a precision plot of 0.351. The first position, meanwhile, is occupied by the Struck method which outperforms the proposed approach thanks to its success rate of 0.443 and a location error threshold of 20 pixels. Meanwhile, DCF, KCF, and TLD are ranked third, fourth, and fifth, with prediction plots of 0.316, 0.316, and 0.294, respectively. Background clutter is a problem that is experienced in object tracking when background in close proximity to the target has the same color or texture as the target itself.

In the case of camera motion, the proposed method ranks first with a precision plot of 0.495. Meanwhile, Struck, TLD, DSST, and MOSSE are ranked second, third, fourth, and fifth with success rates of 0.379, 0.341, 0.326, and 0.295, respectively. In the case of camera motion, results of these experiments show that the proposed method offers better performance than 9 remaining trackers. Furthermore, in the case of fast motion, the second, third, fourth, and fifth ranks are occupied by the proposed method, DSST, CSK, TLD, and DCF, with their precision plot values amounting to 0.341, 0.257, 0.234, 0.197, and 0.175, respectively. In the case of fast motion, the proposed method shows better performance than other tracking algorithms, with its precision plot value differing by 0.084 compared to the second rank.

Figure 3 shows the full occlusion problem encountered in object tracking. In this case, the proposed method ranks first, with a precision plot of 0.435, winning by a margin of 0.098 compared with the second rank occupied by the Struck tracker. Meanwhile, third, fourth, and fifth ranks are occupied by MOSSE, TLD, and CSK approaches, with their precision plots equaling 0.3, 0.296, and 0.274, respectively. In the case of illumination variation, the proposed method ranks first with a success rate of 0.369, while second, third, fourth, and fifth ranks are occupied by Struck, DSST, DCF, and KCF methods, showing precision plot values of 0.325, 0.313, 0.236, and 0.227, respectively. Illumination variation is a problem encountered in object tracking, caused by significant changes in the illumination intensity in the region of the target object.

In the case of low resolution, the Struck approach ranks first with a precision plot of 0.5. Meanwhile, the proposed method occupies rank three, with a precision plot of 0.449. The second, third, and fifth ranks are occupied by TLD, DSST, and ASLA methods, with precision plot values of 0.455, 0.370, and 0.364, respectively. Low resolution is a problem encountered in object tracking due to the number of pixels in the ground-truth bounding box being smaller than 400 pixels. Furthermore, in the out-of-view scenario, the proposed method ranks first, with a precision plot of 0.640, outperforming the second-ranking a precision plot of 0.640, outperforming the second-ranking DSST approach by 0.221. Meanwhile, DCF, KCF, and Struck methods are ranked third, and fourth, respectively.

In the case of partial occlusion, scale variation, presence of a similar object, and viewpoint change, the proposed

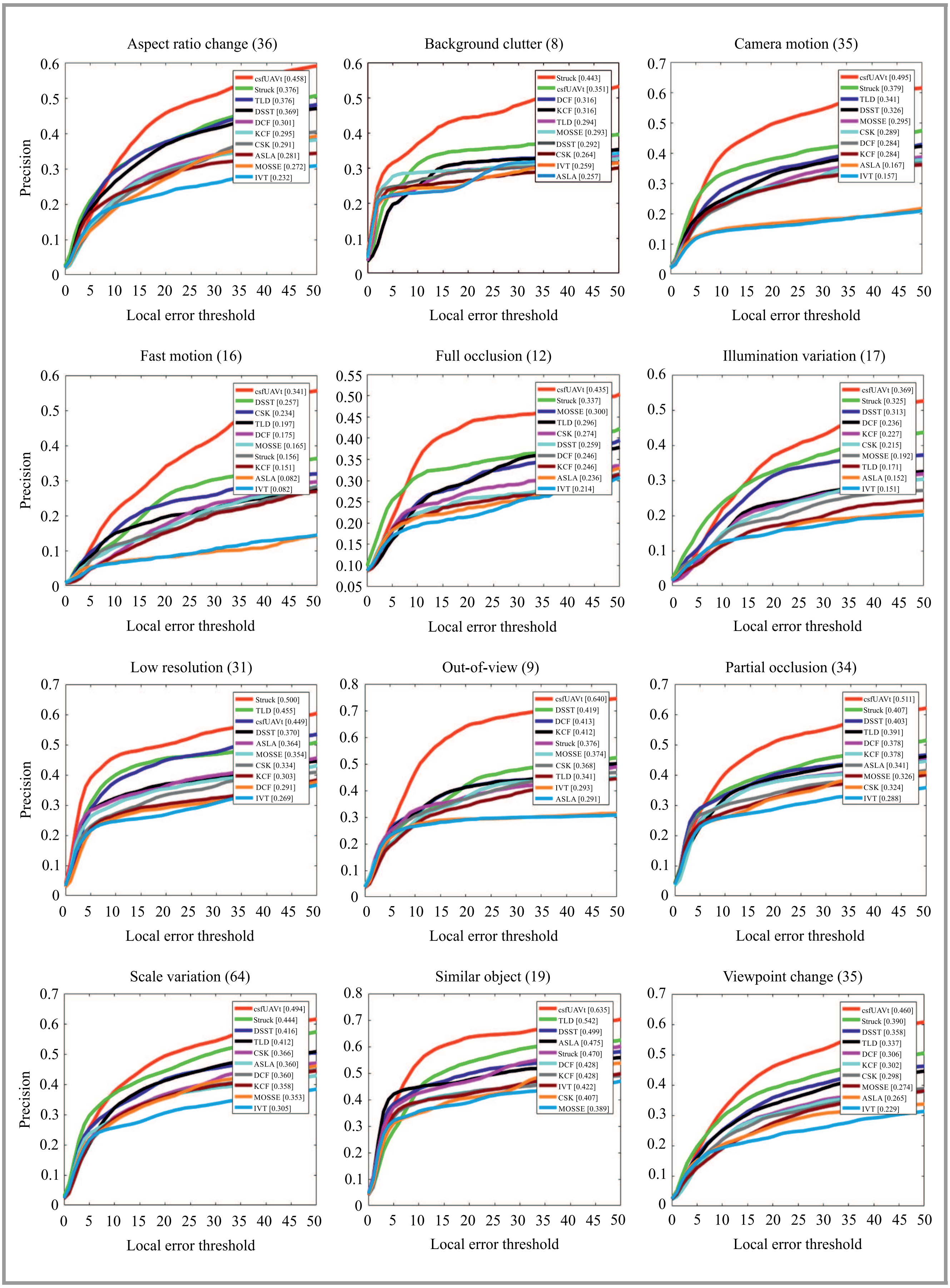


Fig. 3. Precision plots of OPE for all types of problems encountered.

Table 2
Precision plot of OPE

	Precision plots of OPE for all types of problems encountered									
	csfUAVt	TLD	Struck	DSST	ASLA	DCF	KCF	IVT	MOSSE	CSK
Aspect ratio change	**0.458**	0.376	0.376	0.369	0.281	0.301	0.295	0.232	0.272	0.291
Background clutter	**0.351**	0.294	0.443	0.292	0.257	0.316	0.316	0.259	0.293	0.264
Camera motion	**0.495**	0.341	0.379	0.326	0.167	0.284	0.284	0.157	0.295	0.289
Fast motion	**0.341**	0.197	0.156	0.257	0.082	0.175	0.151	0.082	0.165	0.234
Full occlusion	**0.435**	0.296	0.337	0.259	0.236	0.246	0.246	0.214	0.300	0.274
Illumination variation	**0.369**	0.171	0.325	0.313	0.152	0.236	0.227	0.151	0.192	0.215
Low resolution	0.449	0.455	**0.500**	0.370	0.364	0.291	0.303	0.269	0.354	0.334
Out-of-view	**0.640**	0.341	0.376	0.419	0.291	0.413	0.412	0.293	0.374	0.368
Partial occlusion	**0.511**	0.391	0.407	0.403	0.341	0.378	0.378	0.288	0.326	0.324
Scale variation	**0.494**	0.412	0.444	0.416	0.360	0.360	0.358	0.305	0.353	0.366
Similar object	**0.635**	0.542	0.470	0.499	0.475	0.428	0.428	0.422	0.389	0.407
Viewpoint change	**0.460**	0.337	0.390	0.358	0.265	0.306	0.302	0.229	0.274	0.298

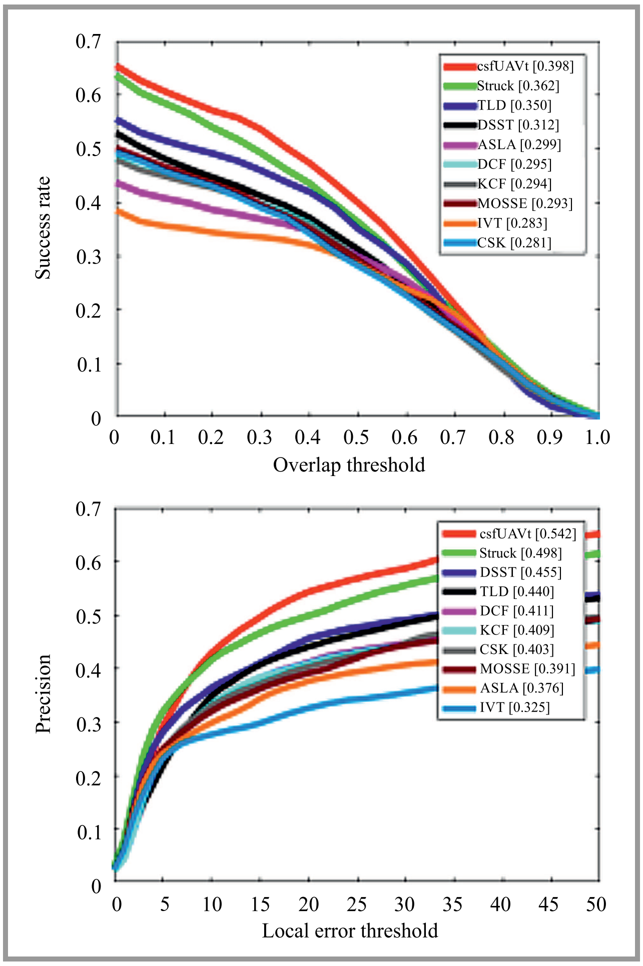


Fig. 4. Success plot and precision plot of OPE.

method offers superb results, ranking first in all the above-mentioned scenarios and exhibiting precision plot values of 0.511, 0.494, 0.635, and 0.460, respectively. The precision plot of OPE is summarized in Table 2. Based on the results of these experiments, it is proved that the proposed method is more precise than the nine algorithms that are used as a benchmark. Scale variation is a problem encountered in object tracking and influenced by the ratio between the bounding box in the first frame and in the latest out-of-range frame. Meanwhile, a similar object involves the presence of a distractor that has a similar color or texture to those of the target, and a viewpoint change is a problem caused by the difference in the target observation point, occurring between the first and the current frame.

After the success rate and precision plot have been calculated, each problem with UAV tracking is obtained. The average of each success rate and precision plot is calculated. The results of those calculations are represented in Fig. 4. In terms of the success rate of OPE, the proposed method ranks first, with a success rate of 0.398. Meanwhile, Struck, TLD, DSST, and ASLA approached occupy second, third, fourth, and fifth places, with their respective success rate values of 0.362, 0.350, 0.312, and 0.299.

In the case of precision plot, the proposed method also ranks first, with a precision plot value of 0.542, outperforming the second-ranking Struck methods by 0.044. Meanwhile, DSST occupies the third place, with a precision plot value of 0.455 and a location error threshold of 20 pixels. The fourth and fifth places are occupied by TLD and DCF, offering precision plot values of 0.440 and 0.411, respectively.

5. Conclusion

In this paper, convolutional features are taken from a CNN pre-trained on the shallow layer and harnessed using framework correlation filters. To solve the problem of changes in the appearance of the target object during tracking, the model is updated by correlation filters. In this updated model, numerator and denominator variables affecting the correlation filters are worked out. To validate the proposed method, an experiment using the UAV123_10fps benchmark dataset was performed.

Based on the results of a quantitative evaluation relying on such parameters as success plots and precision plots, the proposed method ranks first in all scenarios, beating 9 other state-of-the-art tracking algorithms, with the average success plots of OPE equaling 0.398, and the average precision plots of OPE amounting to 0.542.

Acknowledgements

This study was by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia under the *Penelitian Dasar Kompetitif Nasional* 2019–2021 project.

References

[1] X. Qin and T. Wang, "Visual-based tracking and control algorithm design for quadcopter UAV", in *Proc. of the Chinese Control and Decision Conf. CCDC 2019*, Nanchang, China, 2019 (DOI: 10.1109/CCDC.2019.8832545).

[2] Z. Zheng and H. Yao, "A method for UAV tracking target in obstacle environment", in *Proc. Chinese Autom. Congr. CAC 2019*, Nanchang, China, 2019, pp. 4639–4644 (DOI: 10.1109/CCDC.2019.8832545).

[3] M. Mueller, N. Smith, and B. Ghanem, "A benchmark and simulator for UAV tracking", in in *Computer Vision – ECCV 2016. 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. *LNCS*, vol. 9905, pp. 445–461. Cham: Springer, 2016 (DOI: 10.1007/978-3-319-46448-0_27).

[4] K. Nummiaro, E. Koller-Meier, and L. Van Gool, "An adaptive color-based particle filter", *Image and Vis. Comput.*, vol. 21, no. 1, pp. 99–110, 2003 (DOI: 10.1016/S0262-8856(02)00129-4).

[5] S.-K. Weng, C.-M. Kuo, and S.-K. Tu, "Video object tracking using adaptive Kalman filter", *J. of Visual Commun. and Image Represent.*, vol. 17, no. 6, pp. 1190–1208, 2006 (DOI: 10.1016/j.jvcir.2006.03.004).

[6] S. A. Wibowo, H. Lee, E. K. Kim, and S. Kim, "Tracking failures detection and correction for face tracking by detection approach based on fuzzy coding histogram and point representation", in *Proc. of the Int. Conf. on Fuzzy Theory and Its Appl. iFUZZY 2015*, Yilan, Taiwan, 2015, pp. 34–39 (DOI: 10.1109/iFUZZY.2015.7391890).

[7] H. Grabner, M. Grabner, and H. Bischof, "Real-time tracking via on-line boosting", in *Proc. of the 17th British Machine Vision Conf. BMVC 06*, Edinburgh, Scotland, 2006 (DOI: 10.5244/C.20.6).

[8] H. Grabner, C. Leitsner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking", in Computer Vision – ECCV 2008 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part I, D. Forsyth, P. Torr, and A. Zisserman, Eds. *LNCS*, vol. 5302, pp. 234–247. Berlin, Heidelberg: Springer, 2008 (DOI: 10.1007/978-3-540-88682-2_19).

[9] B. Babenko, M.-H. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning", *IEEE Trans. on Pattern Anal. and Mach. Intell.*, vol. 33, no. 8, pp. 1619–1632, 2011 (DOI: 10.1109/TPAMI.2010.226).

[10] K. Zhang and H. Song, "Real-time visual tracking via online weighted multiple instance learning", *Pattern Recogn.*, vol. 46, no. 1, pp. 397–411, 2013 (DOI: 10.1016/j.patcog.2012.07.013).

[11] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection", *IEEE Trans. on Pattern Anal. and Mach. Intell.*, vol. 34, no. 7, pp. 1409–1422, 2012 (DOI: 10.1109/TPAMI.2011.239).

[12] X. Mei, H. Ling, Y. Wu, E. P. Blasch, and L. Bai, "Efficient minimum error bounded particle resampling L1 tracker with occlusion detection", *IEEE Trans. on Image Process.*, vol. 22, no. 7, pp. 2661–2675, 2013 (DOI: 10.1109/TIP.2013.2255301).

[13] C. Bao, Y. Wu, H. Ling, and H. Ji, "Real time robust L1 tracker using accelerated proximal gradient approach", in *Proc. of IEEE Conf. on Comp. Vision and Pattern Recogn.*, Providence, RI, USA, 2012, pp. 1830–1837 (DOI: 10.1109/CVPR.2012.6247881).

[14] W. Zhong, H. Lu, and M.-H. Yang, "Robust object tracking via sparse collaborative appearance model", *IEEE Trans. on Image Process.*, vol. 23, no. 5, pp. 2356–2368, 2014 (DOI: 10.1109/TIP.2014.2313227).

[15] S. A. Wibowo, H. Lee, E. K. Kim, and S. Kim, "Fast generative approach based on sparse representation for visual tracking", in *Proc. of the Joint 8th Int. Conf. on Soft Comput. and Intell. Syst. SCIS and 17th Int. Symp. on Adv. Intell. Sys. ISIS*, Sapporo, Japan, 2016, pp. 778–783 (DOI: 10.1109/SCIS-ISIS.2016.0169).

[16] D. S. Bolme, I. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filter", in *Proc. of the IEEE Conf. on Comp. Vision and Pattern Recogn. CVPR'10*, San Francisco, CA, USA, 2010, pp. 1401–1409 (DOI: 10.1109/CVPR.2010.5539960).

[17] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters", *IEEE Trans. of Pattern Anal. and Mach. Intell.*, vol. 37, no. 3, pp. 583–596, 2015 (DOI: 10.1109/TPAMI.2014.2345390).

[18] S. A. Wibowo, H. Lee, E. K. Kim, and S. Kim, "Multi-scale color features based on correlation filter for visual tracking", in *Proc. of the 1st Int. Conf. on Sig. and Sys. ICSigSys*, Bali, Indonesia, 2017, pp. 272–277 (DOI: 10.1109/ICSIGSYS.2017.7967055).

[19] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking", in *Proc. of the British Mach. Vis. Conf. BMVC'14*, Nottingham, UK, 2014 (DOI: 10.5244/C.28.65).

[20] K. Zhang, L. Zhang, Q. Liu, D. Zhang, and M.-H. Yang, "Fast visual tracking via dense spatio-temporal context learning", in *Computer Vision – ECCV 2014. 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. *LNCS*, vol. 8693, pp. 127–141. Cham: Springer, 2014 (DOI: 10.1007/978-3-319-10602-1_9).

[21] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, and P. H. S. Torr, "Staple: complementary learners for real-time tracking", in *Proc. of the IEEE Conf. on Comp. Vis. and Pattern Recogn. CVPR'16*, Las Vegas, NV, USA, 2016, pp. 1401–1409 (DOI: 10.1109/CVPR.2016.156).

[22] S. A. Wibowo, H. Lee, E. K. Kim, and S. Kim, "Visual tracking based on complementary learners with distractor handling", *Mathem. Probl. in Engin.*, vol. 2017, article ID 5295601, 2017 (DOI: 10.1155/2017/5295601).

[23] D. A. Ross, J. Lim, R. S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking", *Int. J. of Comp. Vision*, vol. 77, no. 1, pp. 125–141, 2008 (DOI: 10.1007/s11263-007-0075-7).

[24] K. Simonyan and A. Zisserman, "Very deep convolutional neural networks for large-scale image recognition", in *Proc. of the 3rd Int. Conf. on Learn. Represent.*, San Diego, CA, USA, 2015 [Online]. Available: https://arxiv.org/pdf/1409.1556.pdf

[25] X. Jia, H. Lu, and M.-H. Yang, "Visual tracking via adaptive structural local sparse appearance model", in *Proc. of the Int. Conf. on Comp. Vis. and Pattern Recogn.*, Providence, RI, USA, 2012 (DOI: 10.1109/CVPR.2012.6247880).

[26] S. Hare, A. Saffari, and P. H. S. Torr, "Struck: Structured output tracking with kernels", in *Proc. of the Int. Conf. on Comp. Vision*, Barcelona, Spain, 2011 (DOI: 10.1109/ICCV.2011.6126251).

[27] F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels", in *Computer Vision – ECCV 2012. 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012. Proceedings, Part IV*, A. Fitzgibbon *et al.*, Eds. *LNCS*, vol. 7575, pp. 702–715. Berlin, Heidelberg. Springer, 2012 (DOI: 10.1007/978-3-642-33765-9_50).

Budi Syihabuddin received his B.Sc. and M.Sc. in Telecommunication Engineering from the School of Electrical Engineering, Telkom University, Bandung, Indonesia, in 2008 and 2012, respectively. In 2010, he joined Telkom University as a lecturer in telecommunication engineering and a microwave laboratory researcher. His interests and publications are in RF microwave devices, wireless communication systems and embedded systems.

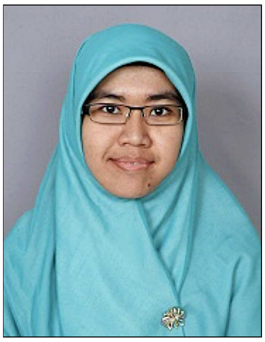
https://orcid.org/0000-0002-2322-2293
E-mail: budisyihab@telkomuniversity.ac.id
School of Electrical Engineering
Telkom University
Bandung, Indonesia, 40257

Suryo Adhi Wibowo received a B.Sc. degree from Telkom Institute of Technology, Indonesia, in 2009, an M.Sc. degree from Telkom Institute of Technology, Indonesia, in 2012, and a Ph.D. from the Department of Electrical and Computer Engineering, Pusan National University, Busan, Korea, in 2018. Now, he is a lecturer in telecommunication engineering and a researcher at the Image Processing & Vision (IMV) Laboratory. His research interests include intelligent systems, computer vision, computer graphics, virtual reality and machine learning.

https://orcid.org/0000-0002-3084-8534
E-mail: suryoadhiwibowo@telkomuniversity.ac.id
School of Electrical Engineering
Telkom University
Bandung, Indonesia, 40257

Agus D. Prasetyo received a B.Sc. degree in 2009 and an M.Sc. degree in 2013, both in Telecommunication Engineering from Telkom Institute of Technology, Bandung, Indonesia. He has been a lecturer in telecommunication engineering and a researcher at the satellite and radar laboratory, Telkom University, Bandung, Indonesia, since 2014. His interests and publications are in electromagnetic devices, antenna design, radar and satellite communication systems.

https://orcid.org/0000-0001-8880-3606
E-mail: adprasetyo@telkomuniversity.ac.id
School of Electrical Engineering
Telkom University
Bandung, Indonesia, 40257



Desti Madya Saputri received a B.Sc. degree in 2009 and an M.Sc. degree in 2012, in Telecommunication Engineering from the School of Electrical Engineering, Telkom Institute of Technology, Bandung, Indonesia. She has been a lecturer in telecommunication engineering and a researcher at the Communication Laboratory, Telkom University, Bandung, Indonesia, since 2012. Her interests and publications are in multiple access, coding theory and signal processing for wireless communications.

https://orcid.org/0000-0002-9200-9816
E-mail: destimadyasaputri@telkomuniversity.ac.id
School of Electrical Engineering
Telkom University
Bandung, Indonesia, 40257

Paper

A Scalable Multicast Routing Protocol for Mobile Ad-Hoc Networks

Liana Khamis Qabajeh

Faculty of Information Technology and Computer Engineering, Palestine Polytechnic University, Hebron, Palestine

https://doi.org/10.26636/jtit.2022.155121

Abstract—The multicasting technique supports a variety of applications that require data to be instantaneously transmitted to a set of destination nodes. In environments with continuously moving nodes, such as mobile ad-hoc networks, the search for efficient routes from sources to the projected destinations is a common issue. Proposed Windmill protocol provides a scalable multicast solution for mobile ad-hoc networks. Windmill aims to improve routing protocol's performance by introducing a hierarchal distributed routing algorithm and dividing the area into zones. Additionally, it attempts to demonstrate better scalability, performance and robustness when faced with frequent topology changes, by utilizing restricted directional flooding. A detailed and extensive simulated performance evaluation has been conducted to assess Windmill and compare it with multicast ad-hoc on-demand distance vector (MAODV) and on-demand multicast routing protocols (ODMRP). Simulation results show that the three protocols achieved high packet delivery rates in most scenarios. Results also show that Windmill is capable of achieving scalability by maintaining the minimum packet routing load, even upon increasing the nodes' speed, the number of sources, the number of group members and the size of the simulated network. The results also indicate that it offers superior performance and is well suited for ad-hoc wireless networks with mobile hosts. The trade-off of using Windmill consists in slightly longer paths – a characteristic that makes it a good choice for applications that require simultaneous data transmission to a large set of nodes.

Keywords—ad-hoc networks, MAODV and ODMRP, position-based multicast routing protocol, simulated performance evaluation.

1. Introduction

A wireless ad-hoc network is a multi-hop self-organizing structure requiring rapid deployment and dynamic reconfiguration [1], [2]. Each participating node has a wireless interface and communicates with other nodes [3]. One of the most important concerns in ad-hoc networks is related to the routing relied upon to forward data packets to the destination [3], [4]. For example, such a network may be implemented to forward packets to students in a university building, soldiers on a battlefield, participants of a conference, and vehicles on the road [3], [5]. The limited number of power nodes and limited bandwidth of the wireless medium require the power consumption and transmission overhead be reduced [1], [2], [4]. Moreover, efficient routing is of key significance, since all nodes act both as hosts and routers and are usually moving rapidly in most cases [6]–[8].

Multicasting is an ideal communication scheme that efficiently supports a wide variety of applications that require collaboration between nodes [9], [10]. Hence, it supports applications that involve simultaneous data transmission to hosts. Military battlefields, disaster recovery, rescue sites and emergency searched are examples of multicast applications for mobile ad-hoc networks [11]. Multicast group members may move, therefore causing random and rapid topology changes at unpredictable times [12]. Thus, tree reconfiguration schemes and membership information logging techniques should be as simple as possible to ensure reduced channel overhead [7], [8], [13]. The constrained power, limited bandwidth, and mobile hosts make the design of a multicast protocol a challenge [14]. Additionally, the need to rely on scalable energy-efficient protocols, along with the existence of inexpensive and low-power positioning instruments, justify the application of position-based routing in mobile ad-hoc networks [7].

In this paper, the Windmill multicast routing model is presented. It introduces a hierarchical distributed routing algorithm to improve performance of the routing protocol and to distribute load by dividing the area into zones. Additionally, the protocol attempts to offer higher scalability, performance and robustness when faced with frequent topology changes, by relying on the idea of restricted directional flooding. Hence, each group member should keep zone leaders (ZL) of its zone updated about its position.

Windmill consists primarily of five phases: network setup, network maintenance and membership update, route instantiation, route maintenance, and, finally, data transmission. The network setup phase includes dividing the area into zones, deciding on initial ZLs, and assigning the interested nodes to different multicast groups. The network maintenance and membership update phase deals with keeping track of the network's structure during node movements and changes.

Whenever a source node has data to be sent to a multicast group, the route instantiation phase is initiated by sending route request packets, mainly with the use of restricted directional flooding. After finishing route discovery and setup, the source begins the data transmission phase by sending the data to the intended destinations. When needed, the route maintenance phase is conducted to repair the broken routes.

We evaluated the performance of the proposed protocol via simulation and compared it with MAODV and ODMRP. Simulation results show that Windmill offers superior performance, regardless of the nodes' mobility speed, the number of sources, the number of group members and network size. Furthermore, Windmill achieved good scalability by maintaining the minimum packet routing load in all presented scenarios, compared to MAODV and ODMRP. The disadvantage of Windmill has the form of slightly longer paths passing through ZLs. Thus, it is suitable for achieving scalability and reducing the overhead of multicast routing in ad-hoc networks established between students of a university, soldiers on a battlefield, rescuers in a disaster area, and sensor-based IoT networks.

The remaining sections of this paper are organized as follows. Related work in presented in Section 2. Section 3 presents the concept behind the Windmill protocol. Section 4 contains a simulated comparison of Windmill, MAODV and ODMRP protocols. Section 5 discusses our findings. The paper is concluded in Section 6, where future directions are discussed as well.

2. Related Work

Multicast routing protocols are classified based on their delivery structure and ability to maintain connectivity between multicast group members. In [9], the authors classified these routing protocols into six categories: flooding, tree-based, mesh-based, hybrid, hierarchical/adaptive multicast, and location-based. Flooding is the easiest way, since it eliminates the need to maintain explicit infrastructure for multicast forwarding. A source initiates a multicast session by broadcasting the packet to its neighbors. Receiving nodes rebroadcast the packet to their neighbors upon receiving the first copy. This process continues until flooding the packet to the whole network. Hence, such a technique offers the lowest control overheads, it is considered to be the most reliable scheme, and data packets are quickly propagated within the network. However, this comes at the expense of generating considerable data traffic in the wireless environment and wasting bandwidth, especially in large networks [15], [16].

In tree-based protocols, the multicast tree is constructed starting from the source of the data and connects all the destinations, i.e. there is only a single path between any source-destination pair. Such protocols are characterized by lower bandwidth consumption than their flooding counterparts. They suffer from low robustness when operating in highly mobile networks, since only a single path between a source-member pair is available. A tree-based protocol can be further categorized into the source-tree and shared-tree varieties. In source-tree protocols, the tree is rooted by the source node itself, whereas in shared-tree protocols, a single tree is shared by all multicast group sources and is rooted at a node known as the core node. The examples of source-tree protocols include the multicast zone routing protocol (MZRP) [17], multicast routing algorithms based on levy flying particle swarm optimization (LPSO) [18], TMRF [19] and multicast opportunistic cooperative routing in mobile ad-hoc networks (MO-CORMAN) [20].

Multicast ad-hoc on-demand distance vector routing protocol (MAODV) [6], shared-tree ad-hoc multicast protocol (STAMP) [21], reliable and energy-aware multicast ad-hoc on-demand distance vector (REA-MAODV) [4], cuckoo search and m-tree-based multicast ad-hoc on-demand distance vector (CS-MAODV) [22] and routing protocol for low-power and lossy networks (RPL) [23] are, in turn, instances of shared-tree protocols.

Mesh-based protocols allow data packets to be forwarded to the same receiver via different paths [24]. Numerous routes between the sender-receiver pair offer better protection against frequent topology changes and increase successful delivery rates [15]. However, the efficiency of mesh-based protocols is lower compared to tree-based protocols, due to multiple routes. Route discovery and mesh building are conducted using broadcasting to discover routes, or using core or central points for mesh building [15]. On-demand multicast routing protocol (ODMRP) [7], core assistant mesh protocol (CAMP) [25], and improved on-demand multicast routing protocol (IODMRP) [26] are examples of mesh-based protocols.

Hybrid multicast protocols combine both tree-based and mesh-based protocols in an attempt to achieve both performance and robustness [15], [16]. Similar to mesh-based approaches, multiple paths are constructed to forward data packets to their destinations. The tree-based approach is used in the route setup process to ensure multicast efficiency. Some examples of hybrid-based protocols include the following: ad-hoc multicast routing protocol (AMRoute) [27], efficient hybrid multicast routing protocol (EHMRP) [28], and zone-based energy aware hybrid multicast routing scheme (ZEHMRP) [29]. Hierarchical routing protocols aim to provide scalability and reduce the number of participating nodes by organizing them into a certain hierarchy. A group of nodes is used to form a cluster or a dominating set of nodes. The examples of cluster-based protocols include LACMQR [30] and EGMP [31]. Adaptive multicast routing protocols adjust their performance taking into account different environmental conditions. For example, the adaptive demand-driven multicast routing protocol (ADMR) [32] is capable of adjusting itself, taking into consideration the mobility state of the network. Once the network mobility level becomes very high, ADMR switches to flooding to overcome link breakages.

Location-based protocols assume that the locations of participating nodes are known. The geographical position of each node is determined using GPS receivers or other positioning services. Moreover, a location service is needed to obtain the positions of destination nodes. Racket forwarding is performed based on the information about the location of the direct neighbor nodes and of the intended destinations. So, nodes offering more efficient progress towards the destinations are selected, resulting in a reduced number of participating nodes. Since location-based multicast routing protocols scale well in large wireless networks, they have recently attracted researchers' attention. The properties of ad-hoc networks, such as constrained power and limited bandwidth, along with the need for scalable and energy efficient protocols, justify utilizing position-based routing in such networks [7]. However, multicasting deals with a group of members and carrying information about the positions of all multicast members in the packet header causes a scalability problem. The positions of a large set of destinations need to be maintained efficiently as well [9]. Some examples of location-based protocols include the following: scalable QoS multicast routing protocol (SQMRP) [33], position-based multicast routing protocol for ad-hoc network using backpressure restoration (PBMRP-BR) [34], scalable and predictive geographic multicast routing scheme in flying ad-hoc networks (SP-GMRF) [35], and location-aware multicasting protocol (LAMP) [36].

It has been observed that most of the existing protocols do not take the issue of scalability into consideration [9]. A crucial problem is that the control overhead may become high if the network is dense, large and/or includes a large number of destinations. Hence, in this research, the scalability and efficiency of multicast routing protocols have been considered.

In such a context, two popular and benchmark protocols have been proposed: MAODV and ODMRP. As the performance of most other protocols is compared to these [9], the rest of this section discusses both protocols in detail.

The MAODV routing protocol [6] uses the broadcast route-discovery approach to discover multicast routes on demand. Nodes participating in the network send a route request (RREQ) packet when they need to join a multicast group, or when they have data to send to a multicast group, and they do not have a route along such a packet could be sent. Only members of the projected multicast group are allowed to respond to a join RREQ. If the RREQ is not a join request, any node having a fresh route to this group can respond. Upon receiving a join RREQ to a group that it is not a member of, or upon receiving a RREQ to a group and not having a route thereto, an intermediate node rebroadcasts this RREQ to its neighbors.

Upon receiving a RREQ packet, the intermediate nodes update their route table. Nodes receiving a join RREQ for a specific multicast group are allowed to reply if they are members of the multicast group tree and the recorded multicast group's sequence number is at least as high as that included in the RREQ. Upon deciding to respond, a node updates its route and multicast route tables by placing the next hop information of the requesting node in the tables. Then, it unicasts a request response (RREP) back to the source node. Upon receiving the RREP, nodes along the path to the source create a forward path by adding a route table along with a multicast route table entry for the node that they received the RREP from.

The source node waits for a specific period of time and enables only the received route with the greatest sequence number and the lowest hop count to the nearest member of the multicast tree. Consequently, it enables the chosen next hop in its multicast route table, then unicasts an activation message (MACT) to the chosen next hop. The next hop, in turn, enables the source node entry in its multicast route table. If this node is a member of the multicast tree, it stops propagating this message. Else, it will have received one or more RREPs from its neighbors. It keeps the best next hop for its route, unicasts MACT to the selected next hop, and enables the correlated entry in its multicast route table. The aforementioned procedure continues until the RREP originating node has been reached. After that, data packets are forwarded only by nodes along the activated routes.

The first member joining the multicast group becomes the group leader. This leader maintains the multicast group sequence number and broadcasts it to the group members via a group hello message. The nodes use the group hello information for updating their request tables. Furthermore, MAODV has to actively track and react to changes in the tree resulting from membership changes and node movements.

ODMRP [7] uses a mesh-based approach. Hence, the multicast tree's drawbacks, such as alternating connectivity, frequent tree reconfiguration, and non-shortest path in a shared tree, are avoided [7]. In ODMRP, the multicast packets are forwarded only by a subset of nodes via scoped flooding. It conducts on-demand procedures to dynamically maintain multicast group membership and build routes. When a source has data to be sent and no already-chosen routes to the group members are available, the source broadcasts a join-query packet to the entire network. Join-query packets are broadcast periodically to update membership information and refresh the routes.

Backward learning for the reverse path back towards the source is used, i.e. routing tables are updated with the appropriate ID of the node from which the message was received. The message is rebroadcast if it is induplicate and TTL is larger than zero.

Upon receiving a join-query packet, a multicast receiver creates and broadcasts a join-reply to its neighbors. Once a node receives a join-reply, it checks if the next hop node ID of one of the entries is the same as its own ID. If so, the node realizes that it is a part of the forwarding group. Hence, it sets the FGFLAG. Accordingly, it broadcasts its join table built upon the matched entries. The next hop node ID field is filled by getting information from the nodes' routing tables. Thus, each forward group member propagates the join-reply until reaching the source via the designated shortest path.

After completing the route construction process and establishing the forwarding group, sources can multicast packets to the receiving nodes via these routes. While the source has data to be sent, it periodically sends join-query packets to keep the forwarding group and the routes fresh. A node forwards a data packet only if it is not a duplicate and the setting of the multicast group FGFLAG has not expired yet. This procedure reduces the traffic overhead and avoids sending packets over expired routes.

ODMRP adopts the soft state approach to maintain multicast group members. Hence, no explicit control packets are sent to leave a multicast group. When a source is about to leave the group, it simply stops sending join-query packets, as it no longer has data to be sent. If a receiver is no longer interested in a particular group, it does not send the join-reply for that group. Nodes in the forwarding group are treated as non-forwarding nodes if not refreshed before they timeout. The relaxed connectivity makes ODMRP more stable for mobile wireless networks [7].

3. Windmill Protocol

The proposed Windmill protocol consists primarily of five phases: network setup, network maintenance and membership update, route instantiation, route maintenance and, finally, data transmission. Table 1 presents the variables and notations used in further discussions.

Windmill assumes that NN cooperative nodes are distributed randomly in a square-shape area and are aware of their positions. During the network setup phase, the nodes collaborate to divide the area into zones and elect an initial ZL for each zone. After that, communication between ZL and the nodes interested in joining a specific group is conducted. In Table 2, the packets exchanged during the Windmill network setup phase are summarized. Upon using RDF, each node receiving a packet forwards the packet only if it is closer to the destination node than its previous

Table 1
Variables and notations used

Notation	Description	Notation	Description
NN	Number of nodes	ZL	Zone leaders
DS_n	Distance between the node and the center of its zone	DS_m	Maximum possible distance between a node and the center of a zone
SP_n	Node speed	SP_m	Maximum possible node movement speed
BT_n	Remaining battery life (time) of a node *n*	BTm	Maximum possible battery life (time)
CP_n	CPU processing power of a node *n*	CPm	Maximum CPU power available
MM_n	Memory usage of a node *n*	MM_m	Maximum memory capacity available
IP_n	IP address of node *n*	SN_n	Sequence number issued by node *n*
Pos_n	Position of node *n*	GID	Group number
$Z_{[x,y]}$	Zone number *x*, *y*	$ZL_{[x,y]}$	Zone leader of zone number *x*, *y*
D_{mov}	Movement distance allowed before sending PosUpdate	$PosZL_{[x,y]}$	Position of ZL of zone number *x*, *y*
DD	Distance between the forwarding node and the destination	D_{TH}	Number of destination nodes in a zone deciding to use RDF or ZBrd
RDF	Restricted directional flooding	ZBrd	Zone broadcast
$ProbL_{nz}$	Probability of node *n* being selected as ZL for its zone *z*	D_{cen}	ZL distance allowed from the zone center before sending ZLElect

Table 2
Packets sent during the network setup phase

Packet identifier	Stands for	Description
ZLProb	ZL probability	• Contains probability of a node to be elected as ZL of its zone • Sent from each node in a specific zone to nodes inside that zone, i.e. ZBrd
ZLPos	ZL position	• ZL of a zone to inform other nodes in it zone about its position • Sent using ZBrd
JoinGroup	Join group	• Nodes in a specific zone to ZL of that zone to inform it that they are interested in joining a specific group • Sent using RDF
PosUpdate	Position update	• Nodes in a specific zone to ZL of that zone to inform it about their position • Sent using RDF

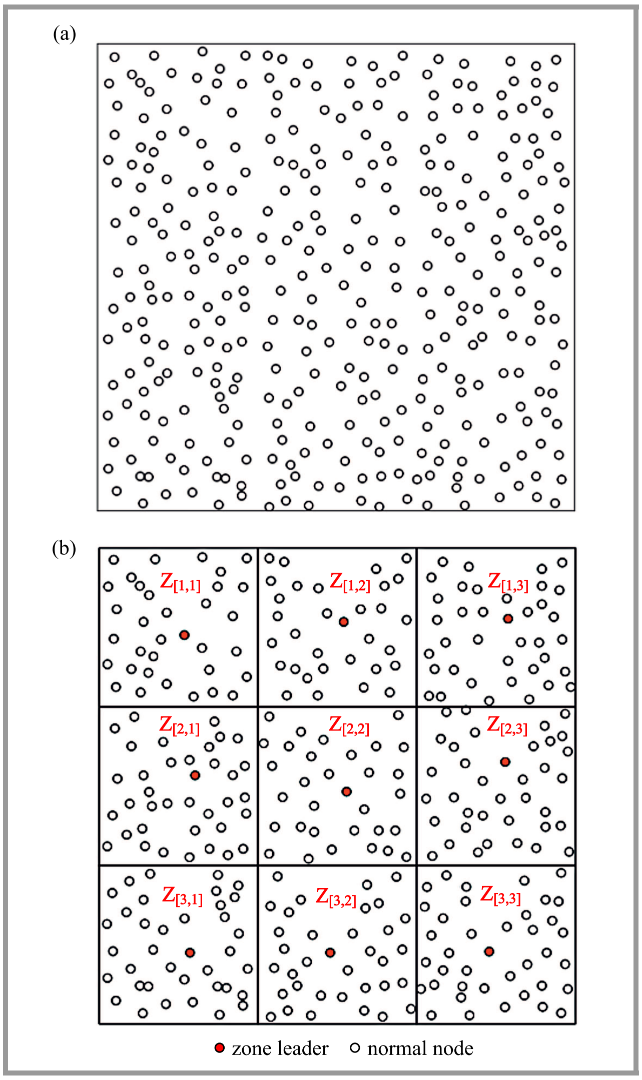


Fig. 1. Network structure at the beginning of the setup phase (a) and after the network setup phase, after division of the area and selection of ZLs (b).

hop. Upon using ZBrd, the nodes process a packet and forward it only if they are within the intended zone.

At the beginning of the network setup phase, the network's area is divided into numerous equal-size square-shape zones and initial ZLs for different zones are elected. Figure 1 shows the network structure at the beginning of the network setup phase, as well as after dividing the area into nine (3 × 3) zones and electing ZLs.

Each node knows the zone it belongs to using its position, the area coordinates, and the number of zones. Node position is known via GPS, while the area coordinates and the number of zones are stored in each node before deployment.

After dividing the area into zones, nodes inside each zone will start electing a ZL. The ZLs are chosen to be near the zone's center, in order to make sure that the time needed for communication between ZL and any node inside the zone is almost the same. Next, each node n inside a zone $Z_{[x,y]}$ is assigned a weight representing its probability of being the ZL of a particular zone. The most important aspects taken into consideration while selecting ZLs are the distance between the node and the center of the zone that the ZL will be responsible for DS_n, the node's speed SP_n and battery remaining life time BT_n. Choosing a ZL that is close to the center of the zone boundary and moving with a low speed increases the probability of the communication between ZLs of different zones being performed in one hop, which helps protect important packets. Choosing ZLs with low movement speeds also increases the probability that the elected ZL will stay in the zone longer, and so there is no need to re-elect a new ZL within a short period of time. Moreover, choosing a node with a high remaining battery life time reduces the likelihood the battery being drained, i.e. reduces the probability of electing a new ZL and transferring important and secure information in its possession.

Two other important factors that should be taken into consideration when electing a ZL are the CPU processing power CP_n and memory capacity MM_n of the nodes. ZLs with high CPU processing power and large memory significantly affect network performance, since these ZLs may be the bottleneck of the position management scheme.

Each node inside a specific zone uses these factors to calculate the probability of itself being elected as a ZL for a specific zone. Probability $ProbL_{nz}$ of node n in zone z being elected as a ZL for that zone is:

$$ProbL_{nz} = 0.2 \times \left(1 - \frac{DS_n}{DSm}\right) + 0.2 \times \left(1 - \frac{SP_n}{SPm}\right) + 0.2 \times \left(\frac{BT_n}{BTm}\right) + 0.2 \times \left(\frac{CP_n}{CPm}\right) + 0.2 \times \left(\frac{MM_n}{MMm}\right) , \quad (1)$$

where: DS_m is the maximum possible distance between a node and the center of a zone, SP_m is the maximum possible node movement speed, BT_m is the maximum possible battery life time, CP_m is the maximum CPU power available, MM_m is the maximum memory capacity available.

Values of the weights of different parameters are chosen equally, since we believe that they are all important when selecting the ZL. DSm is considered to be the distance between two opposite corners of a zone. SP_m is a predefined value that depends on the environment in which the protocol is deployed. BT_m, CP_m and MM_m depend on the current technology found in the market.

After calculating its probability of being elected as a ZL, each node sends a ZLProb message to other nodes in its zone using zone broadcast ZBrd. Upon receiving the packet, each node will process it only if it is in the intended zone $Z_{[X,Y]}$. Otherwise, the packet is dropped. The node with the highest probability in each zone will be the ZL of that zone. At this step, we assume that the network is error free and so all nodes within a specific zone receive the same set of ZLProb messages. Now, the ZL node sends the ZLPos message to inform other nodes in its zone about its position. This message is also sent using ZBrd.

After that, only the interested nodes send JoinGroup and PosUpdate messages to the ZL of their zone to inform it that they are interested in joining a specific group and to tell it about their positions. These packets are sent via

Table 3
Packets sent during the network maintenance and membership update phase

Packet identifier	Stands for	Description
ZLElect	ZL election	• Sent from ZL of a specific zone to nodes inside that zone to initiate a new ZL election process • Sent using ZBrd
ZLQuery	ZL query	• Sent by a node entering a new zone to ask about its ZL • Sent to first hop neighbors
LeaveGroup	Leave group	• Nodes in a specific zone to ZL of that zone to inform it that they are no longer interested in a specific group or when they are leaving the zone • Sent using RDF
ZLProb, ZLPos, JoinGroup, and PosUpdate		• As explained in the network setup phase

Table 4
Packets sent during the route discovery phase

Packet identifier	Stands for	Description
SRREQ	Source route request	• Request sent from the source node using RDF to local ZL to ask about destination nodes for the multicast session to be held • Sent using ZBrd
IRREQ	Internal route request	• Request sent from a specific ZL to the interested local destinations to join the multicast session held • This packet is sent using ZBrd if the number of destination nodes in this zone is greater than D_{TH}, else it is sent using RDF towards each destination
ERREQ	External route request	• Request sent from ZL of a given zone to neighbor ZLs using RDF, to ask about destination nodes in the neighbor zone that are interested in joining the multicast session

RDF. The use of RDF offers a high probability of finding a path compared to the greedy solution. Such an approach also reduces the resulting overhead compared with blind broadcasting to the entire network.

3.1. Network Maintenance and Membership Update

During the network lifetime, nodes may move freely within the network, may move in and out of the network and change their group membership. The proposed protocol tries to cope with these issues. In Table 3, the packets exchanged during the network maintenance and membership update phase of Windmill are summarized.

Let us start with non-ZL nodes. Members joining a specific group can leave it by sending a LeaveGroup packet to the ZL of their zone. Moreover, any node can send JoinGroup and PosUpdate messages to its zone ZL if it becomes interested in a specific group. These packets are sent via RDF and contain the same fields as described in the network setup phase.

Member nodes should also inform their ZLs about their new position if they have moved a predefined distance D_{mov} from their last known position. When a specific member is about to leave the boundaries of its zone, it should send a LeaveGroup message to the previous ZL. Then, it sends a ZLQuery packet to ask about the ZL of the new zone. This packet is sent to first hop neighbors and any node in new zone may reply by sending ZLPos packet containing the IP and position of the responsible ZL. Now the moving node can communicate with the new ZL by sending JoinGroup and PosUpdate messages.

Regarding ZL nodes, a ZL sends a ZLPos message to inform other nodes in its zone about its new position if it has moved D_{mov} from its last known position. This message is sent using ZBrd.

If the ZL decided to depart its zone, its distance from the zone center became higher than a pre-defined distance D_{cen}, or if its battery is about to turn off, it may send a ZLElect packet to initiate a new ZL election. This packet is sent using ZBrd. Upon receiving this packet, each node inside the zone will calculate its probability to become a ZL and a new ZL will be elected, as discussed in the network setup phase.

3.2. Discovery Phase

Table 4 presents the control packets exchanged to handle the route discovery algorithm. When a source node decides to initiate a multicast session, a source route request (SRREQ) packet is first directed to its local ZL node to ask for possible participating nodes in the multicast session held. The SRREQ packet continues to be propagated restrictedly using RDF, until it reaches the intended ZL.

When the source ZL node receives the SRREQ packet, it sends an external route request (ERREQ) packet to the four neighbor ZLs. Here, we consider the case that ZLs of adjacent zones may not be within the transmission range of each other. Hence, multi-hop routing is assumed and packets are sent across zones using RDF.

The source ZL also sends an internal route request (IRREQ) packet only if there are interested nodes within this zone. This packet is sent trying to find routes to the participating nodes within this zone. Upon receiving the packet, each node will process it only if it is in the intended zone $Z_{[X,Y]}$. Otherwise, the packet is dropped. This packet is sent using ZBrd if the number of destination nodes in this zone is greater than D_{TH}. In the zone broadcast, upon deciding to forward the packet, the node stores the IP address of its previous hop IPI to be used in the reverse path. Also, it alters the IPI field to be its own IP address and proceeds with forwarding the packet. On the other hand, if the number of destination nodes in this zone is lower than or equal to D_{TH}, RDF will be used. In this case, $ZL_{[X,Y]}$ will prepare a separate packet for each destination, and each node processing the packet will forward it only if it is closer to that destination.

Upon receiving ERREQ for the first time, the intended neighboring ZL continues the route discovery process by finding a route between itself and the neighbor ZLs (by sending ERREQ), and later between itself and other destinations in its zone (by sending IRREQ). The ERREQ packet is propagated until it reaches all the network zones using the forwarding strategy, as discussed later on.

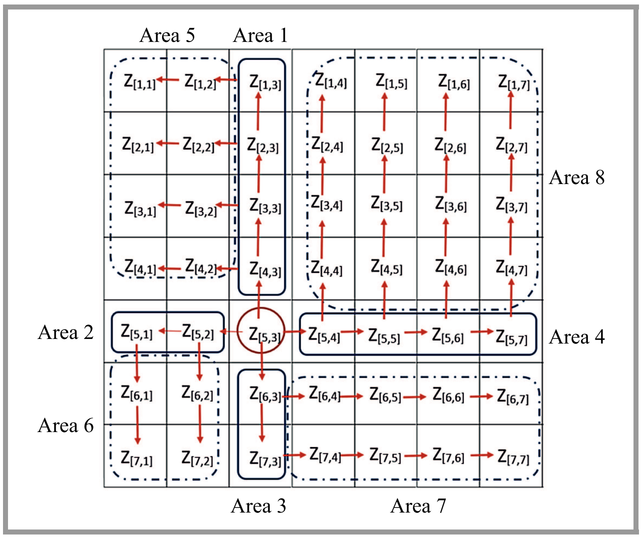


Fig. 2. Forwarding ERREQ packet in Windmill protocol.

The proposed protocol utilizes the network division to forward the ERREQ packets to discover the anticipated group members with very low overhead, as well as to prevent sending duplicate packets. In this subsection, the forwarding of ERREQ packets between the network zones is explained – see Fig. 2. The decision to forward the ERREQ packet to the neighbor zones is the responsibility of the ZL node.

The source node resides in zone $Z_{[5,3]}$. Firstly, the ERREQ packet is forwarded towards the border of the four neighbor zones as the first forwarding step (in our example, there are zones $Z_{[4,3]}$, $Z_{[5,2]}$, $Z_{[6,3]}$ and $Z_{[5,4]}$ are present).

If each zone receiving the ERREQ packet resends it to all 4 of its neighbors, meaning that a lot of duplicate packets are produced. To overcome this, an efficient forwarding strategy is proposed. This algorithm enables the ZL of each zone to take part in delivering the packet to two neighbor zones at the most. In this forwarding scheme, the ZL is based on the number of the source zone $Z_{[X,Y]}$, and the coordinates of the intermediate zone that is currently forwarding the packet $Z_{[x,y]}$. This forwarding strategy ensures that the ERREQ packet is propagated through the network with no duplicates and all the network zones are visited only once (see to Fig. 2).

For example, assume that the packet is sent out from zone $Z_{[5,3]}$. Here, the ZL node of zones $Z_{[4,3]}$, $Z_{[3,3]}$, $Z_{[2,3]}$ and $Z_{[1,3]}$ (area 1) forwards the packet to the zones that are above and to the left of the current zone (if any). In the following step, zones $Z_{[1,2]}$, $Z_{[2,2]}$, $Z_{[3,2]}$, $Z_{[4,2]}$, $Z_{[1,1]}$, $Z_{[2,1]}$, $Z_{[3,1]}$ and $Z_{[4,1]}$ (area 5) send the packet only towards zones to their left (if any). A similar strategy is used for packets forwarding to other network parts to eliminate duplicate packets.

The pseudocode of the forwarding strategy is illustrated below, considering that the source zone is $Z_{[X,Y]}$, and the current zone to forward the packet is zone $Z_{[x,y]}$:

- if $x = X$ and $y = Y$ (source zone), then forward to zones $Z_{[X-1,Y]}$, $Z_{[X,Y-1]}$, $Z_{[X+1,Y]}$ and $Z_{[X,Y+1]}$,
- if $y = Y$ and $x < X$ (area 1), then forward to zones $Z_{[X,Y-1]}$ and $Z_{[X-1,Y]}$,
- if $x = X$ and $y < Y$ (area 2), then forward to zones $Z_{[X,Y-1]}$ and $Z_{[X+1,Y]}$,
- if $y = Y$ and $x > X$ (area 3), then forward to zones $Z_{[X,Y+1]}$ and $Z_{[X+1,Y]}$,
- if $x = X$ and $y > Y$ (area 4), then forward to zones $Z_{[X,Y+1]}$ and $Z_{[X-1,Y]}$,
- if $x < X$ and $y < Y$ (area 5), then forward to zone $Z_{[X,Y-1]}$,
- if $x > X$ and $y < Y$ (area 6), then forward to zone $Z_{[X+1,Y]}$,
- if $x > X$ and $y > Y$ (area 7), then forward to zone $Z_{[X,Y+1]}$,
- if $x < X$ and $y > Y$ (area 8), then forward to zone $Z_{[X-1,Y]}$.

Figure 3 shows the control packets exchanged during the route discovery phase of the Windmill protocol.

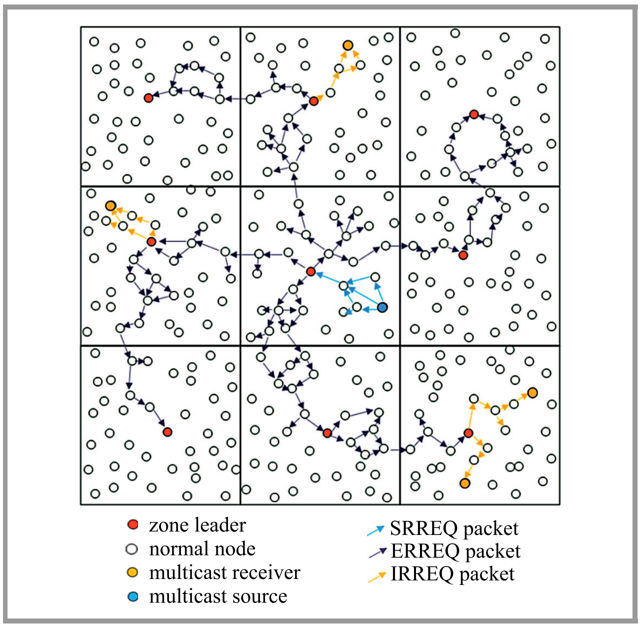


Fig. 3. Packets sent during the route discovery phase.

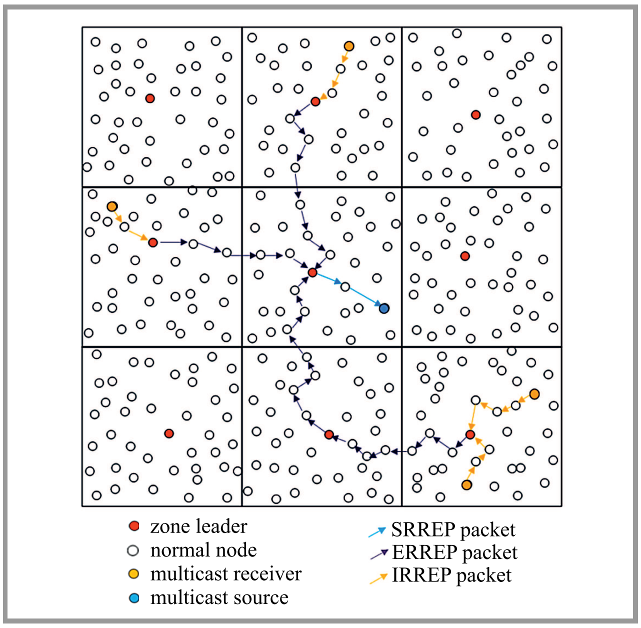


Fig. 4. Packets sent during the route setup phase.

3.3. Route Setup

The next step, after propagating the request packets, is to setup the routes by sending the reply packets. Table 5 contains the control packets exchanged to handle the route setup phase.

After forwarding the IRREQ packet and if it is interested in participating in the session, node *J* commences the process of setting up a route from the local ZL to itself by sending an internal route reply (IRREP) packet. Each intermediate node forwards this packet to the node from which it received the corresponding IRREQ packet. This process continues until the packet reaches the intended ZL.

To reduce the network overhead, each zone leader $\mathrm{ZL}_{[x,y]}$ sends only one external route reply (ERREP) to the neighbor ZL that forwarded the original ERREQ to it. This packet is sent using the reverse path, until the ZL node that issued the original ERREQ packet is reached.

To further reduce the overhead in the network, the source zone leader $\mathrm{ZL}_{[X,Y]}$ sends only one source route reply (SRREP) to the source node *S*. Each node sends this packet to the previous hop from which it received the original SRREQ packet, until the packet reaches node *S*.

Figure 4 shows the control packets exchanged during the route setup phase.

3.4. Route Maintenance

During data transmission, some nodes may not receive data packets due to broken links caused by failure or movement of the nodes. When a link break is detected, the node located upstream of the broken link sends a route error (RERR) packet backwards to the upstream nodes to inform them about this failure. Intermediate upstream nodes, upon receiving this packet, clear the information related to the downstream nodes, and re-forward the packet towards their upstream nodes. Also, the nodes located downstream of the broken link will clear the related entries and free the resources when a predefined time has elapsed without receiving data from the upstream nodes.

When a ZL receives the RERR packet, it deletes the related entry from its routing table and initiates a new route discovery process towards the affected destinations. Also, if the source receives a RERR packet, it discovers that the link between itself and the local zone leader is no longer

Table 5
Packets sent during the route setup phase

Packet identifier	Stands for	Description
SRREP	Source route reply	• Reply sent from ZL of the zone of the source node indicating that there are nodes in the source zone want to join the multicast session held
IRREP	Internal route reply	• Reply from a given node to its local ZL setting up a route to itself • Nodes reply to the first IRREQ they receive
ERREP	External route reply	• Reply sent from ZL of a given zone to the ZL of the zone from which it received the ERREQ packet. This packet indicates that there should be a route passing through this ZL

Table 6
Packets sent during the route maintenance phase

Packet identifier	Stands for	Description
RERR	Route error	When a broken link is encountered during data transmission, the node that discovers the broken link informs its upstream nodes about this failure using RERR packet

available. Accordingly, the source node deletes the related entry from its routing table and initiates a new route discovery process to reconstruct the broken route towards the local ZL. Table 6 shows the control packet exchanged to ensure route maintenance.

3.5. Data Transmission

The source node waits for a predefined time to setup the routes to the nodes that want to participate in the multicast session. Then, it starts sending data packets to the multicast group members using the chosen routes. The multicast data packets are sent along the multicast tree, from the source to the ZL nodes. Whenever a data packet reaches the ZL nodes, the ZL nodes forward a copy of the received data packet to the members in their zone. Each intermediate node simply re-forwards data packets to its successor in the route determined during the route initiation process.

4. Performance Evaluation

In this section, a simulated performance evaluation of MAODV, ODMRP and Windmill is presented. MAODV and ODMRP protocols are considered for comparison purposes, since they were proposed by the mobile ad-hoc networks working group at the IETF and are often considered as benchmarks for evaluating performance of ad-hoc multicast routing protocols [8].

Global Mobile Simulation (GloMoSim) [37] is used as a simulation tool to evaluate the performance of the three protocols under consideration. A network with 60 mobile nodes located within an area of 1000 m × 1000 m that is divided into 4×4 zones is considered. The nodes' transmission range of 250 m and channel capacity of 2 Mbit/s are used. The initial positions of the nodes are chosen randomly. After that all nodes are allowed to move in accordance with the random waypoint mobility principle, i.e. each node travels to a randomly selected location at a configured speed and then pauses for a configured pause time, before choosing another random location and repeating the same steps. A pause time between 0 and 10 s is simulated. The maximum node mobility speed is 40 km/h.

The 802.11 MAC layer and constant bit rate (CBR) traffic over user datagram protocol (UDP) have been used. For either protocol, a routing packet processing delay of 1 ms is assumed. In order to minimize collisions, a random delay between 0 and 10 ms is introduced before retransmitting the broadcast packets. Sources and destinations are chosen randomly. One multicast group with a single source and 20 members is simulated. The source sends data at the rate of 20 packets/s. The size of data payload is 512 bytes. Multicast group members are allowed to join and leave the multicast group at any time during the simulation. Member nodes are selected randomly with uniform probabilities. Each simulation is performed for 300 s.

4.1. Performance Metrics

Five important parameters related to ad-hoc network multicast transmissions have been tested. These parameters include the following: node mobility speed, number of sources, multicast group size (members), network size and number of zones. For each parameter, five performance metrics are evaluated. The metrics were derived from the ones suggested by the IETF mobile ad-hoc network working group for the purpose of evaluating routing/multicast protocols [38]:

1. **Packet delivery fraction (PDF)**. The ratio of the number of data packets actually really delivered to the multicast receivers versus the number of data packets supposed to reach them. This evaluates the protocol's ability to discover and maintain routes, as well as its effectiveness in delivering data to the intended receivers.
2. **Number of control packets transmitted per data packet delivered (CPD)**. Instead of using a pure control overhead, we choose to use a ratio of control packets transmitted to data packets delivered in order to investigate how efficiently control packets are utilized in delivering data to the intended receivers. Packets used for route instantiation and maintenance are considered upon calculating this metric. Furthermore, packets sent to construct and maintain the network's structure, update node positions and maintain membership are considered as well. The transmission at each hop along the paths is included in the calculation of this metric.
3. **Number of control and data packets transmitted per data packet delivered (CDPD)**. This metric shows the efficiency in terms of channel access and is very important in ad-hoc networks, since link layer protocols are typically contention-based.
4. **Average path length (APL) [hop]**. The average length of the paths discovered by the protocol. It is calculated by taking the average number of hops taken by each data packet to reach the destination.
5. **Average route latency (ARL) [ms]**. The average delay needed for discovering a route to the destination. It is defined as the average delay between sending

a route request/discovery packet by a source and receiving the first corresponding route reply packet. If a request is timed out and requires to be retransmitted, the sending time of the first transmission is used in calculating the latency.

Each point in the following figures is obtained by averaging the results of five simulation runs with similar configurations but various, randomly generated numbers.

4.2. Node Mobility Speed Effect

The node mobility speed has been varied to evaluate the ability of the protocols to deal with route changes. Figure 5a shows the PDF of the three protocols as a function of mobility speed. ODMRP is more effective than AODV and Windmill in PDF, as the maximum node speed is increased from 0 to 80 km/h. This is caused by ODMRP mesh topology which allows for alternative paths and makes ODMRP more robust compared to MAODV and Windmill which rely on a single path in their multicast tree. PDF for the three protocols decreases with increasing mobility speed, due to higher probability of link breakages and data packet drops.

Since most ad-hoc network medium access control protocols are contention based, having less packets transmitted per data packet delivered is very important [7]. As shown in Fig. 5b-c, CPD and CDPD for ODMRP are higher than those for MAODV and Windmill. The increased ODMRP

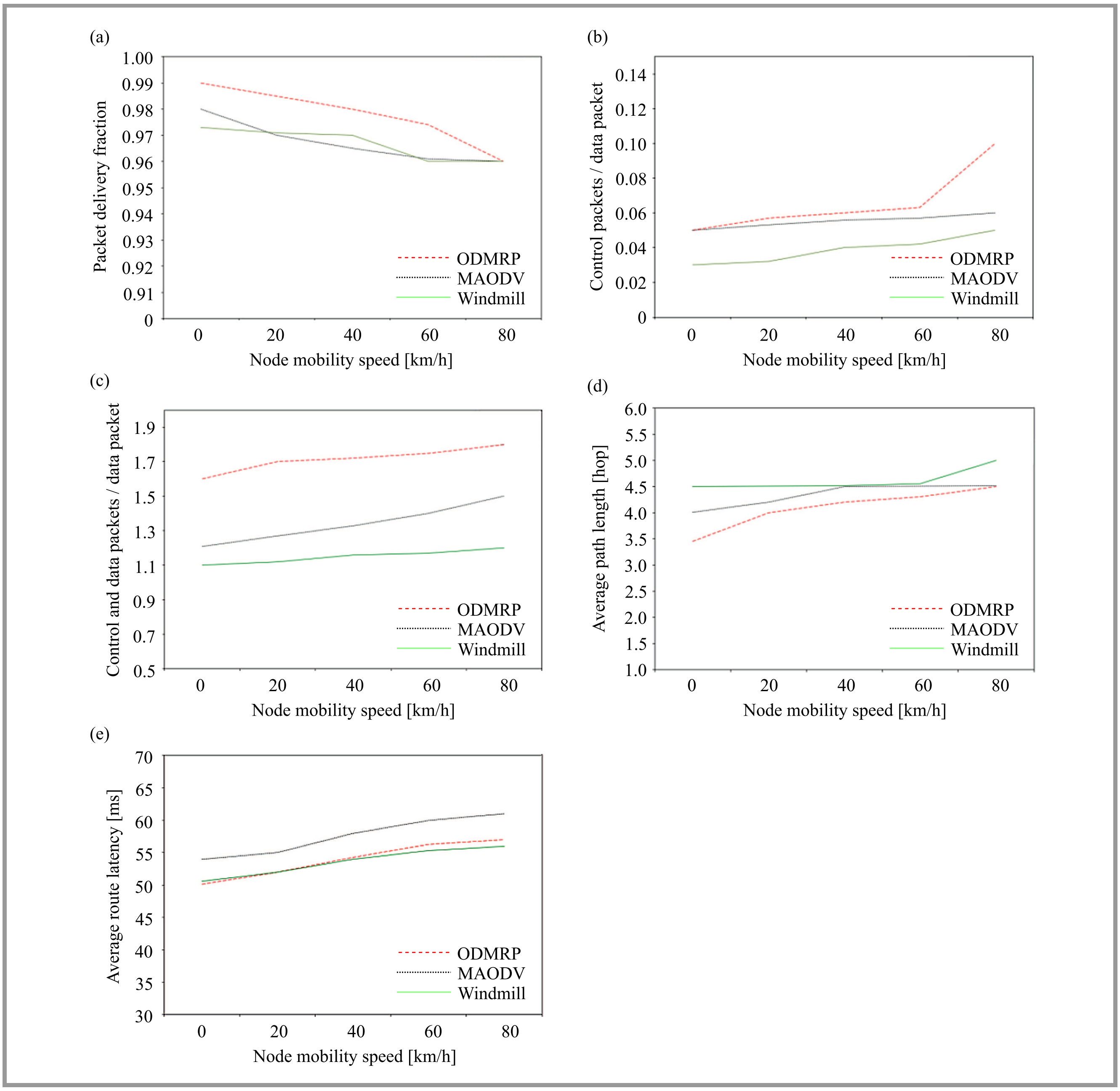


Fig. 5. Node mobility speed simulations.

CPD is due to its broadcast of the route request and reply and the periodic refresh of routes from source to different destinations. Moreover, bidirectional trees used in MAODV and Windmill are more efficient, compared to mesh, and avoid sending numerous copies of data packets to receivers, i.e. lower MAODV and Windmill CDPD.

In MAODV, the multicast group leader maintains updated multicast tree information by sending periodic group hello messages. Windmill, on the contrary, does not require sending periodic group hello messages. Moreover, MAODV sends the request packet to the entire network, whereas in Windmill, the request packets between zones are sent using RDF, and RDF or ZBrd are used only inside zones having destinations inside them. These two points justify the lower value of CPD of Windmill compared to MAODV. As far as network structure maintenance is concerned, in Windmill, the process of dividing the area into zones and initial ZL election is conducted once, at the beginning of the network setup phase. After that, any updates such those concerning nodes joining and leaving groups, position updates and new ZL election processes, are performed locally, inside the intended zone and most properly using RDF. Hence, the impact of network structure maintenance group membership on the control overhead is not noticeable. CPD and CDPD for the three protocols slightly increase with an increase in mobility speed, due to higher probability of link breakages and route repairs.

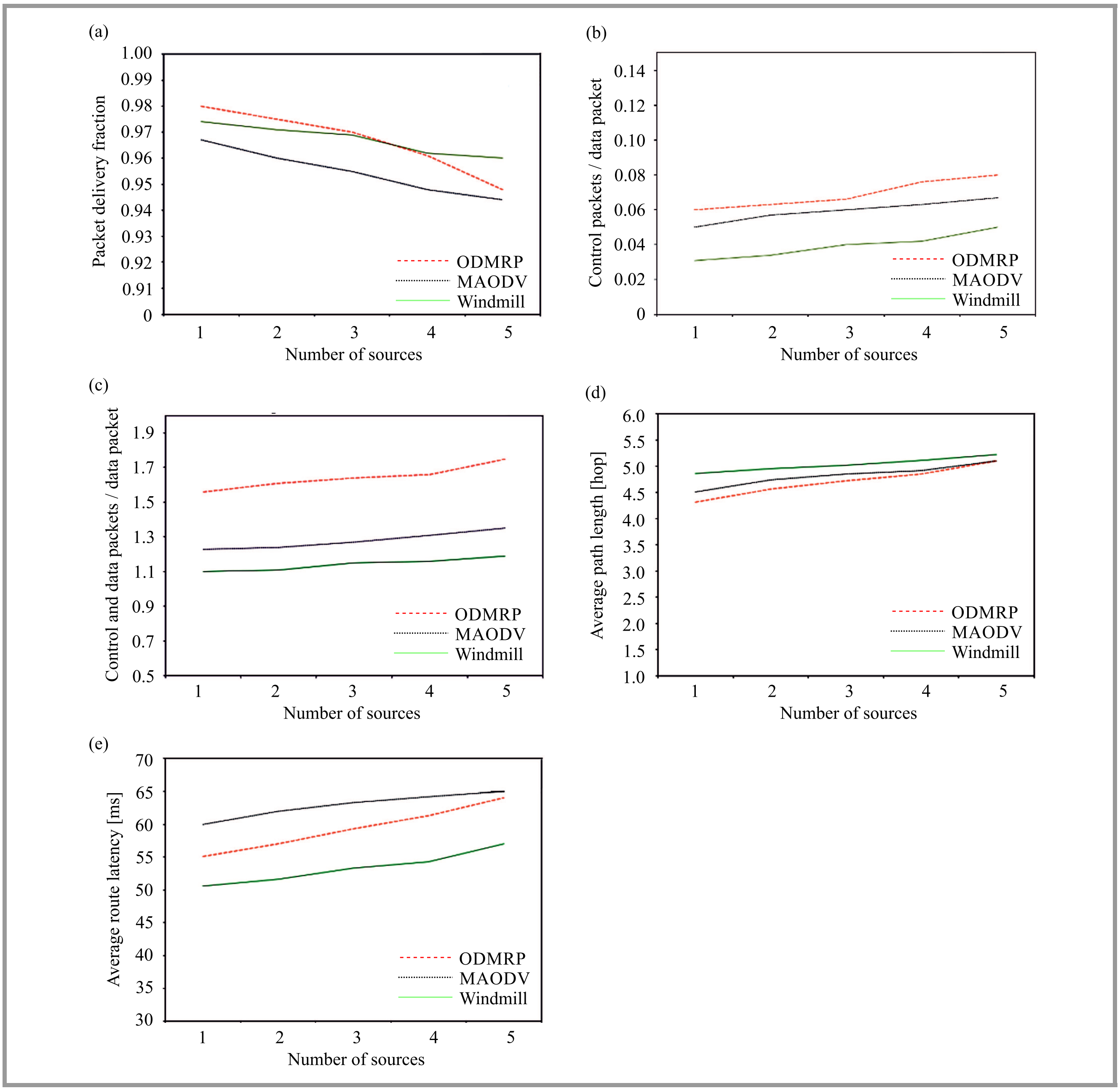


Fig. 6. Number of sources.

Regarding APL of the selected routes, Fig. 5d shows that routes in Windmill are a little bit longer than those in MAODV and ODMRP, since the routes are forced to pass through ZLs.

Figure 5e demonstrates that ARL of MAODV is higher than in the case of two other protocols. MAODV does not activate a multicast route immediately. A potential multicast receiver waits for a specified period of time, allowing to receiving numerous replies before sending an activation message along the chosen multicast route. On the contrary, ODMRP and Windmill activate the routes immediately. Moreover, ARL of the proposed protocol is a little bit lower than that of ODMRP, since the number of request and reply packets received by each node in Windmill is lower, which reduces the time spent by these nodes on processing these packets.

APL and ARL for the three protocols slightly increase along with increasing mobility speed, due to the higher probability of link breakages and choosing other, longer routes.

4.3. Number of Sources Effect

Next, the number of senders in the multicast group has been varied in order to evaluate the scalability of different protocols with respect to source nodes and the resulting effective traffic load. Figure 6a presents the PDF of the

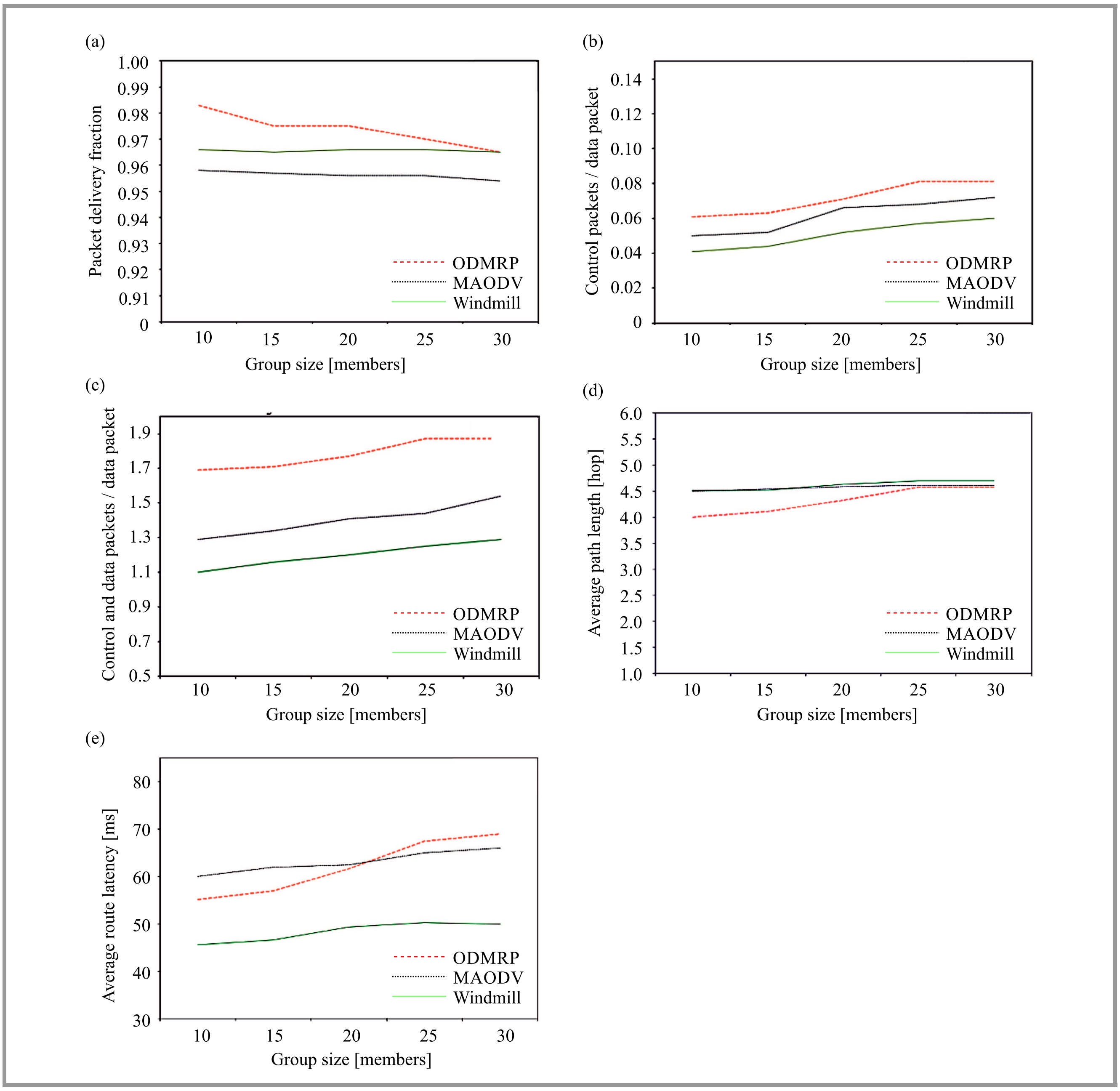


Fig. 7. Simulation results with varying group sizes.

three protocols as a function of the number of senders. ODMRP is more effective compared to MAODV in PDF when the number of senders is low due to its mesh topology. However, upon increasing number of senders from 1 to 5, ODMRP in particular does not scale well for PDF. In ODMRP, every source node sends out, periodically route requests through the network. When the number of source nodes becomes larger, this causes congestion in the network and the PDF drops significantly. MAODV, on the other hand, maintains only one multicast group leader that periodically sends group hellos through the network. Therefore, MAODV is more scalable compared to ODMRP [8]. As far as the proposed protocol is concerned, no periodic packets are sent. However, the network structure will be constructed for each source, as this justifies the moderate decrease of PDF in Windmill.

Figures 6b-c show that CPD and CDPD for the three protocols increase slightly with the increasing number of senders, due to congestion resulting from packets being sent periodically in ODMRP and MAODV, and from the network structure of Windmill. This congestion also justifies the increase in ARL shown in Fig. 6e.

Regarding APL of the selected routes, Fig. 6d shows that in Windmill are a little bit longer than those in both MAODV and ODMRP, since the routes are forced to pass through ZLs. However, this difference decreases as the number of sources increases, due to congestion forcing MAODV and ODMRP to choose longer paths.

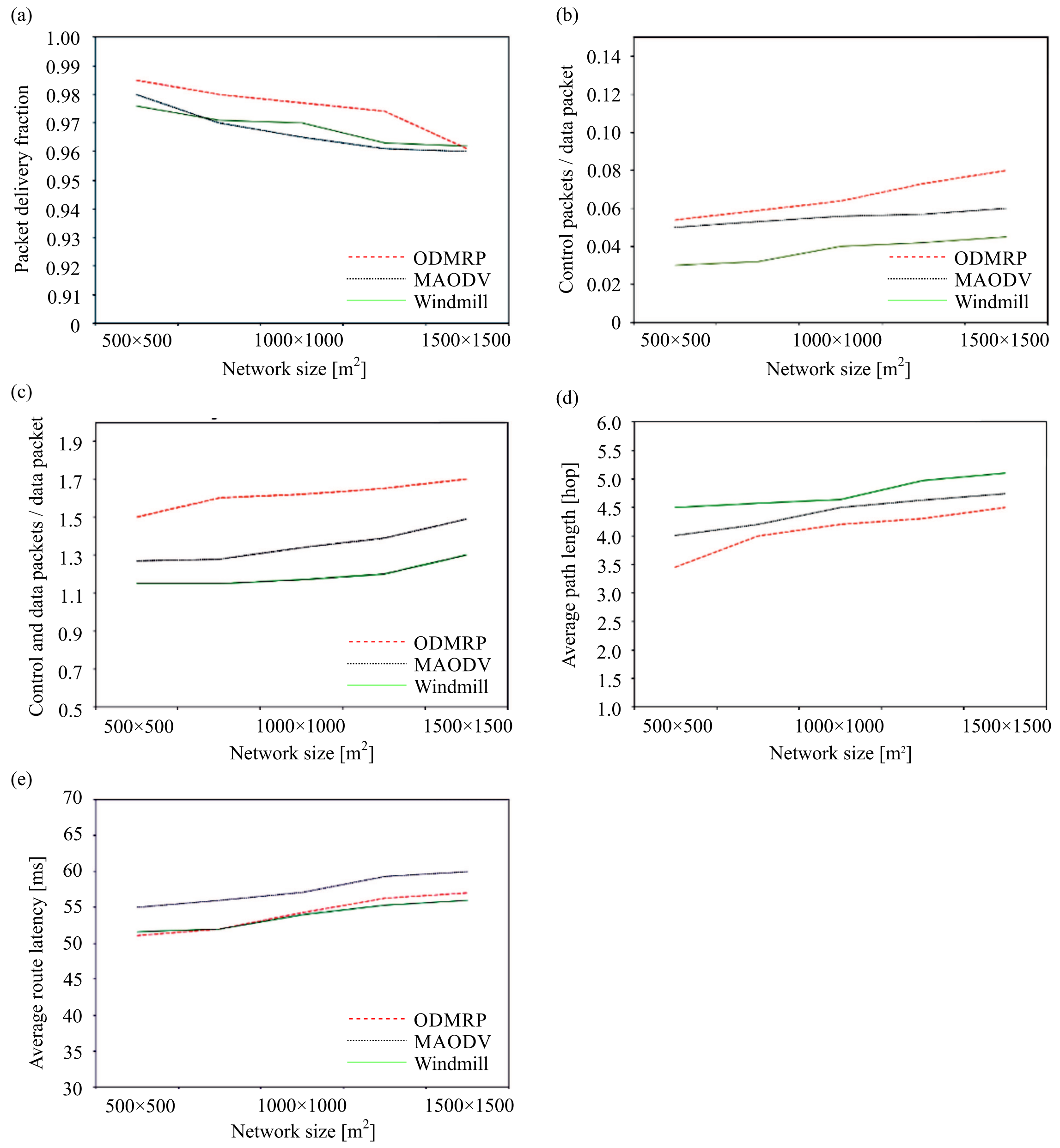


Fig. 8. Simulations with a varying network size.

4.4. Multicast Group Size Effect

In this scenario, the size of the multicast group is varied to examine the scalability of the protocol regarding number of members. Figure 7a shows that ODMRP is more effective than MAODV regarding PDF when the number of multicast group members is low. However, ODMRP does not scale well with multicast group size. There is a noticeable decline in PDF as the multicast group increases to 30 members. This can be attributed to collisions that occur from the frequent broadcasts through the network [8].

MAODV and Windmill scale better in terms of CPD and CDPD compared to ODMRP (Figs. 7b–c). This is due to ODMRP broadcast of route request and reply packets, periodic refresh of routes from the source to different destinations, and sending multiple copies of data packets to receivers. This increased number of packets also contributes to an increase in ARL and APL due to the time spent by participating nodes on processing these packets and the increased copies of data packets passing through longer paths.

4.5. Network Size Effect

For the fourth set of simulations, we varied the network size in order to evaluate the protocols' scalability for larger network areas. Different network sizes have been considered with node density of 60 nodes/km^2. Hence, the studied networks are 500 ×500 m with 15 nodes, 750 ×750 m with

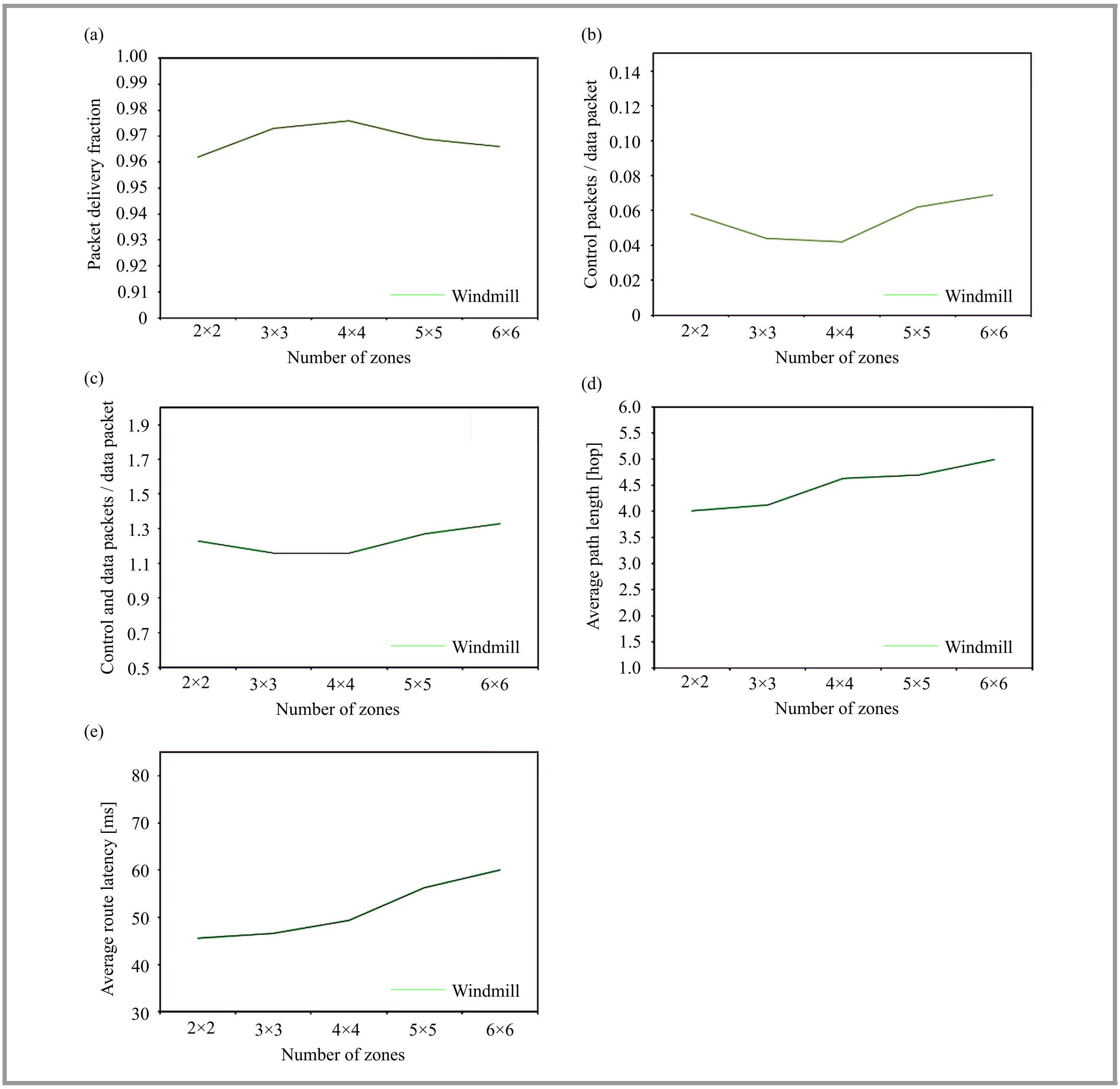


Fig. 9. Simulations results with varying number of zones.

34 nodes, 1000 × 1000 m with 60 nodes, 1250 × 1250 m with 94 nodes, and 1500 × 1500 m with 135 nodes.

Figure 8a shows that PDF of the three protocols decreases with an increase in network size. Larger network size increases the probability of having the source and destination nodes far away from each other, which means that longer routes are established and a higher probability of link breakages and data packet drops exists.

Figures 8b–c reveal that CPD and CDPD for the three protocols increase with an increase in network size due to longer routes and higher probability of link breakages that require route repairs and high control over the packets. CPD and CDPD for ODMRP are still higher than those for MAODV and Windmill, due to increased control and data packets.

As far as APL of the selected routes is concerned, Fig. 8d shows that routes in the proposed protocol are still a little longer than in MAODV and ODMRP, due to the routes passing through ZLs. APL and ARL for the three protocols increase with an increase in the network size, due to increased probability of having the source and destination nodes far away from each other, i.e. longer routes and extended setup times.

4.6. Number of Zones Effect

This parameter has been studied only for the Windmill protocol, since it is the only protocol dealing with the network as zones. To examine the effect of the number of zones, a network of 1 × 1 km is considered. This network is divided into 4 zones, each having the dimensions of 500 × 500 m, 9 zones – each of 333.33 × 333.33 m, 16 zones each of 250 × 250 m, 25 zones each of 200 × 200 m and, finally, 36 zones – each of 166.67 × 166.67 m.

Figure 9a shows that Windmill's PDF is always above 96%. This is an indication that it is highly effective in discovering and maintaining routes, regardless of zone size. Nevertheless, the highest PDF is obtained upon dividing the network into 9 and 16 zones.

Figures 9b–c reveal that the minimum CPD and CDPD values are obtained upon dividing the network into 9 and 16 zones. A large number of zones, i.e. with a small zone size, results in a higher probability of nodes moving from one zone to another, which means a higher control overhead required to maintain the network structure and group membership information, as well as to maintain the routes. A large number of zones also means a higher control overhead needed to discover external routes. On the other hand, an increase in zone size results in a higher probability of having destination nodes in each zone, which means a higher control overhead required to discover internal routes, especially when using ZBrd.

APL and ARL increase along with the increasing number of zones (Figs. 9d–e). A large number of zones means longer routes due to forcing the routes to pass through ZLs.

The analysis shows better performance in terms of PDF, CPD and CDPD for Windmill, when the network is divided into 9 and 16 zones. Moreover, moderate performance in terms of APL and ARL is achieved upon dividing the area into 9 or 16 zones. Hence, it is recommended to divide the network into 3 × 3 or 4 × 4 zones.

5. Results Summary and Discussion

Numerous conclusions may be drawn from the simulation results presented in the previous section:

- PDF for the three protocols is above 95% in most scenarios. This indicates that the three protocols are effective in discovering and maintaining routes for data delivery, even with fairly high node mobility levels and large area networks.
- The proposed protocol performs well in terms of scalability, as it maintains the minimum CPD and CDPD levels in all scenarios. The main reason behind the gap between CPD and CDPD levels typical of Windmill and those of MAODV and ODMRP is that nodes in MAODV and ODMRP are unaware of their and other nodes' positions. Hence, all request packets are sent using broadcasts to the entire network. Additionally, both protocols require sending periodic messages. Windmill, however, does not rely on sending periodic messages. Furthermore, request packets are sent between zones using RDF, with RDF or ZBrd being only used inside zones, with destinations located inside them.
- Slightly longer routes (higher APL) compared to MAODV and ODMRP are the only expense of using the new protocol, since routes in Windmill are forced to go through ZLs.
- Roughly speaking, an increase in node mobility speed, number of sources, multicast group size, and network size results in decreasing PDF and increasing CPD, CDPD, APL and ARL for the three protocols. This is mainly due to higher probability of link breakages which require route repairs and numerous control packets.
- When using Windmill, it is recommended to divide the network into 3 × 3 or 4 × 4 zones, since better performance in terms of PDF, CPD and CDPD, as well as moderate performance in terms of APL and ARL is achieved when dividing the area into 9 or 16 zones.

6. Conclusions and Future Works

The establishment of efficient routes between sources and the anticipated destinations is an important issue in mobile ad-hoc networks. This paper proposes Windmill, a hierarchal multicast routing protocol that seeks to enhance performance and scalability by dividing the network into zones and by relying on RDF. The novel protocol has been assessed and compared with its MAODV and ODMRP counterparts. In MAODV and ODMRP, the nodes are unaware of their and other nodes' positions. Hence, all request packets are sent using a broadcast to the entire network. Additionally, both protocols require sending periodic messages. Windmill, however, does not involve sending periodic messages. Furthermore, the request packets are sent between the zones using RDF, with RDF or ZBrd being only used inside the zones, with destinations located inside them.

A detailed performance evaluation has been conducted. Simulation results illustrate the efficiency of the three protocols in discovering and maintaining routes. Moreover, Windmill performs well in terms of scalability, as it maintains the minimum CPD and CDPD levels, even with high node mobility levels, large number of sources, large multicast groups, and large networks. Windmill's reduced CPD and CDPD levels are a consequence of using restricted directional flooding to send request packets. On the other hand, the proposed protocol's reduced overhead comes at the price of slightly longer routes.

There are still many open research issues related to adhoc networks, such as quality of service and energy efficiency. Security-related aspects stemming from the existence of malicious nodes performing different types of attacks are an interesting area of research as well. This work considered nodes that were evenly distributed from the geographical point of view. So, it is one of our future tasks to study scenarios with dense and sparsely populated regions of the network. Moreover, some improvements may be introduced to Windmill as well, such as turning into a dynamic/adaptive protocol by changing some details concerning the current state of the network. Lastly, we aim to implement and test the proposed protocol in real world conditions.

References

[1] A. Goyal, V. Sharma, S. Kumar, and K. Kumar, "Modified local link failure recovery multicast routing protocol for MANET", *J. of Inform. and Optimiz. Sci.*, vol. 41, no. 2, pp. 669–677, 2020 (DOI: 10.1080/02522667.2020.1733202).

[2] S. Ghasemnezhad and A. Ghaffari, "Fuzzy logic based reliable and real time routing protocol for mobile ad hoc networks", *Wirel. Personal Commun.*, vol. 98, pp. 593–611, 2018 (DOI: 10.1007/s11277-017-4885-9).

[3] M. Qabajeh and L. Qabajeh, "A survey and comparative study for performance evaluation technologies for wireless networks", *Int. J. of Comp. Sci. and Netw. Secur.*, vol. 19 no. 10, pp. 43–50, 2019 [Online]. Available: http://paper.ijcsns.org/07_book/201910/20191008.pdf

[4] A. Tavizi and A. Ghaffari, "Tree-based reliable and energy-aware multicast routing protocol for mobile ad hoc networks", *The J. of Supercomput.*, vol. 74, no. 11, pp. 6310–633, 2018 (DOI: 10.1007/s11227-018-2562-8).

[5] L. Qabajeh and M. Qabajeh, "Detailed performance evaluation of ARANz, ARAN and AODV protocols", *J. of Theoret. and Applied Inform. Technol.*, vol. 98. no. 12, pp. 2109–2131, 2020 [Online]. Available: http://www.jatit.org/volumes/Vol98No12/10Vol98No12.pdf

[6] E. Royer and C. Perkins, "Multicast operation of the ad hoc on demand distance vector routing protocol", in *Proc. 5th Ann. ACM/IEEE Int. Conf. on Mob. Comput. and Network. MOBICOM 1999*, Seattle, VA, USA, 1999, pp. 207–218 (DOI: 10.1145/313451.313538).

[7] S. Lee, M. Gerla, and C. Chiang, "On-demand multicast routing protocol", *IEEE Wirel. Commun. and Network. Conf. WCNC 1999*, New Orleans, LA, USA, 1999 pp. 1298–1302 (DOI: 10.1109/WCNC.1999.796947).

[8] T. Kunz and E. Cheng, "Multicasting in ad-hoc networks: Comparing MAODV and ODMRP", *Workshop on Ad Hoc Commun.*, Bonn, Germany, 2001 [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.5446&rep=rep1&type=pdf

[9] M. Qabajeh, A. Abdalla, O. Khalifa, and L. Qabajeh, "A survey on scalable multicasting in mobile ad hoc networks", *Wirel. Personal Commun.*, vol. 80, pp. 369–393, 2015 (DOI: 10.1007/s11277-014-2016-4).

[10] A. Rajeswari, "A mobile ad hoc network routing protocols: A comparative study", in *Recent Trends in Communication Networks*, P. Mitra, Ed. IntechOpen, 2020. pp. 1–24 (DOI: 10.5772/intechopen.92550).

[11] B. Yang, Z. Wu, Y. Shen, X. Jiang, and Sh. Shen, "On delay performance study for cooperative multicast MANETs", *Ad Hoc Networks*, vol. 102, pp. 1–14, 2020 (DOI: 10.1016/j.adhoc.2020.102117).

[12] S. Balaji, A. Rocha, and Y.-N. Chung (Eds.), *Intelligent Communication Technologies and Virtual Mobile Networks. ICICV 2019. Lecture Notes on Data Engineering and Communications Technologies*, vol. 33 pp. 169–178 and 607–617. Springer, Cham, 2020 (ISBN: 9783030283636).

[13] V. Vinya and G. Rao, "An energy efficient multicast route establishment using AODV with PSO algorithm and RSA for secured transmission", *Int. J. of Intell. Engin. and Syst.*, vol. 12, no. 5, pp. 257–266, 2019 (DOI: 10.22266/ijies2019.1031.26).

[14] A. Goyal and V. Sharma, "Design and implementation of modified local link repair multicast routing protocol for MANETs", *Int. J. of Scient. & Technol. Res.*, vol. 9, no. 2, pp. 2316–2321, 2020 [Online]. Available: http://www.ijstr.org/final-print/feb2020/Design-And-Implementation-Of-Modified-Local-Link-Repair-Multicast-Routing-Protocol-For-Manets.pdf

[15] L. Junhai, X. Liu, and Y. Danxia, "Research on multicast routing protocols for mobile ad-hoc networks", *Comp. Networks*, vol. 52, no. 5, pp. 988–997, 2008 (DOI:10.1016/j.comnet.2007.11.016).

[16] C. Morais, H. Gossain, and D. Agrawal, "Multicast over wireless mobile ad hoc networks: Present and future directions", *IEEE Network*, vol. 17, no. 1, pp. 52–59, 2003 (DOI: 10.1109/MNET.2003.1174178).

[17] Z. Xiaofeng and L. Jacob, "Multicast zone routing protocol in mobile ad hoc wireless networks", in *Proc. 28th Ann. IEEE Int. Conf. on Local Computer Networks LCN 2003*, Bonn/Konigswinter, Germany, 2003, pp. 50–159 (DOI: 10.1109/LCN.2003.1243122).

[18] Ch. Zhi, J. Cui, and L. Zhu, "Multicast routing algorithms based on Levy flying particle swarm optimization", *J. of Physics: Conf. Series 1453*, 2020 (DOI: 10.1088/1742-6596/1453/1/012005).

[19] A. Sufian, A. Banerjee, and P. Dutta, "A tree multicast routing based on fuzzy mathematics in mobile ad-hoc networks", in *Applications of Internet of Things*, J. Mandal, S. Mukhopadhyay, A. Roy, Eds. *Lecture Notes in Networks and Systems*, vol 137, pp. 107–117. Springer, 2020 (DOI: 10.1007/978-981-15-6198-6_10).

[20] K. Vanisrsee and V. Reddy, "Multicast cooperative routing for opportunistic data transfer in mobile ad hoc network", *J. of Electron. and Commun. Engin.*, vol. 12, no. 3, pp. 26–33, 2017 (DOI: 10.9790/2834-1203012633).

[21] L. Canourgues, J. Lephay, L. Soyer, and A. Beylot, "STAMP: Shared-tree ad hoc multicast protocol", *in Proc. of Military Commun. Conf. MILCOM 2006*, Washington, DC, USA, 2006 (DOI: 10.1109/MILCOM.2006.302031).

[22] D. Babu and M. Ussenaiah, "CS-MAODV: Cuckoo search and M-tree-based multiconstraint optimal multicast ad hoc on-demand distance vector routing protocol for MANETs", *Int. J. of Commun. Syst.*, vol. 33, no. 16, pp. 1–17, 2020 (DOI: 10.1002/dac.4411).

[23] T. Winter and P. Thubert, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", IETF ROLL working group, Feb. 2010 [Online]. Available: https://datatracker.ietf.org/doc/html/rfc6550 (retrieved 2021-9-30).

[24] C. Murthy and B. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols*. Prentice Hall PTR, 2004 (ISBN: 9780133007060).

[25] J. Garcia and E. Madruga, "The core-assisted mesh protocol", *IEEE J. on Selec. Areas in Commun.*, vol. 17, no. 8, pp. 1380–1394, 1999 (DOI: 10.1109/49.779921).

[26] I. Al-Mejibli, "Improve on-demand multicast routing protocol in mobile ad-hoc networks", *J. University of Kerbala*, vol. 16, no. 1, pp. 351–363, 2018 [Online]. Available: http://iraqjournals.com/article_143909_b8dbdbc05b91967f48d085bf69d3a977.pdf

[27] J. Xie, R. Talpade, A. McAuley, and M. Liu, "AMRoute: Ad hoc multicast routing protocol", *Mob. Netw. and Appl.*, vol. 7, no. 6, pp. 429–439, 2002 (DOI: 10.1023/A:1020748431138).

[28] J. Biswas, M. Barai, and S. Nandy, "Efficient hybrid multicast routing protocol for ad-hoc wireless networks", in *Proc. 29th Ann. IEEE Int. Conf. on Local Comp. Networks*, Tampa, FL, USA, 2004 (DOI: 10.1109/LCN.2004.47).

[29] G. Walikar and R. Biradar, "Energy aware hybrid multicast routing in mobile ad hoc networks: zone-based approach", *Int. J. of Mob. Netw. Design and Innov.*, vol. 8, no. 2, pp. 80–100, 2018 (DOI: 10.1504/IJMNDI.2018.092344).

[30] T. Shih, C. Shih, and C. Chen, "Location-based multicast routing protocol for mobile ad hoc networks", *WSEAS Trans. on Computers*, vol. 7, no. 8, pp. 1270–1279, 2008 (DOI: 10.5555/1457999.1458015).

[31] X. Xiang, X. Wang, and Y. Yang, "Supporting efficient and scalable multicasting over mobile ad hoc networks", *IEEE Trans. on Mob. Comput.*, vol. 10, no. 4, pp. 544–559, 2010 (DOI: 10.1109/TMC.2010.176).

[32] J. Jetcheva and D. Johnson, "Adaptive demand-driven multicast routing in multi-hop wireless ad hoc networks", in *Proc. ACM Int. Symp. on Mob. AdHoc Network. & Comput. MobiHoc'01*, Long Beach, CA, USA, 2001 (DOI: 10.1145/501422.501423).

[33] M. Qabajeh, A. Hashim, O. Khalifa, and L. Qabajeh, "Geographical multicast quality of service routing protocol for mobile ad-hoc networks", *J. of Engin. Lett.*, vol. 18, no. 3, pp. 212–225, 2010 [Online]. Available: http://www.engineeringletters.com/issues_v18/issue_3/EL_18_3_02.pdf

[34] A. Daniel, "Position based multicast routing protocol for ad-hoc wireless network using backpressure restoration", in *Proc. 2nd Int. Conf. on Comp. Engin. and Technol.*, Chengdu, China, 2010 (DOI: 10.1109/ICCET.2010.5485544).

[35] H. Hussen, S. Choi, J. Park, and J. Kim, "Predictive geographic multicast routing protocol in flying ad hoc networks", *Int. J. of Distrib. Sensor Netw.*, vol. 15, no. 7, pp. 1–20, 2019 (DOI: 10.1177/1550147719843879).

[36] R. Shankar and E. Ilavarasan, "Scalable multicasting through hexagonal zone based structure over mobile adhoc networks", *J. of Internet Technol.*, vol. 19, no. 7, pp. 2111–2124, 2018 (DOI: 10.3966/160792642018121907014).

[37] "GloMoSim: A Scalable Simulation Environment for Wireless and Wired Network Systems", UCLA Parallel Computing Laboratory and Wireless Adaptive Mobility Laboratory [Online]. Available: http://pcl.cs.ucla.edu/projects/glomosim/

[38] S. Corson and J. Macker, "Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations", RFC 2501, 1999 [Online]. Available: http://www.ietf.org/rfc/rfc2501.txt

Liana Khamis Qabajeh received her B.Sc. in Computer Engineering from Palestine Polytechnic University (PPU), Palestine, in 2000 and joined the Engineering and Technology Faculty, PPU, as a research assistant. She received her M.Sc. in Computer Engineering from Jordan University of Science and Technology, Jordan, in 2005. Between 2005 and 2008, before pursuing her study, she was primarily involved in academic teaching and research at PPU. She has secured her Ph.D. in Computer Science in 2012 from University of Malaya, Malaysia. In 2012 she was appointed assistant professor at the Information Technology and Computer Engineering Faculty, PPU. Her current research interests include distributed systems and ad-hoc networks.

https://orcid.org/0000-0002-2801-0070
E-mail: liana_tamimi@ppu.edu
Faculty of Information Technology
and Computer Engineering
Palestine Polytechnic University
Hebron, Palestine

Paper

An Attribute-Based Encryption Method Using Outsourced Decryption and Hierarchical Access Structure

Tabassum N. Mujawar[1,2] and Lokesh B. Bhajantri[3]

[1] Research Scholar, Department of CSE, Basaveshwar Engineering College, Bagalkot, Karnataka, India
[2] Department of Computer Engineering, Ramrao Adik Institute of Technology, D Y Patil deemed to be University, Navi Mumbai, Maharashtra, India
[3] Department of ISE, Basaveshwar Engineering College, Bagalkot, Karntaka, India

https://doi.org/10.26636/jtit.2022.158421

Abstract—Cloud computing is being rapidly adopted by many organizations from different domains and large amounts of data is stored in the cloud. In order to ensure data security, the attribute-based access control mechanism has been emerging recently as a fine-grained access control model that grants access based on the data user's attributes. In this model, the data owner builds the access policy using the attributes of the data users and access to the data is granted only if the requirements of such an access policy are satisfied. Ciphertext policy-based attribute-based encryption (CPABE) is one of the most widely used methods for providing encrypted access control. Complex, time consuming and costly paring operations are the major issue with the CPABE method. Hence, another efficient method is needed to reduce the data user's overhead while decrypting data. This paper presents an efficient method consisting in outsourcing decryption operations to a third-party server, so that complex operations may be performed by that machine with only some simple calculations left on the data user's side. The concept of a hierarchical access structure is also integrated with the traditional CPABE technique. The hierarchical approach enables the data owner to encrypt multiple data using a single common hierarchical access structure. This allows the user to decrypt only the relevant part of ciphertext, depending on which fragment of the hierarchical access structure is satisfied. The paper evaluates also the performance of the proposed model in terms of time and storage cost.

Keywords—cloud computing, CPABE, hierarchical access structure.

1. Introduction

In order to maintain authenticity of the data stored on cloud servers, it is necessary to apply appropriate security mechanisms. The data must be accessible to authorized users only and access must be denied to other parties. In traditional models, access permissions will be granted by the server on which the data is stored. However, the server itself is an untrusted entity and if it is compromised, the data may be accessed by any unauthorized person. Also, in this approach, the data owner is completely dependent on a third-party server for enforcing the data access policies. Hence, another approach in which the data owner is capable of controlling the access policies and deciding who can access the data is required. The modern attribute-based access control model provides access to data based on considering attributes of the data users and allows data owners to establish access policies by combining attributes and specifying which data user is allowed to access specific data.

Attribute-based encryption (ABE) is a scheme that offers encrypted access control by considering the user's attributes [1]. The access policy is built using different logic gates and data user's attributes. The scheme provides fine-grained access control and relies on a one-to-many encryption mechanism. Two different categories of the ABE scheme may be distinguished: ciphertext policy-based attribute-based encryption (CPABE) [2] and key policy-based attribute-based encryption (KPABE) [3]. In the case of CPABE, the user's attributes are bound with a private key used for decryption and the access structure is associated with the ciphertext. The ciphertext is decrypted only when the attributes associated with the decryption key satisfy the access structure. In the case of KPABE, the ciphertext is combined with the attributes and the access structure is integrated with the decryption key.

CPABE is the most widely adopted scheme providing access control based on specific attributes. The traditional CPABE scheme proposed in [2] includes a rather costly decryption process. The degree of complexity is increased with the complexity of the access policy. This is a significant the drawback of the scheme, as the data user incurs a lot of overhead. In order to deal with this issue, an outsourcing mechanism is applied, allowing to hand over the costly operations to a third-party server. This machine generates a transformed ciphertext which can be further decrypted by the data user. While decrypting the transformed ciphertext, the user has to apply simple computations and this reduces the burden on the data user's side. In the traditional scheme, each message is encrypted separately, using the relevant access structure, and the ciphertext is generated. Sometimes, the different access structures are hierarchically related and that can be combined to form one common access structure which, in turn, may be relied upon to encrypt multiple pieces of data together, instead of encrypting them separately. Hence, the time required for encryption and the storage cost related to the generated ciphertext will be reduced by applying such hierarchical access structures.

In this paper, an efficient CPABE scheme that utilizes the hierarchical access structure to encrypt data is proposed. Hierarchical access structures are built by combining multiple hierarchically related access structures. This approach helps to encrypt multiple pieces of data together, and generate a common ciphertext. If the access structure is satisfied fully, then the entire ciphertext is decrypted. If only a specific portion of the access structure is satisfied, then only the relevant portion of the ciphertext is decrypted. All major pairing operations are outsourced to an outsourcing server. This third-party outsourcing server returns the partially decrypted ciphertext to the data user. Then, the user can apply simple computations and can retrieve the original data. This scheme takes advantage of both the hierarchical access structure and the outsourcing approach. The proposed scheme is less costly in terms of storing the ciphertext and less time-consuming when it comes to performing the decryption process.

The remaining part of this paper is organized as follows. Section 2 describes the existing attribute-based encryption methods that rely on the outsourcing mechanism. Section 3 elaborates on the proposed outsourced decryption and hierarchical access structure-based CPABE scheme. The experimental analysis is described in Section 4. Conclusions are presented in Section 5.

2. Related Work

The complexity of the decryption operation in a traditional attribute-based encryption scheme depends primarily on how complex the access policy used for encryption is. The size of the generated ciphertext is quite large as well, meaning that the scheme requires more time for decryption. The overhead is incurred by the data user intending to access the encrypted data. An outsourcing-based ABE scheme is proposed in [4], eliminating the overhead stemming from decryption operations. In this scheme, the ABE ciphertext is converted into an ElGamal style ciphertext by the cloud. Then, the data user may transform this ElGamal style ciphertext to plaintext with less processing. The drawback of such an approach is that the correctness of the transformed ciphertext is not verified. It may be the case that a malicious cloud node may perform an incorrect transformation and, hence, the data user will not receive correct data. A scheme that performs an outsourced decryption along with a verifiable transformation is proposed in [5]. The correctness of transformation is verified by computing the checksum, which is a combination of one random message and the original message to be encrypted. This checksum and the random message are added to the ciphertext. The data user may verify correctness by computing the checksum again. An efficient outsourced ABE scheme with verifiability of transformation based on the key encapsulation mechanism is proposed in [6]. Here, the data is encrypted using symmetric encryption and the ABE approach is applied for encrypting the symmetric key. The hash value of the key is computed and is concatenated with ciphertext. The hash function is once again applied on this concatenated message and this hash is used to check whether the transformation has been performed correctly or not. The first hash value is used to check the integrity of the encryption key. The ABE scheme with outsourced decryption and verification that is CPA- and RCCA-secure is presented in [7]. This scheme utilizes a random value to check whether the cloud node has performed partial decryption correctly or not. The message and the random value are encrypted together. The scheme requires fewer computation resources and the ciphertext is small in size as well.

The standard model for a CPABE scheme with verifiable outsourced decryption and with a constant ciphertext length is presented in [8]. In this scheme the size of the ciphertext does not increase, despite the growing number of attributes or the complexity of access structure. The costly paring operations are outsourced to the cloud node and the data user can recover the message by applying very simple computations. The scheme proposed in [9] provides a fully verifiable outsourced decryption facility. The different access policies are designed for authorized and unauthorized users. The MAC is integrated with the ciphertext to ensure that the transformation performed by the cloud is correct. The scheme first verifies the correctness of the transformed ciphertext and then decrypts the ciphertext to obtain the plain text.

A scheme that outsources both encryption and decryption operations to a third-party untrusted server is proposed in [10]. Here, the exponentiation modulo computations are outsourced to a third-party encryption server, so that the overhead of the data owner can be reduced. An efficient method for generating the transformation key is proposed as well. The scheme verifies the correctness of the transformed messages and, hence, ensures their verifiability. An ABE scheme that supports hidden access policies is proposed in [11]. In this scheme, the attribute name is kept public, whereas the attribute value is kept hidden, so that privacy can be maintained. The scheme also outsources the costly paring operations to cloud nodes and data users can verify the calculations performed thereby.

An online-offline scheme for resource constrained devices operating in the cloud environment is proposed in [12]. It uses the Chameleon hash function to generate an immediate ciphertext and blind it with the offline ciphertext. In order to eliminate the overhead of decryption without authorization, a ciphertext test is performed before decryption. A verifiable and multiple authority-based ABE scheme is presented in [13]. In this scheme, the majority of encryption and decryption operations are transferred to fog devices. This will reduce the overhead on the part of the data owner and user. The scheme also presents a verification method to check the computations performed by fog devices. An efficient revocation method for users and attributes is also implemented. A fully outsourced ABE scheme is presented in [14]. The CPABE scheme is implemented by outsourcing all major tasks, such as encryption, decryption and key generation. In order to manage the additional communication overhead, key generation and encryption operations are performed offline. In [15], an efficient CPABE scheme that supports outsourced encryption and decryption is presented. The scheme utilizes the fog computing environment to outsource complex encryption and decryption operations. The fog nodes are responsible for partially encrypting the message and partially decrypting the ciphertext. The scheme also includes an efficient attribute revocation mechanism.

In [16], the authors present two different CPABE schemes, where the encryption and decryption operations are outsourced. In the first approach, the untrusted service provider performs the complex operations related to encryption and decryption. On the other hand, the other approach includes a key encapsulation mechanism, with the encryption and decryption operations being outsourced to a semi-trusted service provider. A traceable and outsourced decryption-based ABE scheme is presented in [17]. The scheme provides an outsourced decryption facility to reduce burden of the data user. A mechanism for updating policies while encrypting data is also included. A traceability feature is also incorporated in this scheme, so that malicious users can be detected.

3. Proposed Work

The CPABE scheme that supports hierarchical access structure and outsourced decryption is presented in this paper. The hierarchically related access structures are combined together to form a common hierarchical access structure. This feature will allow encrypting multiple messages using one common hierarchical access structure. The major issue associated with traditional CPABE schemes is that the decryption operation is very costly computationally. The decryption of a CPABE ciphertext requires complex paring operations and this incurs computational overhead for the data user. Hence, an outsourcing mechanism is proposed permitting to transfer complex operations to a third-party server. This unit will generate a partially decrypted ciphertext without understanding any details about the original message. Therefore, privacy of the original message will be preserved. This outsourcing concept will reduce the burden of performing complex tasks on the data user's side. The model of the proposed scheme is presented in Fig. 1.

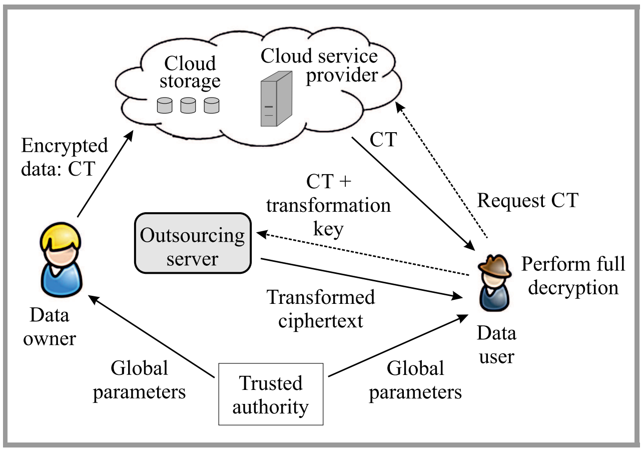


Fig. 1. System architecture.

The data owner generates access policies for the files that they want to store along with cloud storage. The access policies are prepared by combining attributes and logic gates. These decide who can access the data. A set of attributes of those users who are going to access the data is maintained and is based on the sensitivity of data. A tree-based access structure is generated to represent these conditions. Hierarchically related access structures are combined and one hierarchical access structure is formed. The data owner encrypts the data by using such a hierarchical access structure. This allows the data owner to encrypt multiple pieces of data together and generate a common ciphertext. This will reduce the efforts related to encrypting multiple pieces of data separately by using different access structures. This will reduce the encryption time and also the storage cost required to store separate ciphertexts. As the CPABE scheme is implemented, the access structure is integrated with the ciphertext and is then stored in the cloud. Global parameters of the system needed for encryption are issued by a trusted authority.

The cloud service provider is responsible for handling all requests of the data owner for storing the encrypted data in cloud storage and also deals with the users who are requesting access to the encrypted data. The data user must possess the necessary attributes to access the data from cloud storage. The service provider will transfer the ciphertext to the outsourcing server for partial decryption upon receiving a request from the data user.

The outsourcing server generates the transformed ciphertext if the attributes of data users satisfy the access structure. The hierarchical access structure is applied while encrypting the data. Hence, decryption is performed based on how big a portion of the access structure is satisfied. If the complete access structure is satisfied, then the user is allowed to access all of information. On the other hand, if only a specific portion of the access structure is satisfied, then only the data associated with that portion are accessible. Thus, the hierarchical access structure allows decrypting of all data or of its relevant portion. The outsourcing server only partially decrypts the data and generates the transformed ciphertext. The transformation algorithm will be applied in such a way that the outsourcing server cannot read or understand anything about the original data. The actual data can be retrieved only by the data user, by performing necessary computations over the transformed ciphertext. The transformation key is given to the outsourcing server to perform the partial decryption step. The partially decrypted ciphertext is sent to the data user. Now, the data user can decrypt it using the secret key and can recover the actual data.

3.1. Proposed Scheme

The proposed model comprises different phases, such as system setup, encryption, key generation, transformation key generation, partial decryption and decryption. The traditional CPABE system [2] is relied upon to implement these phases.

The system setup phase is responsible for generating the public key (PK) and the master key (MK). The trusted authority generates these keys by taking global system parameters as input. Let $\mathbb{G}$ be the bilinear group of prime order p with generator g and there is a bilinear map as e : $\mathbb{G} \times \mathbb{G}$. Then, the two elements α and β are selected over $\mathbb{Z}$p. PK is computed as PK $= \left[\mathbb{G}_0, g, g^{\beta}, e(g,g)^{\alpha}\right]$. MK is computed as MK $= g^{\alpha}$.

The hierarchical access structure $\mathbb{A}$ serves as the input for encryption, in the form of access tree, as well as PK and the messages to be encrypted $\{m_1, \dots m_n\}$. Multiple messages are encrypted by applying the hierarchical access structure and the ciphertext is generated. The polynomial q_{node} and threshold value k are associated with each node of the access tree. In the access tree, the attributes are present at leaf nodes and their threshold value is 1. The threshold values for AND and OR gate are set to 2 and 1, respectively. The degree of the polynomial associated with each node is set to one less than the corresponding threshold value.

Let us assume there are n messages to be encrypted and m attributes. The secret value s for the root node of the hierarchical access structure is selected randomly, such that $s \in \mathbb{Z}p$. Let s_i be the secret associated with the root node of the access structure meant for message m_i. The ciphertext components for i-th message are computed as $CT_i' = m_i.e(g,g)^{\alpha s_i}$ and $CT_i^* = h^{s_i}$, where $h = g^{\beta}$.

In access tree the attributes are present at leaf nodes. For $k = 1, \dots, m$ attributes and for each such node we compute $C_{i,k} = g^{q_k(0)}$ and $C_{i,k}' = H\left[\text{attribute}(k)\right]^{q_k(0)}$, where H is the hash function.

Finally, the generated ciphertext is represented as:

$$\begin{aligned} CT = \Big\{ & \mathbb{A}, \{m_1, \dots, m_n\} \ , \\ & CT_i', CT_i^*, C_{i,k}, C_{i,k}' \ , \\ & \forall i \in (1,n) \ , \\ & \forall k \in (1,m) \Big\} \ . \end{aligned}$$

In the key generation phase, SK is sent to the user by taking MK and set of attributes A as input. Let there be m attributes in set $A = \{a_1, a_2, \dots, a_m\}$. The random element x over $\mathbb{Z}p$ is selected and for each attribute a_j, a random element y_j over $\mathbb{Z}p$ is also selected. The secret key component is computed as $SK' = (g^{\alpha+x})^{\frac{1}{\beta}}$. For each attribute a_j, $S_j = g^x.H(a_j)^{y_j}$ and $S_j' = g^{y_j}$ is computed. The secret key is represented as $SK = \left\{SK', S_j, S_j', \forall j \in (1,m)\right\}$.

Next, the transformation key (TK) is needed to support the outsourced decryption. It is used by the outsourcing server to partially decrypt the ciphertext. PK and SK are taken as input and TK is generated as output by this phase. The random element z over $\mathbb{Z}p$ is selected and for each attribute a_j random element y_j over $\mathbb{Z}p$ is selected. Then, $D = (SK')^{\frac{1}{z}}$ is computed and for each attribute a_j, $D_j = g^{\frac{x}{z}}.H(a_j)^{\frac{y_j}{z}}$ and $D_j' = g^{\frac{y_j}{z}}$.

The partial decryption phase takes the ciphertext (CT) and TK as input and generates the partially decrypted ciphertext CT''. Complex paring computations are performed in this phase by the outsourcing server. Partial decryption is performed only when the user's attributes satisfy the necessary access structure. If the access structure is not satisfied by the user's attributes, then the phase returns null. If the access structure is satisfied, then this phase recovers value $e(g,g)^{\frac{xs_i}{z}}$ for each message. The transformed ciphertext for each ciphertext CT_i is computed as $CT_i'' = (CT1_i, CT2_i)$. Here, $CT1_i = CT_i'$ and

$$CT2_i = \frac{e(CT_i^*, D)}{e(g,g)^{\frac{xs_i}{z}}} \ .$$

In decryption phase the transformed ciphertexts CT_i'', CT and the TK are taken as input and the actual data is recovered. Now, for complete decryption, a simple operation is required as:

$$m_i = \frac{CT1_i}{(CT2_i)^z} \ .$$

4. Performance Analysis

The proposed hierarchical access structure and the outsourced decryption-based CPABE scheme are implemented using the Java JPBC library. Table 1 presents a comparison of the proposed scheme and some existing solutions with respect to the features adopted in each of the approaches. As per this comparison, only the proposed scheme supports a hierarchical access structure and, hence, it is more efficient in terms of encryption time and storage cost.

Table 1
Comparison of features

Scheme	Access structure	Encryption time	Outsourced operation	Storage cost
[14]	LSSS	Standard	Encryption, key generation and decryption	Standard
[10], [15], [16]			Encryption and decryption	
Proposed	Hierarchical	Less	Decryption	Less

Evaluation of the performance of the proposed scheme consists in comparing it with schemes with and without outsourcing. The symmetric encryption technique is used to encrypt the messages and the key used for encryption is encrypted using the proposed CPABE scheme. The data owner generates different hierarchical access structures, as required. These access structures are used to encrypt multiple files.

Various experiments are carried out by varying the number of attributes and the time needed to encrypt and decrypt is measured. Also, the number of messages to be encrypted is varied and the performance is measured with respect to time required for encryption and decryption. The hierarchical access structure eliminates the need for generating multiple ciphertexts and, hence, considerably improves storage cost as well. Storage cost is also compared by varying the number of messages.

Comparison of the decryption time needed by the proposed scheme, the traditional CPABE approach and CPABE with a hierarchical access structure, with a varying the number of attributes, is carried out and is shown in Fig. 2. As the proposed scheme outsources the decryption to a third-party server, the time required by the user to recover the original message is much shorter when compared with the traditional system. Therefore, the computation overhead of the data user is also reduced. The number of attributes is varied from 2 to 20. It can be observed from the Fig. 2 that the proposed scheme needs less time for decryption than the two remaining methods, for any number of attributes.

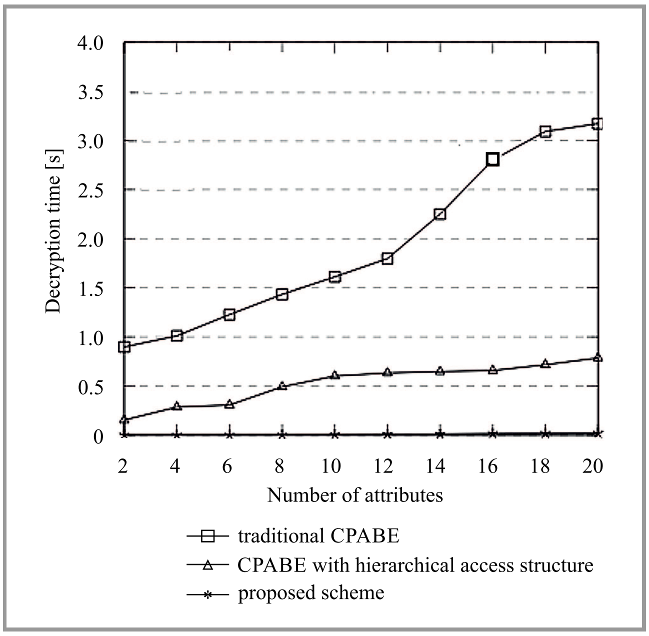


Fig. 2. Comparison of decryption time with respect to the number of attributes.

The encryption time required by the proposed method is also analyzed by varying the number of attributes. Performance is compared with the traditional CPABE approach without a hierarchical access structure (Fig. 3). It can be observed that the proposed method outperforms the other approach. This is possible because one common hierarchical access structure is used to encrypt multiple messages, instead of encrypting them separately.

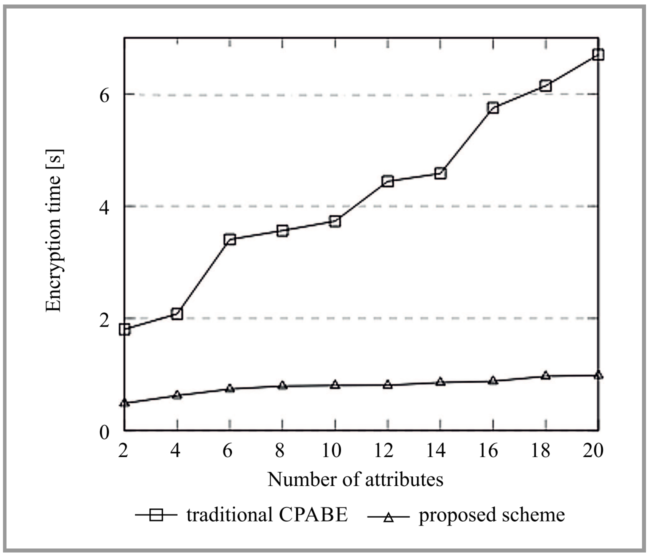


Fig. 3. Comparison of encryption time with respect to the number of attributes.

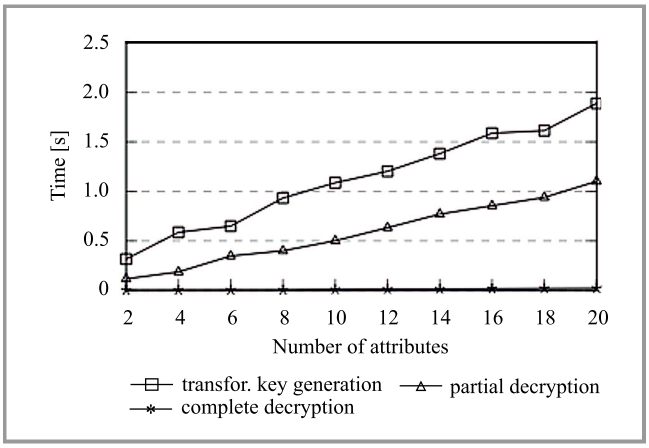


Fig. 4. Comparison of time required for decryption components with respect to the number of attributes.

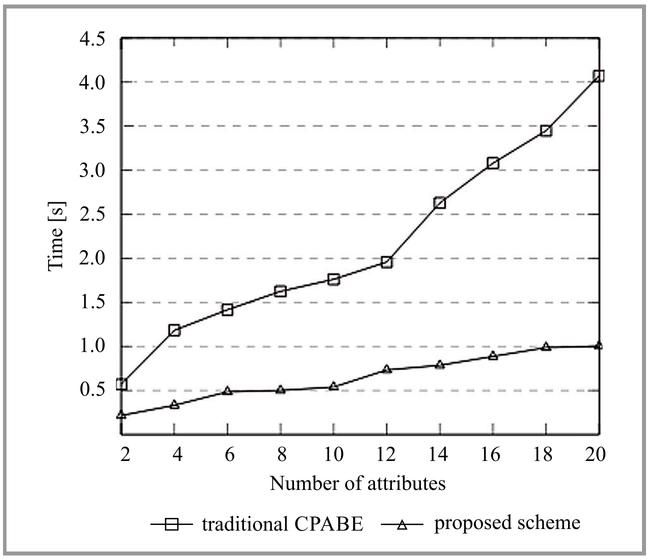


Fig. 5. Comparison of time required for private key generation with respect to the number of attributes.

In the proposed system, decryption is outsourced to an outsourcing server. The decryption process involves various components, such as transformation key generation, partial decryption and complete decryption. The comparison of time required for all these components, by varying the number of attributes, is shown in Fig. 4. It can be observed that the time required for transformation key generation and partial decryption is linearly dependent on the number of attributes. The time needed for complete decryption is almost constant, as very simple computations are performed by the data user for decrypting the partially decrypted ciphertext.

Figure 5 shows the comparison of time required for secret key generation, with different numbers of attributes, by the traditional approach and the proposed scheme. The proposed scheme takes less time to complete the operation compared with the traditional approach. The time required by the proposed scheme increases at a slower rate, compared to other methods.

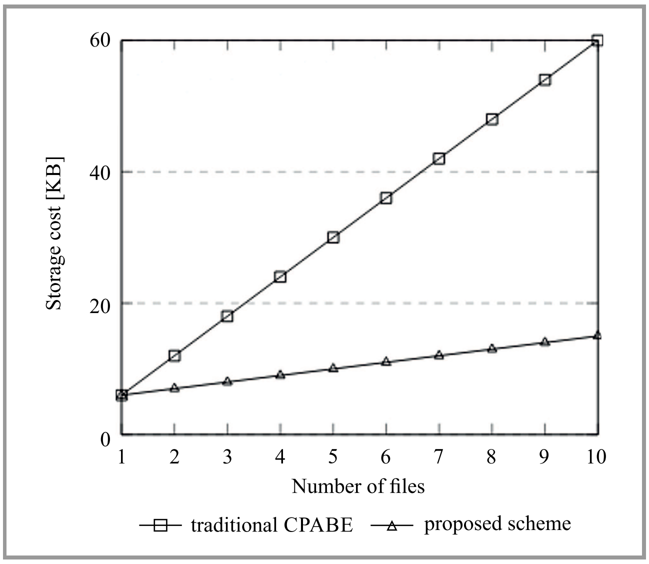


Fig. 6. Comparison of storage cost required for ciphertext with respect to the number of files.

The comparison of storage cost needed for storing ciphertext by the proposed outsourcing model and the traditional scheme, for a different number of files, is shown in Fig. 6. Ten different files with size of 1 KB each are considered for evaluating the performance. The storage cost of the proposed scheme is significantly lower compared with other methods.

5. Conclusion

The paper presents an attribute-based encryption method that utilizes the hierarchical access structure and the outsourcing mechanism. In the proposed scheme, the common hierarchical access structure is generated by combing hierarchically related access structures. Hence, multiple data can be encrypted with this common access structure and an appropriate portion of data is decrypted by satisfying the relevant portion of the access structure. This reduces the encryption time and storage space required for ciphertext. In order to deal with the complexity of the decryption operation, the outsourced decryption mechanism is presented. Here, the cloud server performs the partial decryption process and generates some intermediate ciphertext. Further, this transformed ciphertext is decrypted by the user completely. The outsourcing mechanism reduces the burden of complex operation on the data user's side and makes decryption process more efficient. Thus, the proposed scheme constitutes an efficient approach for implementing attribute-based access control for cloud data. The proposed method is efficient in terms of encryption time, decryption time and storage cost when compared with the traditional approach.

References

[1] A. Sahai and B. Waters, "Fuzzy identity based encryption", in *Advances in Cryptology – EUROCRYPT 2005. 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings*, R. Cramer, Ed. *LNCS*, vol. 3494, pp. 457–473. Berlin, Heidelberg: Springer, 2005 (DOI: 10.1007/11426639_27).

[2] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext policy attribute based encryption", in *Proc. IEEE Symp. on Secur. and Priv. SP'07*, Berkeley, CA, USA, 2007, pp. 321–334 (DOI: 10.1109/SP.2007.11).

[3] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute based encryption for fine-grained access control of encrypted data", in *Proc. of the 13th ACM Conf. on Comp. and Commun. Secur.*, Alexandria, VA, USA, 2006, pp. 89–98, 2006 (DOI: 10.1145/1180405.1180418).

[4] M. Green, S. Hohenberger, and B. Waters, "Outsourcing the decryption of ABE ciphertexts", in *Proc. of the 20th USENIX Conf. on Secur.*, San Francisco, CA, USA, 2011 [Online]. Available: https://www.usenix.org/legacy/event/sec11/tech/full_papers/Green.pdf

[5] M. Green, S. Hohenberger, and B. Waters, "Attribute-based encryption with verifiable outsourced decryption", *IEEE Trans. on Inform. Foren. and Secur.*, vol. 8, no. 8, pp. 1343–1354, 2013 (DOI: 10.1109/TIFS.2013.2271848).

[6] B. Qin, R. H. Deng, S. Liu, and S. Ma, "Attribute-based encryption with efficient verifiable outsourced decryption", *IEEE Trans. on Inform. Foren. and Secur.*, vol. 10, no. 7, pp. 1384–1393, 2015 (DOI: 10.1109/TIFS.2015.2410137).

[7] X. Mao, J. Lai, Q. Mei, K. Chen, and J. Weng, "Generic and efficient constructions of attribute-based encryption with verifiable outsourced decryption", *IEEE Trans on Depend. and Secure Comput.*, vol. 13, pp. 533–546, 2016 (DOI: 10.1109/TDSC.2015.2423669).

[8] J. Li, F. Sha, Y. Zhang, X. Huang, and J. Shen, "Verifiable outsourced decryption of attribute-based encryption with constant ciphertext length", *Secur. Commun. Netw.*, vol. 2017, pp. 1–11, 2017 (DOI: 10.1155/2017/3596205).

[9] J. Li, Y. Wang, Y. Zhang, and J. Han, "Full verifiability for outsourced decryption in attribute based encryption", *IEEE Trans. on Serv. Comput.*, vol. 13, pp. 478–487, 2017 (DOI: 10.1109/TSC.2017.2710190).

[10] Z. Li, W. Li, Z. Jin, H. Zhang, and Q. Wen, "An efficient ABE scheme with verifiable outsourced encryption and decryption", *IEEE Access*, vol. 7, pp. 29023–29037, 2019 (DOI: 10.1109/ACCESS.2018.2890565).

[11] J. Yu, G. He, X. Yan, Y. Tang, and R. Qin, "Outsourced ciphertext-policy attribute-based encryption with partial policy hidden", *Int. J. of Distrib. Sensor Netw.*, vol. 16, pp. 1–14, 2020 (DOI: 10.1177/1550147720926368).

[12] L. Li, Y. Zhang, X. Chen, and Y. Xiang, "Secure attribute-based data sharing for resource-limited users in cloud computing", *Computers & Secur.*, vol. 72, pp. 1–12, 2018 (DOI: 10.1016/j.cose.2017.08.007).

[13] K. Fan, J. Wang, X. Wang, H. Li, and Y. Yang "A secure and verifiable outsourced access control scheme in fog-cloud computing", *Sensors*, vol. 17, no. 7, Article no. 1695, 2017 (DOI: 10.3390/s17071695).

[14] R. Zhang, H. Ma, and Y. Lu, "Fine-grained access control system based on fully outsourced attribute-based encryption", *J. of Syst. Softw.*, vol. 125, pp. 344–353, 2017 (DOI: 10.1016/j.jss.2016.12.018).

[15] J. Zhao, P. Zeng, and K. R. Choo, "An efficient access control scheme with outsourcing and attribute revocation for fog-enabled e-health", *IEEE Access*, vol. 9, pp. 13789–13799, 2021 (DOI: 10.1109/ACCESS.2021.3052247).

[16] H. E. Gafif and A. Toumanari, "Efficient ciphertext-policy attribute-based encryption constructions with outsourced encryption and decryption", *J. Secur. and Commun. Netw.*, vol. 2021, pp. 1–17, 2021 (DOI: 10.1155/2021/8834616).

[17] K. Sethi, A. Pradhan, and P. Bera, "Practical traceable multi-authority CP-ABE with outsourcing decryption and access policy updation", *J. of Inform. Secur. and Appl.*, vol. 51, pp. 1–16, 2020 (DOI: 10.1016/j.jisa.2019.102435).

Tabassum N. Mujawar earned her M.E. in Computer Engineering from the University of Mumbai, Maharashtra, India in 2012. Currently, she is pursuing a Ph.D. in Computer Science and Engineering from Basaveshwar Engineering College, Bagalkot, affiliated to Visvesvaraya Technological University (VTU), Belgaum, Karnataka, India. She is working as an Assistant Professor at the Ramrao Adik Institute of Technology, D Y Patil deemed to be University. She has total 15 years of teaching experience. Her areas of interests include cloud computing, security and machine learning.

E-mail: tabbu3002@gmail.com
Research Scholar,
Department of Computer Science
and Engineering
Basaveshwar Engineering College,
Bagalkot, Karntaka, India

Department of Computer Engineering
Ramrao Adik Institute of Technology,
D Y Patil deemed to be University
Navi Mumbai, Maharashtra, India

Lokesh B. Bhajantri received his Ph.D. degree in Computer Science and Engineering from the Visvesvaraya Technological University (VTU), Belgaum, Karnataka, in 2015. He has been working, for the past 16 years, as an Associate Professor at the Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkot, India. His areas of interests include distributed/wireless sensor networks, cognitive Internet of Things, mobile computing and communications, networking protocols, genetic algorithms, applications of agents, as well as real time systems.

https://orcid.org/0000-0002-3947-4292
E-mail: lokeshcse@yahoo.co.in
Department of ISE
Basaveshwar Engineering College
Bagalkot, India

Paper

Modeling and Parameter Estimation of Radar Sea-Clutter with Trimodal Gamma Population

Zakía Terki[1], Amar Mezache[2,3], and Fouad Chebbara[1]

[1] *Laboratoire de Génie Electrique, LAGE, Département d'Electronique, Université Kasdi Merbah Ouargla, Ouargla, Algéria*
[2] *Département d'Electronique, Université Mohamed Boudiaf M'sila, M'sila, Algéria*
[3] *Laboratoire SISCOM, Université de Constantine, Constantine, Algéria*

https://doi.org/10.26636/jtit.2022.160422

Abstract—Real radar data often consist of a mixture of Gaussian and non-Gaussian clutter. Such a situation creates one or more inflexion points in the curve of the empirical cumulative distributed function (CDF). In order to obtain an accurate fit with sea reverberation data, we propose, in this paper, a trimodal gamma disturbance model and two parameter estimators. The non-linear least-squares (NLS) fit approach is used to avoid computational issues associated with the maximum likelihood estimator (MLE) and moments-based estimator for parameters of the mixture model. For this purpose, a combination of moment fit and complementary CDF (CCDF) NLS fit methods is proposed. The simplex minimization algorithm is used to simultaneously obtain all parameters of the model. In the case of a single gamma probability density function, a zlog(z) method is derived. Firstly, simulated life tests based on a gamma population with different shape parameter values are worked out. Then, numerical illustrations show that both MLE and zlog(z) methods produce closer results. The proposed trimodal gamma distribution with moments NLS fit and CCDF NLS fit estimators is validated to be in qualitative agreement with different cell resolutions of the available IPIX database.

Keywords*—*CCDF, estimation, least squares, MLE, modeling, trimodal Gamma model, zlog(z).

1. Introduction

Modeling of unknown and non-stationary radar sea-clutter statistics is a serious research subject for target detection with a constant false alarm rate (CFAR). Compound Gaussian class distributions are useful models for sea-clutter observed by high resolution radars [1]. It is shown that gamma, inverse gamma, lognormal, and inverse Gaussian are efficient disturbances for the texture component characterizing the variability of both sea surface conditions and selected radar parameters [2]. A two-parameter family of continuous probability distributions on the positive real line is such a class of models.

Popular K distribution is widely applied in many disciplines of radar signal processing and is obtained from a gamma distributed texture component. The well-known Pareto type II model has been shown to occur as intensity distribution of the compound Gaussian process with an inverse gamma texture [3]. The compound Gaussian inverse Gaussian (CGIG) distribution is constructed if the modulation component follows the inverse Gaussian law [4]. However, in situations when we have a sequence of sea clutter with two or more distributions, compound-Gaussian models cited above fail to fit in with empirical data. This is particularly true for intelligent pixel X-band (IPIX) backscatter obtained from a small grazing angle and/or a low-range cell surface, using horizontal antenna polarization for transmit and receive [5], [6].

Various distributions that are probabilistic mixtures of other distributions have been proposed in the available literature [7]–[9]. Rosenberg *et al.* [7] analyzed the KK distribution for modeling the Ingara radar database with different scenarios. The addition of multiple looks and a thermal noise component is considered to produce greater accuracy of the mean and underlying shape parameters. In [8], a mixture of K and lognormal distributions is proposed to model the clutter data, the target data, or the mix of clutter and target data. The ML method using the expectation-maximization approach is presented for estimating the parameters of the mixture model. Experiments including synthetic aperture radar (SAR) data are conducted to show the effectiveness of the mixture model against KK and lognormal-lognormal distributions. In [9], a trimodal discrete (3MD) radar clutter model is utilized for modeling radar sea-clutter. A six-parameter mixture model is considered in paper [9], involving a multi-look Gaussian clutter scenario.

Parameter estimation of clutter models involving shape and scale parameters is an essential task, particularly if coherent or non-coherent detection processors are required to comply with the CFAR property. In [10], the authors presented moment-based methods including higher order moments, fractional order moments and log-moments for estimating K distribution parameters. A method for synthesizing correlated K-distributed random fields is also reported. The Gauss quadrature method based on Legendre and Laguerre polynomials is applied to constrained non-integer order moments, zlog(z) and MLE approaches for the estimation of Pareto type II, K and CGIG clutter parameters [11], [12]. Convergence is verified by using the first two integer order moments, in which the estimation problem is reduced to one dimension. Computer simulations are illustrated with known and unknown clutter-to-noise ratio (CNR).

In [9], four estimation methods based on a mixture of Gaussian distributed parameters are derived: two method-of-moments estimators (i.e. integer order moments and fractional order moments) and two other, based on NLS fit to the CCDF and MLE procedures. Accuracy and the computational time of the estimators is compared in terms of sample size, number of pulses and clutter parameter values. In [13], the Wilson-Hilferty normal-based approximation method is proposed to estimate the parameters of a gamma mixture model. The methodology uses a popular Gaussian mixture clustering algorithm, namely the CLUSTring (MCLUST) method and a confidence interval-based search approach to obtain the estimates. Performance comparisons with the existing expectation maximization (EM) approach are performed using both simulated and real-life datasets.

In this paper, we extend the recent work presented in [9] by using a mixture gamma distribution in a single-look transmission. It is an alternative solution proposed as a more accurate model in some scenes of the IPIX data (low range resolution) labeled trimodal gamma disturbance. Because gamma and incomplete gamma functions are presented in this model, the NLS fit approach is applied in this work to avoid computational issues associated with ML and moments estimators of the parameters of mixture models. For this purpose, the matching of moments and the CCDF approach is proposed, in which simplex minimization is used to obtain all model parameters simultaneously.

In the case of single gamma PDF, the zlog(z) method is derived. Simulated life tests performed with the use of the gamma population with different values of the shape parameters are worked out at first. Numerical illustrations show that both MLE and zlog(z) methods produce closer results. The proposed trimodal gamma distribution with moment NLS fit and CCDF NLS fit estimators is validated to be in qualitative agreement with different cell resolutions of the available IPIX database.

The paper is organized as follows. In Section 2, we initially review the mixture gamma distribution, the mixture gamma CDF and the mixture r-th raw moment's expression. Then, in Section 3, we present the MLE- and zlog(z)-based estimators for a single gamma distribution. After that, estimation procedures based on NLS CCDF fit and NLS moments fit are presented for the mixture of two and three gamma distributions. A series of numerical illustrations is given in Section 4. Finally, conclusions are outlined in Section 5.

2. Mixture Gamma Distribution

Analysis of high-resolution surveillance radar shows that statistical mixture models fit accurately with land and sea reverberation data. Gamma distribution is a two-parameter family of continuous probability distributions. Exponential distribution, Erlang distribution, and chi-square distribution are special cases of the gamma model. The gamma PDF of the random variable x is given by [14], [15]:

$$f(x;\beta,\alpha)=\frac{\beta^{\alpha}x^{\alpha-1}}{\Gamma(\alpha)}\mathrm{e}^{-\beta x}\ , \tag{1}$$

where x denotes clutter intensity (power), $\Gamma(.)$ is the gamma function, α is the shape parameter and β is the scale parameter. Its CDF is:

$$F(x;\alpha,\beta)=\gamma(\beta x,\alpha)\ , \tag{2}$$

where $\gamma(x,a)=\frac{1}{\Gamma(a)}\int\limits_0^x t^{\alpha-1}\mathrm{e}^{-t}\mathrm{d}t$ is the lower incomplete gamma function [16]. The r-th raw moment can be defined as:

$$E[x^r]=\beta^{-r}\frac{\Gamma(\alpha+r)}{\Gamma(\alpha)}\ . \tag{3}$$

As discussed, many models, including gamma distribution, fail to fit in well with real data that follow two or more densities. This occurs when radar echoes are observed from small range cells. To this effect, a general mixture gamma distribution could be capable of describing the majority of data scenarios given by [13]:

$$f(x;p_i,\beta_i,\alpha_i)=\sum_{i=1}^{n}p_i\frac{\beta_i^{\alpha_i}x^{\alpha_i-1}}{\Gamma(\alpha_i)}\mathrm{e}^{-\beta_i x}\ , \tag{4}$$

where p_i, $i=1,\ldots,n$ is the probability, and n is the number of gamma distributions. The corresponding CDF of Eq. (4) is written as:

$$F(x;p_i,\alpha_i,\beta_i)=\sum_{i=1}^{n}p_i\gamma(\beta_i x,\alpha_i)\ . \tag{5}$$

The r-th raw moment is given as a function of the gamma function with fractional variables:

$$E[x^r]=\sum_{i=1}^{n}p_i\beta_i^{-r}\frac{\Gamma(\alpha_i+r)}{\Gamma(\alpha_i)}\ . \tag{6}$$

The mixture model has $3n-1$ parameters, n discrete scale parameters, β_i, n discrete shape parameters α_i and $n-1$ probabilities p_i. Note that, Eqs. (4)–(6) reduce to the mixture expressions given by Bocquet *et al.* [9] with $\alpha_i=1$, $\beta_i=\frac{1}{b_i}$ and a single look scenario transmission.

3. Parameter Estimation

In this section, we consider some estimators based on MLE, log-moments, CCDF fit and moments fit of parameters for gamma distribution and a mixture of two and three gamma distributions.

3.1. MLE and zlog(z) Estimators of Gamma PDF

The MLE method has been used to estimate shape and scale parameters from a finite number of independent samples, $x_1, x_2, \dots, x_N$ observed [18]. This was achieved by maximizing the likelihood function (LF), so that:

$$\begin{cases} \ln(\hat{\alpha}) - \ln\left(\langle x\rangle\right) = \dfrac{\partial}{\partial\hat{\alpha}} \ln\left[\Gamma(\hat{\alpha}\right] + \left\langle \ln(x)\right\rangle = 0 \\ \hat{\beta} = \dfrac{\hat{\alpha}}{\langle x\rangle} \end{cases} . \quad (7)$$

Simplifying (7) allows to numerically determine $\hat{\alpha}$ as:

$$\begin{cases} \ln(\hat{\alpha}) - \ln\left(\langle x\rangle\right) - \psi(\hat{\alpha}) + \left\langle \ln(x)\right\rangle = 0 \\ \hat{\beta} = \dfrac{\hat{\alpha}}{\langle x\rangle} \end{cases} , \quad (8)$$

where $\langle . \rangle$ denotes the empirical mean and $\psi(.)$ is the psi function [16]. From formula (8), it can be easily seen that the resulting MLE for estimating α is non-linear and has no closed form solutions. Many papers reveal that MLE and zlog(z) estimators have approximate results, i.e. [17]. To obtain a short execution time, a closed form zlog(z) estimator is derived below, but with some mathematical manipulations of log-moments. Using the fact that:

$$\Gamma(x+1) = x\Gamma(x)\,, \quad \frac{\partial\Gamma(x)}{\partial x} = \Gamma(x)\psi(x)$$

and

$$psi(x+1) = \psi(x) + \frac{1}{x} \quad [16]\,,$$

the derivative of Eq. (3) with respect to the moment order r is found to be:

$$\begin{aligned} \frac{\partial\left\langle x^r\right\rangle}{\partial r} &= \left\langle x^r \ln(x)\right\rangle \\ &= \beta^{-r} \ln\left(\frac{1}{\beta}\right) \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} + \beta^{-r} \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} \psi(\alpha+r)\,. \end{aligned} \quad (9)$$

Using the first integer moment, $\langle x\rangle$ and Eq. (9) using $r = 0$ and $r = 1$, we write:

$$\begin{cases} \langle x\rangle = \beta^{-1}\alpha \\ \left\langle \ln(x)\right\rangle = \ln\left(\frac{1}{\beta}\right) + \psi(\alpha) \\ \left\langle x\ln(x)\right\rangle = \beta^{-1} \ln\left(\frac{1}{\beta}\right)\alpha + \beta^{-1}\alpha\psi(\alpha+1) \end{cases} , \quad (10)$$

Manipulating formula (10) finally produces the closed form zlog(z) estimator:

$$\begin{cases} \hat{\alpha} \left[\dfrac{\left\langle x\ln(x)\right\rangle}{\langle x\rangle} - \left\langle \ln(x)\right\rangle \right]^{-1} \\ \hat{\beta} = \dfrac{\hat{\alpha}}{\langle x\rangle} \end{cases} . \quad (11)$$

Equation (11) has no special functions making it easier to implement compared to (8).

3.2. CCDF Fit and Moments Fit Estimators

Recurrence relations of gamma and incomplete gamma functions with real variables given in Eqs. (5) and (6) are not easily tractable analytically. For the case of $n = 2$, the corresponding CCDF and moments expressions are:

$$\begin{cases} CCDF(T; p_i, \alpha_i, \beta_i) = p_i\left[1 - \gamma(\beta_1 T, \alpha_1)\right] \\ \qquad + (1-p_1)\left[1 - \gamma(\beta_2 T, \alpha_2)\right] \\ E[x^r] = p_1\beta_1^{-r}\dfrac{\Gamma(\alpha_1+r)}{\Gamma(\alpha_1)} + p_2\beta_2^{-r}\dfrac{\Gamma(\alpha_2+r)}{\Gamma(\alpha_2)} \end{cases} , \quad (12)$$

where T is the normalized detection threshold. Note that the manipulation of equations in formula (12), having five parameters, could not reduce the complexity of the estimation in one or two dimensions. Moreover, the use of log-moments does not resolve such an estimation issue. Therefore, the curve fitting technique is the most suitable approach for simultaneously estimating all model parameters. Its convergence is related to the application of an effective optimization algorithm. Based on expressions (5), (6) and (12), the fitness functions considered in this work are written as the sum of quadratic errors between empirical and theoretical quantities:

$$\begin{cases} Fitness_{CCDF} = \displaystyle\sum_{j=1}^{m_1} \left[real_{CCDF\,j} - \sum_{i=1}^{n} p_i\left[1 - \gamma(\beta_i T_j, \alpha_i)\right] \right]^2 \\ Fitness_{Moment} = \displaystyle\sum_{j=1}^{m_2} \left[\left\langle x^{r_j}\right\rangle - \sum_{i=1}^{n} p_i\beta_i^{-r_j} \frac{\Gamma(\alpha_i + r_j)}{\Gamma(\alpha_i)} \right]^2 \end{cases} , \quad (13)$$

subject to $\alpha_i > 0$, $\beta_i > 0$ and $\sum\limits_{i=1}^{n} p_i = 1$. m_1 and m_2 denote numbers of points used in the curves of objective functions. To reduce the research space dimension to $3n - 2$, β_1 expression may be inserted in (13) as a function of the empirical mean $\langle x\rangle$ and the model parameters, so that:

$$\beta_1 = \frac{p_1\alpha_1}{\langle x\rangle - \displaystyle\sum_{i=2}^{n} p_i \frac{\alpha_i}{\beta_i}} . \quad (14)$$

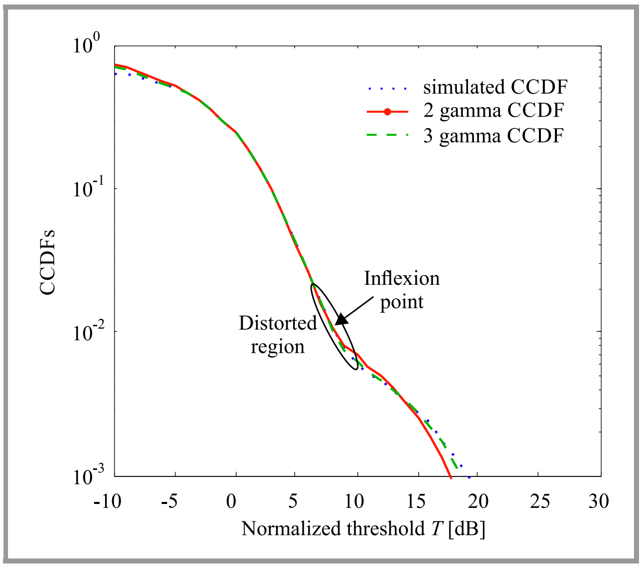


Fig. 1. Simulated CCDF from mixture of gamma distributed samples for: $p_1 = 0.3$, $p_2 = 0.5$, $p_3 = 0.2$, $\alpha_1 = 0.01$, $b_1 = 0.006$, $\alpha_2 = 1$, $b_2 = 1$, $\alpha_3 = 1.2$, and $b_1 = 0.48$.

To show the incapability of modeling radar sea-clutter with a single gamma distribution, we simulate, in Fig. 1, a mixture of Gaussian (i.e. $\alpha = 1$) and gamma (i.e. $\alpha \neq 1$) distributed clutter using the composition method [20]. The distorted region of the CCDF curve exhibited by one or more inflexion points is observed with the presence of Gaussian and non-Gaussian samples in the data sets.

From [5], it has been shown that several data scenes of the IPIX database have a similar nature as the CCDF curve shown in Fig. 1, i.e. a mixture of two or three distributions with different parameter values. Some of these scenarios will be presented in Section 4 by means of two estimators. Based on this, expression (13) is treated as a non-linear optimization problem involving four or seven parameters with a mixture of two and three gamma distributions, respectively.

Many optimization algorithms have been proposed in the literature [14], [19]. Some of them include genetic algorithms, biography-based optimization, simplex minimization, and particle swarm optimization [23]. For the purpose of fast convergence, one has to resort to iterative simplex minimization search based on the Nelder-Mead (NM) algorithm [19].

3.3. Simplex Minimization Algorithm

Simplex minimization was widely used for solving a variety of optimization problems. The NM simplex algorithm [21], [22] is an enormously popular search method for multidimensional unconstrained optimization. No derivative of the cost function is required, which makes the algorithm interesting for noisy problems. The NM algorithm falls in the more general class of direct search algorithms. It maintains simplexes being approximations of the optimal point. The vertices are sorted according to objective function values. The algorithm attempts to replace the worst vertex with a new point which depends on the worst point and the center of the best vertices. The goal of this part is to provide an NM direct search optimization method to solve the above constrained optimization problem given by expression (13).

This algorithms is based on the iterative update of a simplex made up of $m+1$ points $S = \{Vi\}$, $i = 1, 2, \ldots, m+1$. Each point in the simplex is called a vertex and is associated with a function value $f_i = f(V_i)$. It uses four parameters: coefficient of refection $\rho > 0$, expansion $\chi > 1$ with $\chi > \rho$, contraction $0 < \gamma < 1$ and shrinkage $0 < \sigma < 1$. The standard values of these coefficients are $\rho = 1$, $\chi = 2$, $\gamma = 0.5$ and $\sigma = 0.5$. Moves of the NM simplex are executed according to five main operators: reflection, expansion, inside contraction, outside contraction, shrink after inside contraction, and shrink after outside contraction. These components are interpreted by mathematical equations that are detailed in [21], [22].

The cost function of each estimator is given in expression (13). In the NM optimizer, the search of unknown parameters is carried out with constraints, because all model parameters are real positive. In particular, the two probabilities of trimodal gamma distribution must be limited between 0 and 1. In the estimation procedure, scale and shape parameters in (13) are restricted, in the following manner, to the range of $[0, 10]$:

$$\begin{cases} \alpha_i = \max\left[\min(\alpha_i, 10), 0\right] \\ \beta_i = \max\left[\min(\beta_i, 10,)0\right] \\ p_1 = \max\left[\min(p_1, 1), 0\right] \\ p_2 = \max\left[\min(p_2, 1), 0\right] \\ p_2 = \max\left[\min(p_2, 1), 0\right] \\ p_2 = \max\left[\min(p_2, 1 - p_1), 0\right] \end{cases}, \qquad i = 1, \ldots, n. \quad (15)$$

The constraint of the probabilities, p_1 and p_2 in the interval of $[0, 1]$ is also considered in (15). Figure 2 summarizes different steps of the optimization of formula (13) with the constraint procedure given in (15) using an SM based on the NM algorithm.

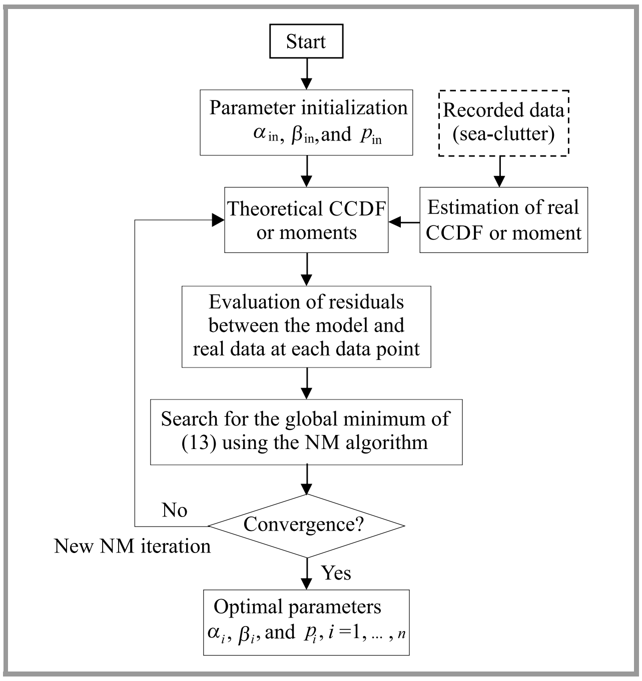


Fig. 2. Flowchart of the simplex minimization algorithm of expression (13).

4. Numerical Illustrations

In this section, we assess the above MLE, zlog(z), SM CCDF fit and SM moments fit estimators with the use of both simulated and real IPIX databases. One, two and three component gamma distributions are considered for estimation and modeling of some scenarios of the radar IPIX database.

4.1. Gamma Model Case

Consider a random sample of size $N = 1000$ from a finite gamma distribution with its PDF as defined in Eq. (1). A number $n = 1000$ of Monte Carlo runs is used to average MSE and bias criterion tests. MLE and zlog(z) methods given by formulas (8) and (11) are executed to estimate gamma PDF parameters.

Through this simulation study, the performance of zlog(z) estimator is assessed vis-à-vis the MLE method in terms of MSE and bias estimates of the shape parameter as depicted in Fig. 3. From these plots, a comparison of estimates reveals that both approaches produce closer results.

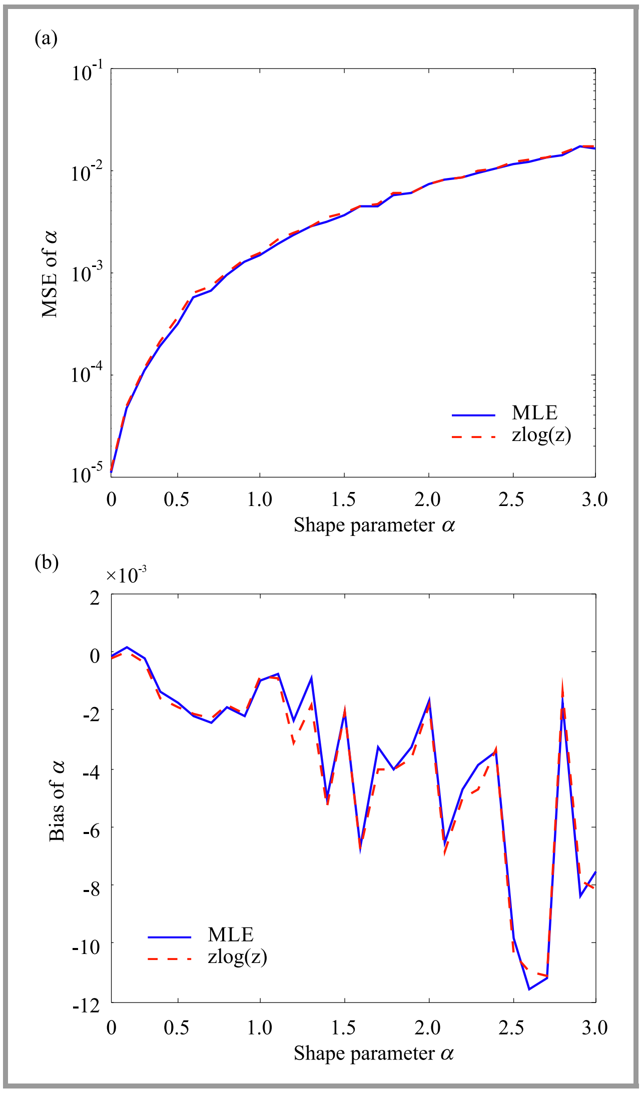


Fig. 3. Estimation comparison of MLE and zlog(z) methods of α for $N = 1000$ and $n = 1000$: (a) MSE metric test, (b) bias metric test.

An implementation of the zlog(z) methodology will provide estimates with a shorter execution time than the MLE method. The zlog(z) procedure may be attributed to the fact that it uses, in the computation of the adaptive CFAR detection, a threshold for radar targets embedded in gamma distributed clutter with unknown parameters. The modeling of IPIX data with a resolution of 30 m, VV polarization and 13th range cell is checked, as shown in Fig. 4, using the two estimators. Figure 4a shows a comparison of empirical and estimated moments with the moments' orders between $r = 0.1$ and $r = 2$. Figure 4b highlights the fit of CCDF with real data always using the two estimators.

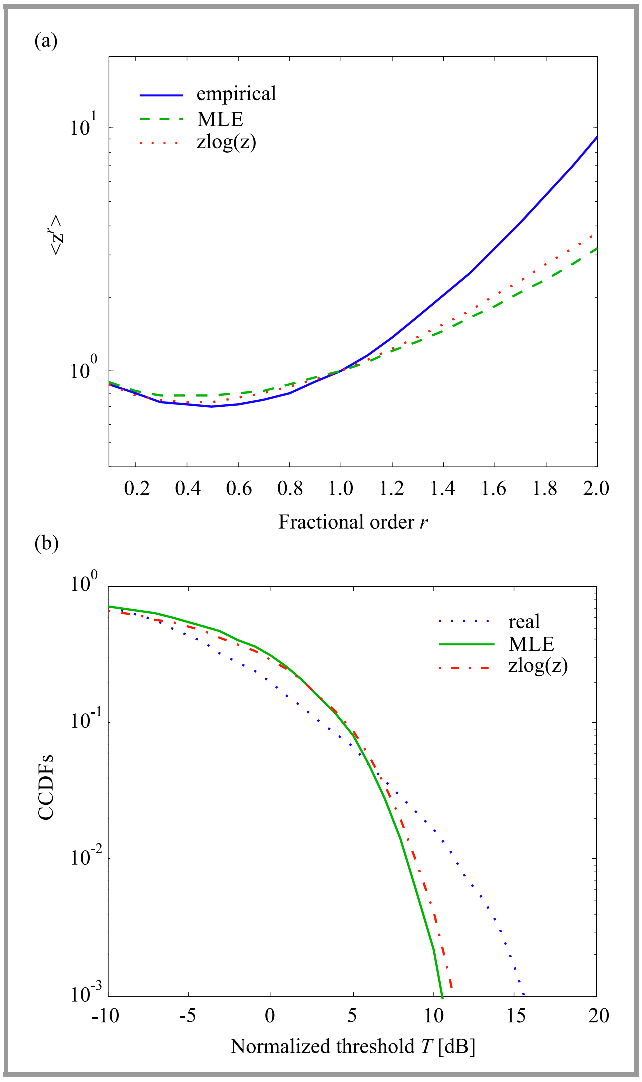


Fig. 4. Modeling of IPIX data with a resolution of 30 m, VV polarization and 13th range cell: (a) moments fit test, (b) CCDF fit test.

It is clearly seen that a single gamma PDF is not capable of modeling such a data scene. Unfortunately, several scenarios of IPIX data cannot be fitted by a single gamma PDF as well. For the purpose of a goodness-of-fit with real data, we consider in the following subsection a mixture of two and three gamma distributions.

4.2. Mixture Gamma Model Case

In this subsection, SM moment fit and SM CCDF fit approaches for the estimation of a mixture of gamma distributions are implemented on IPIX real-life datasets, as described in [12], [24]. The IPIX radar experimental data we processed were collected at Grimsby, Ontario, Canada, from the Communications Research Laboratory, McMaster University.

Sea clutter sets to be used here are measured with the use of the McMaster IPIX radar, a fully coherent X-band radar, with advanced features, such as dual transmit/receive polarization, frequency agility, and stare/surveillance mode.

It is extremely versatile, as each feature is highly adjustable through software in the control computer. Originally, the IPIX radar was shorthand for “ice multi-parameter imaging X-band” radar, called that way as the radar was designed for the detection of growlers, i.e. small pieces of ice breaking away from icebergs. After major upgrades introduced between 1993 and 1998, the high-resolution data collected by the IPIX radar became a benchmark for testing intelligent detection algorithms. Accordingly, the adjustable meaning of the IPIX acronym was changed to “intelligent pixel processing X-band” radar, where the term “pixel” refers to a picture element [24]. The IPIX radar is equipped with computer control and digital data acquisition capabilities.

In 1998, a database of high-resolution radar measurements was collected in Grimsby, on the shore of Lake Ontario, between Toronto and Niagara Falls, Canada. The 222 data sets in this database focus specifically on the presence of floating objects (targets) of varying size, observed under varying weather conditions. A graphical representation of all 222 datasets and images including radar return plots and time Doppler spectra are available in [24]. The characteristic features of the IPIX radar and the environmental conditions under which the radar data was collected are also presented in [24].

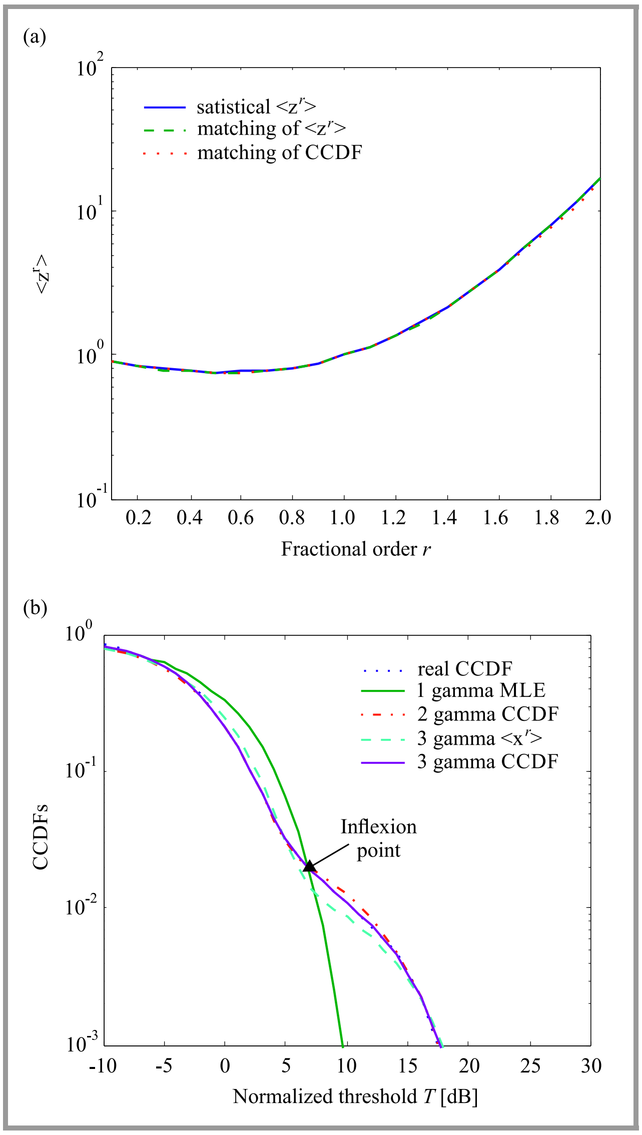


Fig. 5. Modeling of IPIX data with a resolution of 3 m, HH polarization and 17th range cell: (a) moments fit test for the case of 3 gamma PDFs, (b) CCDF fit test of 1 gamma, 2 gamma and 3 gamma PDFs.

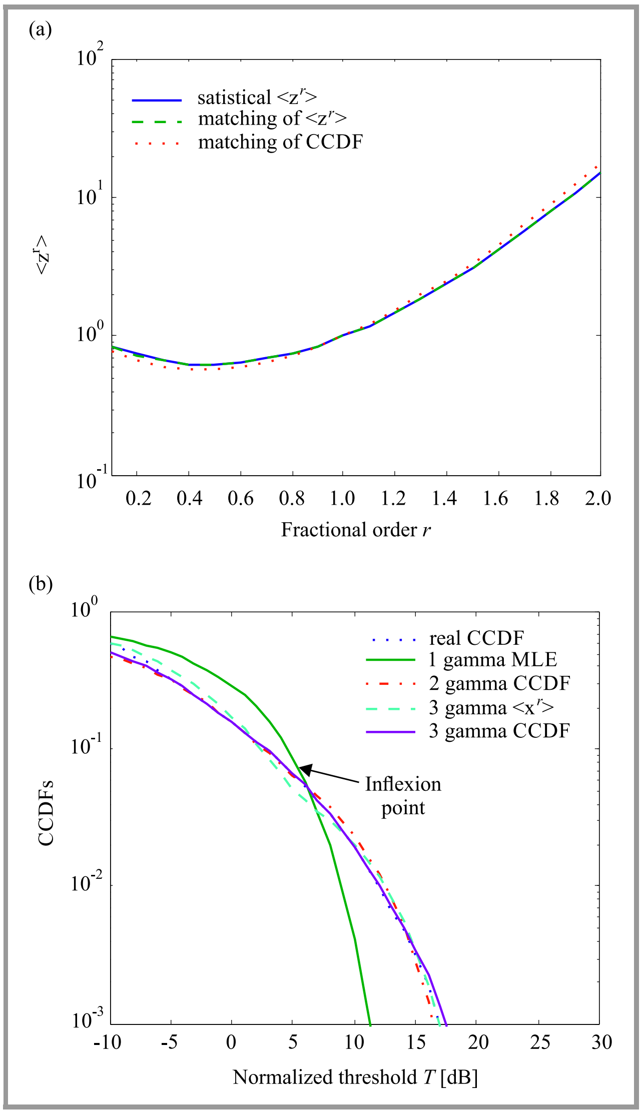


Fig. 6. Modeling of IPIX data with a resolution of 15 m, HH polarization and 5th range cell: (a) moments fit test for the case of 3 gamma PDFs, (b) CCDF fit test of 1 gamma, 2 gamma and 3 gamma PDFs.

The above estimators are executed according to the flowchart presented in Fig. 2. To start the optimization, CCDF values in (13) are considered between $10^{-0.55}$ and 10^{-3}. On the other side, 200 moments values with orders between $r = 0.1$ and $r = 2$ are taken in formula (13). Our first study concerns modeling of IPIX data with a high resolution of 3 m, HH (horizontal-horizontal) antennas polarization and 17th range cell using one gamma, two gamma and three gamma PDFs defined by Eqs. (1) and (4), as shown in Fig. 5. Moment fit curves are depicted for the case of 3 gamma PDFs – see Fig. 5a – and CCDF fit curves are presented for the case of 1 gamma, 2 gamma and 3 gamma PDFs (see Fig. 5b). From Fig. 5a, it is observed that the mixture of gamma PDF moments has smaller errors between empirical and estimated moment curves. This indicates that this scenario of IPIX data has probably a mixture of Gaussian and non-Gaussian clutter. From Fig. 5b, the best tail fitting with such data is obtained by a mixture of three gamma distributions using the CCDF fit estimator. Moreover, it is clearly seen in Fig. 6 that the IPIX data with a resolution of 15 m, HH polarization and 5th range cell are well modeled by a mixture of three gamma components with the moment fit estimator. This means that the proposed estimators converge to the best estimates of respective 7 parameters.

Figure 7, taking into consideration the modeling study of real data with a low resolution of 30 m, HH polariza-

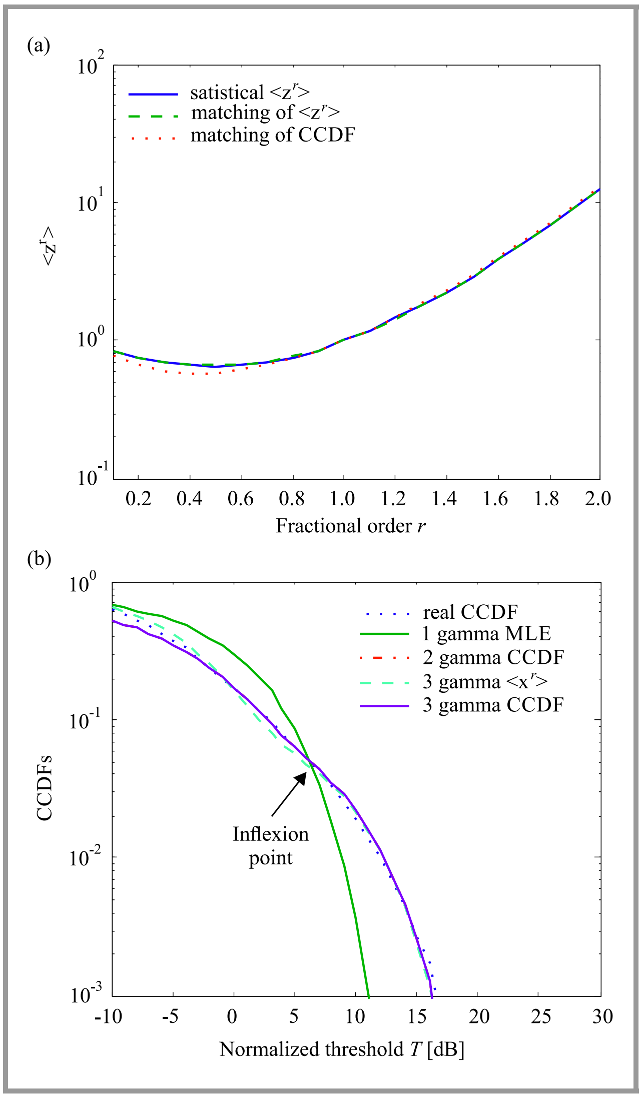


Fig. 7. Modeling of IPIX data with a resolution of 30 m, HH polarization and 11th range cell: (a) moments fit test for the case of 3 gamma PDFs, (b) CCDF fit test of 1 gamma, 2 gamma and 3 gamma PDFs.

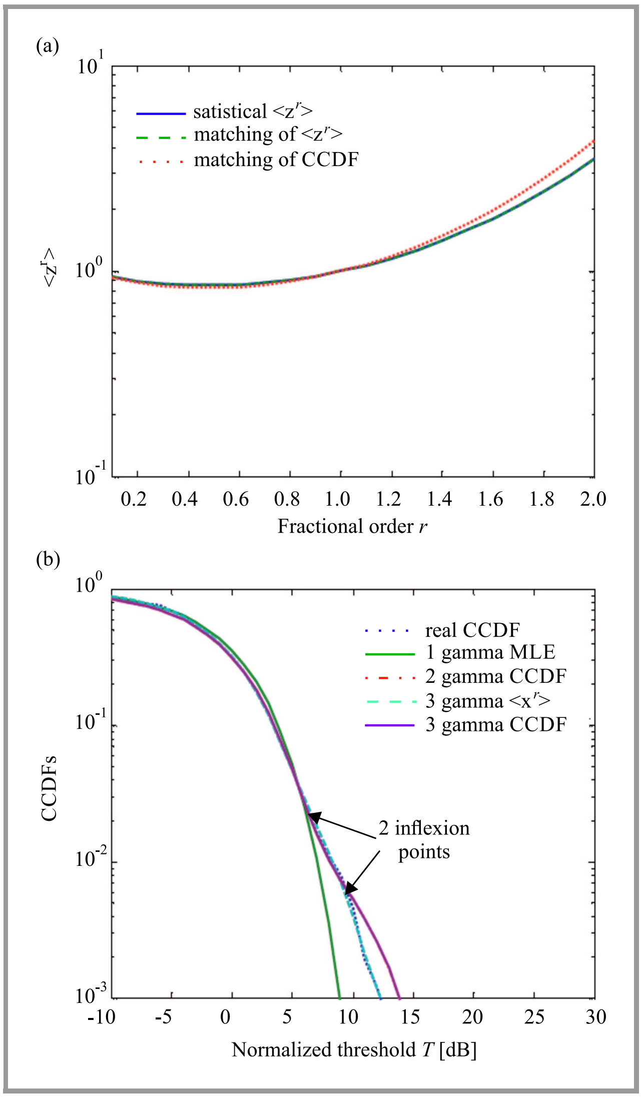


Fig. 8. Modeling of IPIX data with a resolution of 3 m, VV polarization and 21st range cell: (a) moments fit test for the case of 3 gamma PDFs, (b) CCDF fit test of 1 gamma, 2 gamma and 3 gamma PDFs.

Table 1
Comparison of complexity (number of FLOP) and performance of different model structures

Estimator / IPIX data	SQE using 2 gamma CCDF	SQE using 3 gamma CCDF	SQE using 3 gamma $\langle x^r \rangle$
HH, 3 m and 17th cell	1.98×10^{-4}	2.32×10^{-5}	0.0195
HH, 15 m and 5th cell	8.48×10^{-4}	3.57×10^{-5}	0.0086
HH, 30 m and 11th cell	5.34×10^{-4}	5.86×10^{-4}	1.98×10^{-4}
VV, 3 m and 21st cell	5.40×10^{-4}	5.37×10^{-4}	5.87×10^{-7}

tion and 11th range cell, shows the efficiency of both SM CCDF fit and SM moment fit estimators. Here, two and three gamma component models offer almost similar results. For the case of modeling IPIX data with a resolution of 3 m, VV (vertical-vertical) antenna polarization and 21st range cell, Fig. 8 depicts the different curves of moments and CCDFs respectively. It is remarkable that the mixture of three gamma PDFs additionally ensures goodness-of-fit with real data. This model has the ability to track empirical CCDF in the presence of different inflexion points. From the above, it is shown that the SM moment fit and SM CCDF fit ensure proper estimation results.

From the above results, the trimodal gamma distribution fits IPIX data in most cases in terms of different range resolutions. Finally, the comparison of fitness function values, i.e. the sum of quadratic errors (SQE) given in (13) is highlighted in Table 1. Residuals corresponding to the CCDF NLS fit method are calculated for $10^{-0.55} < CCDF < 10^{-3}$. On the other hand, residuals corresponding to the moments, i.e. $\langle x^r \rangle$, the NLS fit method, are calculated for $0.01 < r < 2$. From Table 1, improved results are offered by the CCDF NLS fit method for a mixture of 3 gamma populations.

5. Conclusion

Two methods for estimating parameters of a mixture of gamma PDFs were introduced in this paper. One, two and three gamma component models were considered for the purpose of parameter estimation and modeling tasks. The proposed zlog(z) approach does not involve the use of numerical calculi, but it is used for the case of a single gamma PDF.

Experiments showed that the latter is not suitable for modeling IPIX data. In order to show the efficiency of the proposed SM moment fit and SM CCDF fit, modeling of several data scenes have been tested by means of two and three gamma component models. From numerical illustrations, it was shown that the best fit with radar echoes was achieved by a mixture of three gamma PDFs.

The computational time associated with the proposed moment and CCDF fits is relatively high owing to the fact that they search seven dimensional spaces. To conclude, the application of the proposed mixture of three gamma PDFs with SM CCDF fit and SM moments fit methods is a novel approach to modeling IPIX radar echoes. In fact, this model produces excellent tail fitting of the recorded data.

References

[1] V. Anastassopoulos, G. A. Lampropoulos, A. Drosopoulos, and M. Rey, "High resolution radar clutter statistics", *IEEE Trans. on Aerosp. and Electron. Syst.*, vol. 35, no. 1, 1999 (DOI: 10.1109/7.745679).

[2] I. Chalabi and A. Mezache, "Estimators of compound Gaussian clutter with log-normal texture", *Remote Sensing Lett.*, vol. 10, no. 7, pp. 709–716, 2019 (DOI: 10.1080/2150704X.2019.1601275).

[3] V. G. Weinberg, S. D. Howard, and C. Tran, "Bayesian framework for detector development in Pareto distributed clutter", *IET Radar Sonar & Navig.*, vol. 13, no. 9, pp. 1548–1555, 2019 (DOI: 10.1049/iet-rsn.2018.5635).

[4] H. Yu, P. L. Shui, and Y. T. Huang, "Low-order moment-based estimation of shape parameter of CGIG clutter model", *Electron. Lett.*, vol. 52, no. 18, pp. 1561–1563, 2016 (DOI: 10.1049/el.2016.2248).

[5] A. Mezache, F. Soltani, M. Sahed, and I. Chalabi, "Model for non-Rayleigh clutter amplitudes using compound inverse Gaussian distribution: An experimental analysis", *IEEE Trans. on Aerosp. and Electron. Syst.*, vol. 51, no. 1, 2015 (DOI: 10.1109/TAES.2014.130332).

[6] A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "High resolution sea clutter data: statistical analysis of recorded live data", *IEE Proc. – Radar, Sonar & Navig.*, vol. 144, no. 3, pp. 121–130, 1997 (DOI: 10.1049/ip-rsn:19971107).

[7] L. Rosenberg, D. J. Crisp, and N. J. Stacy, "Analysis of the KK-distribution with medium grazing angle sea-clutter", *IET Radar, Sonar & Navig.*, vol. 4, no. 2, pp. 209–222, 2010 (DOI:10.1049/iet-rsn.2009.0096).

[8] X. Zhou, R. Peng, and C. Wang, "A two-component K-lognormal mixture model and its parameter estimation method", *IEEE Trans. on Geosci. and Remote Sens.*, vol. 53, no. 5, 2015 (DOI: 10.1109/TGRS.2014.2363356).

[9] S. Bocquet, L. Rosenberg, and C. H. Gierull, "Parameter estimation for a compound radar clutter model with trimodal discrete texture", *IEEE Trans. on Geosci. and Remote Sens.*, vol. 58, no. 10, 2020 (DOI: 10.1109/TGRS.2020.2979449).

[10] J. R. Nicholas, "Estimating the parameters of the K distribution in the intensity domain", Rep. DSTO-TR-0839, DSTO Electronics and Surveillance Research Laboratory, pp. 1–76, 1999, South Australia [Online]. Available: https://apps.dtic.mil/sti/pdfs/ADA368069.pdf

[11] S. Bocquet, "Parameter estimation for Pareto and K distributed clutter with noise", *IET Radar, Sonar & Navig.*, vol. 9, no. 1, pp. 104–113, 2015 (DOI: 10.1049/iet-rsn.2014.0148).

[12] A. Mezache, A. Gouri, and H. Oudira, "Parameter estimation of CGIG clutter plus noise using constrained NIOME and MLE approaches", *IET Radar, Sonar & Navig.*, vol. 12, no. 2, pp. 176–185, 2018 (DOI: 10.1049/iet-rsn.2017.0234).

[13] R. Vani Lakshmi and V. S. Vaidyanathan, "Parameter estimation in gamma mixture model using normal-based approximation", *J. of Statis. Theory and Appl.*, vol. 15, no. 1, pp. 25–35, 2016 (DOI: 10.2991/jsta.2016.15.1.3).

[14] Q. Xianxiang, Z. Shilin, Z. Huanxin, and G. Gui, "A CFAR detection algorithm for generalized gamma distributed background in high-resolution SAR images", *IEEE Geosci. and Remote Sens. Lett.*, vol. 10, no. 4, 2013 (DOI: 10.1109/LGRS.2012.2224317).

[15] É. Magraner, N. Bertaux, and P. Réfrégier, "A new CFAR detector in gamma-distributed non homogeneous backgrounds", in *Proc. of 16th Eur. Sig. Process. Conf. EUSIPCO 2008*, Lausanne, Switzerland, 2008 [Online]. Available: https://www.eurasip.org/Proceedings/Eusipco/Eusipco2008/papers/1569104566.pdf

[16] M. Abramowitz and I. A. Stegun, Ed., *Handbook of Mathematical Functions*. New York: Dover Publications, Inc., 1970 (ISBN: 978-0486612720).

[17] D. Blacknell and R. J. A. Tough, "Parameter estimation for the K-distribution based on [z log(z)]", *IEE Proc. – Radar, Sonar & Navig.*, vol. 148, no. 6, 309–312, 2001 (DOI:10.1049/ip-rsn:20010720).

[18] T. P. Minka, "Estimating a Gamma distribution", 2002 [Online]. Available: https://tminka.github.io/papers/minka-gamma.pdf

[19] S. A. Hamadi, A. Chouder, M. M. Rezaoui, S. Motahhir, and A. M. Kaddouri, "Improved hybrid parameters extraction of a PV module using a moth flame algorithm", *Electronics*, vol. 10, 2021 (DOI: 10.3390/electronics10222798).

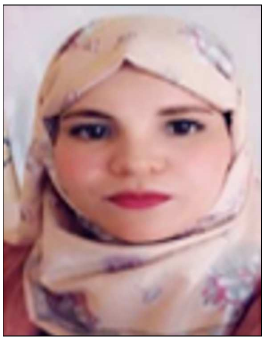
[20] D. P. Kroese, T. Taimre, and Z. I. Botev, Handbook of Monte Carlo Methods. New York, NY, USA: Wiley, 2011 (ISBN: 9780470177938).

[21] J. A. Nelder and R. Mead, "A simplex method for function minimization", *The Computer J.*, vol. 7, no. 4, pp. 308–313, 1965 (DOI: 10.1093/comjnl/7.4.308).

[22] M. Baudin, "Nelder-Mead user's manual" [Online]. Available: https://www.scilab.org/sites/default/files/neldermead.pdf

[23] Shu-Kai, S. Fan, and E. Zahara "A hybrid simplex search and particle swarm optimization for unconstrained optimization", *European J. of Oper. Res.*, vol. 181, no. 2, pp. 527–548, 2007 (DOI:10.1016/j.ejor.2006.06.034).

[24] R. Bakker and B. Currie, "The McMaster IPIX radar sea clutter database", 2001 [Online]. Available: http://soma.mcmaster.ca/ipix/



Zakía Terki received her M.Sc. degree in Telecommunications in 2019 from M'sila University. Since, 2019–2020, she has been a Ph.D. student at the University of Ouargla, Algeria. Her current research interests include modeling of high resolution sea clutter, estimating parameters of statistical distributions and CFAR detection based on Neayman-Pearson and Bayesian approaches.

E-mail: terki.zakia@univ-ouargla.dz
Laboratoire de Génie Electrique, LAGE
Département d'Electronique
Université Kasdi Merbah Ouargla
Ouargla, Algéria

Amar Mezache received his B.Eng. degree in Electronics/Systems Control, the M.Sc. degree and the Ph.D. degree, both in Signal Processing, from the University of Constantine, Algeria in 1997, 2000 and 2007, respectively. He joined the Department of Electronics, University of M'sila, in 2004 as a full-time professor, where he has been teaching radar signal detection and estimation, digital signal processing, power electronics and control of DC/AC motors. His current research interests include modeling and estimating parameters of high-resolution sea clutter, radar CFAR detection, and application of artificial intelligence in radar signal processing. He holds the position of chairman of the Department's Scientific Committee.

E-mail: amar.mezache@univ-msila.dz
Département d'Electronique
Université Mohamed Boudiaf M'sila
M'sila, Algéria

Algérie Laboratoire SISCOM
Université de Constantine
Constantine, Algéria

Fouad Chebbara received his baccalaureate certificate in 1996, and his communication engineer diploma in 2001. He earned his M.Sc. degree in 2006, and a Ph.D. in 2011. He completed his university habilitation program in 2016 and became a professor in 2020. His areas of research are in telecommunications, microband antennas, signal processing and computer science. He holds the position of chairman of the Department's Scientific Committee.

E-mail: chebbara.fouad@univ-ouargla.dz
Laboratoire de Génie Electrique, LAGE
Département d'Electronique
Université Kasdi Merbah Ouargla
Ouargla, Algérie

Paper

RED-LE: A Revised Algorithm for Active Queue Management

Samuel O. Hassan

Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Nigeria

https://doi.org/10.26636/jtit.2022.160022

Abstract—The random early detection (RED) algorithm was developed in 1993. Nearly three decades later, several improved variants have been proposed by scientists. The use of a (pure) linear function for computing packet drop probability has turned out to be a disadvantage, leading to the problem of large delays. Such a problem may be addressed by using linear and non-linear (i.e. as exponential) packet drop probability functions. This paper proposes a revised RED active queue management algorithm named RED-linear exponential (RED-LE). This variant involves an interplay of linear and exponential drop functions, in order to improve the performance of the original RED algorithm. More importantly, at low and moderate network traffic loads, the RED-LE algorithm employs the linear drop action. However, for high traffic loads, RED-LE employs the exponential function for computing the packet drop probability rate. Experimental results have shown that RED-LE effectively controls congestion and offers an improved network performance under different traffic loads.

Keywords—active queue management, network congestion, routers, RED-LE, simulation.

1. Introduction

Network congestion may be described as a condition in which the amount of incoming data packets (generated traffic) is greater than the amount that the network's available resources are capable of accommodating [1]–[5]. The problem of network congestion affects the quality of service (QoS), as it leads to high packet delays, loss rates, and low throughput [2], [4], [5]–[9].

A router plays an important role in the process of controlling network congestion, as it allows to achieve improved network performance rates [9]. Router-based congestion control algorithms, such as active queue management (AQM), effectively circumvent network congestion by dropping packets at an early stage, before the buffer becomes full and sends a feed- back signal to sources in order to reduce their transmission rates [10]–[12].

The most prominent type of an AQM algorithm is random early detection (RED), developed by Floyd and Jacobson in 1993 [11]. RED continues to serve as a basis for many new AQM algorithms [12]. Upon the arrival of each packet at the router, RED updates the average queue size (denoted *avg*) which is used as an indicator for congestion detection. To perform this computation, the current status of the queue is examined.

If the router's queue is non-empty, *avg* value is determined using the exponential weighted moving average (EWMA) mechanism in the following manner:

$$avg = (1 - W_q)avg' + (W_q \times q_{cur}) \ , \tag{1}$$

where $W_q \in [0, 1]$ represents a preset weighting factor, avg' represents the previously computed average queue size, and q_{cur} represents the current queue size.

However, if the router's queue is empty, *avg* is determined as:

$$avg = (1 - W_q)^n \times avg' \ , \tag{2}$$

with

$$n = f(q\_current\_time - q\_idle\_time) \ , \tag{3}$$

where $q\_current\_time$ denotes the current time, q_idle_time denotes the beginning of queue idle time, and $f(t)$ denotes a linear function of time t.

The probability of dropping a packet in RED depends on *avg* in the following manner:

$$P_b = \begin{cases} 0, & \text{if } avg \in [0, minTH) \\ maxP(\frac{avg - minTH}{maxTH - minTH}), & \text{if } avg \in [minTH, maxTH) \\ 1, & \text{if } avg \geq maxTH \end{cases} \ , \tag{4}$$

where $minTH$ is the router's minimum queue threshold, $maxTH$ represents the router's maximum queue threshold, $maxP$ represents the maximum packet drop probability, and P_b stands for the initial packet dropping probability.

In RED, if $avg \in [0, minTH)$ then no packet will be dropped and if $avg \in [minTH, maxTH)$, then the packets are randomly dropped with the probability of:

$$P_b = maxP(\frac{avg - minTH}{maxTH - minTH}) \ . \tag{5}$$

Finally, if $avg \geq maxTH$, then the packet is forced to be dropped, with a probability of one. The final packet drop probability P_a therefore given by:

$$P_a = \frac{P_b}{1 - count \times P_b} \ , \tag{6}$$

where *count* represents the number of packets that arrived since the last dropped packet.

There are several models in literature that modify the linear probability function of RED algorithm in an attempt to overcome its weaknesses. In this paper, another improvement-aiming modification is suggested, known as the random early detection-linear exponential (RED-LE) algorithm. The RED-LE algorithm uses both linear and exponential packet drop functions instead of a (pure) linear packet drop probability function of RED, while retaining RED's other characteristics.

The rest of the paper is organized as follows. A review of related works on the RED algorithm is provided in Section 2. A description of RED-LE is given in Section 3. The simulation configuration is presented and the results are discussed in Section 4. Finally, a brief conclusion is given in Section 5.

2. Related Works

To increase the throughput of RED, Floyd developed, in [13], the gentle RED (GRED) variant in which a linear function is employed to compute the packet drop probability when *avg* lies within the *minTH* and *maxTH* queue threshold range – Eq. (5) – while another linear function is employed to compute the packet drop probability when *avg* is within the *maxTH* and 2 × *maxTH* threshold range – Eq. (7):

$$P_b = maxP + (1 - maxP)\frac{avg - minTH}{maxTH} \ . \tag{7}$$

To ensure higher stability, Giménez *et al.* developed a new RED variety called BetaRED, which involves a beta distribution function to compute the packet drop probability instead of a linear function when *avg* value is within the *minTH* and *maxTH* threshold range [12].

In [14], an attempt to reduce the number of input parameters for RED was made by Abdel-Jaber, known as Exponential RED (RED_E) in which a (pure) exponential drop function given in Eq. (8) is employed to compute the packet drop probability when *avg* value is between the *minTH* and *maxTH* queue thresholds.

$$P_b = \begin{cases} 0, & \text{if } avg \in [0, minTH) \\ \dfrac{e^{avg} - e^{minTH}}{e^{maxTH} - e^{minTH}}, & \text{if } avg \in [minTH, maxTH) \\ 1, & \text{if } avg \geq maxTH \end{cases} \ . \tag{8}$$

To increase RED's throughput, Zhang *et al.* proposed, in [15], the MRED variety in which a quadratic function is employed to compute the packet drop probability when *avg* is within the *minTH* and *maxTH* queue threshold range given by Eq. (9), while a linear function is employed to compute the packet drop probability when *avg* is within the *maxTH* and 2 × *maxTH* queue threshold range as stated in Eq. (10).

$$P_b = maxP\frac{avg^2 - minTH^2}{maxTH^2 - minTH^2} \ , \tag{9}$$

$$P_b = maxP + (1 - maxP)\frac{avg - minTH}{maxTH} \ . \tag{10}$$

To achieve a trade-off between delay and throughput performance metrics, Paul *et al.* suggested, in [16], the Smart RED (SmRED) scheme, given in Eq. (11), in which a quadratic function is employed to compute the packet drop probability when *avg* lies within the *minTH* and *Target* queue threshold range, while a linear function is employed to compute the packet drop probability when *avg* lies within the *Target* and *maxTH* queue threshold range:

$$P_b = \begin{cases} 0, & \text{if } avg \in [0, minTH) \\ maxP\left(\dfrac{avg - min_{TH}}{max_{TH} - min_{TH}}\right)^2, & \text{if } avg \in [minTH, Target) \\ maxP\sqrt{\dfrac{avg - min_{TH}}{max_{TH} - min_{TH}}}, & \text{if } avg \in [Target, maxTH) \\ 1, & \text{if } avg \geq maxTH \end{cases} \tag{11}$$

in which

$$Target = minTH + \frac{maxTH - minTH}{2} \ . \tag{12}$$

In order to obtain improved throughput, Suwannapong and Khunboa developed, in [17], yet another variety of RED, named Congestion Control RED (CoCo-RED) which involves both linear and an exponential drop functions. The linear function is employed when *avg* value is within the *minTH* and *maxTH* queue threshold range, while the exponential function is employed when *avg* value is within the *maxTH* and *K* queue threshold range:

$$P_b = \begin{cases} 0, & \text{if } avg \in [0, minTH) \\ maxP\dfrac{avg - minTH}{maxTH - minTH}, & \text{if } avg \in [minTH, maxTH) \\ ab^{avg}, & \text{if } avg \in [maxTH, K) \end{cases} \tag{13}$$

in which

$$a = \frac{1}{\left(e^{\frac{\ln(1/maxP)}{K - maxTH}}\right)^{maxTH}} \times maxP \tag{14}$$

and

$$b = e^{\frac{\ln(1/maxP)}{K - maxTH}} \ . \tag{15}$$

Feng *et al.* [18] proposed a three-section RED (TRED) which employs the usage of a non-linear drop action, a linear drop action, and a non-linear drop function for low, moderate, and high buffer occupancy rates, respectively. TRED results in high throughput at high traffic loads and achieves a reduced delay at high traffic loads.

In order to increase throughput, Zhou *et al.* proposed, in [19], another variant named non-linear RED (NLRED) in which a quadratic function is employed to compute the packet drop probability when *avg* value is between the *minTH* and *maxTH* queue thresholds.

To reduce the packet loss rate, Kumhar *et al.* developed, in [20], quadratic RED (QRED) in which a quadratic function is deployed to compute the packet drop probability when *avg* value lies within the $minTH - maxTH$ queue threshold range:

$$P_b = \left(\frac{avg - minTH}{K - minTH}\right)^2 \tag{16}$$

or

$$P_b = 1 - \left(\frac{K - avg}{K - minTH}\right)^2 , \tag{17}$$

in which K represents the buffer size.

To achieve increased throughput, Adamu *et al.* developed, in [21], the flexible RED (FXRED) algorithm. At low and moderate traffic loads, i.e. when *avg* value lies within the $minTH$ and Δ queue threshold range, FXRED uses a non-linear function to drop packets. However, at high traffic loads, i.e. when *avg* value lies within the Δ and $maxTH$ queue threshold range, FXRED switches to a linear pattern for aggressive drop action. In FXRED, Δ was chosen in the following manner:

$$\Delta = \frac{minTH + maxTH}{2} . \tag{18}$$

Furthermore, to improve throughput performance, the self-adaptive RED (SARED) algorithm developed by Adamu *et al.* in [4] is quite similar to RED, except that either a quadratic or linear drop function is employed when *avg* lies within the *minTH* and *maxTH* threshold range. The quadratic drop function is employed for lower and moderate buffer occupancies, while a linear drop function is employed for higher buffer occupancies, respectively.

3. Random Early Detection-Linear Exponential (RED-LE)

The proposed algorithm is named random early detection-linear exponential (simply denoted by RED-LE). This revised RED algorithm involves an interplay of linear and exponential drop functions in order to increase the performance of the original RED algorithm. The improved packet drop probability is:

$$P_b = \begin{cases} 0, & \text{if } avg \in [0, minTH) \\ 2maxP\dfrac{avg - minTH}{maxTH - minTH}, & \text{if } avg \in [minTH, Target) \\ \mathrm{e}^{\log(maxP)\frac{2(maxTH - avg)}{maxTH - minTH}}, & \text{if } avg \in [Target, maxTH) \\ 1, & \text{if } avg \geq maxTH \end{cases} \tag{19}$$

in which

$$Target = \frac{maxTH + minTH}{2} . \tag{20}$$

Considering the curve given in Fig. 1, RED-LE breaks the section between $minTH$ and $maxTH$ queue thresholds into two parts which include both a linear drop function and an exponential drop function, such that:

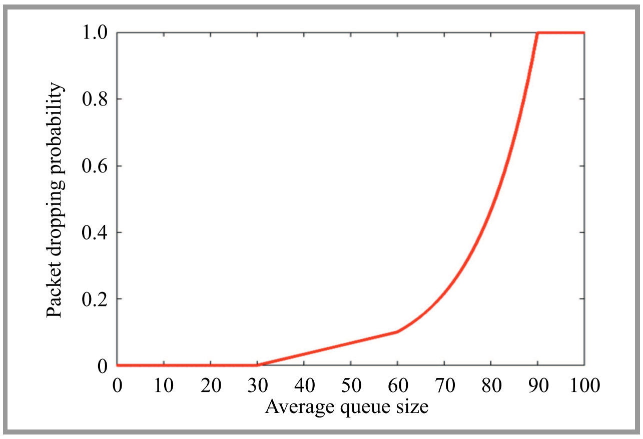


Fig. 1. RED-LE's drop probability function curve.

- At low and moderate traffic loads which account for cases with $minTH \leq avg < Target$, the packet drop function is expressed as:

  $$P_b = 2maxP\frac{avg - minTH}{maxTH - minTH} . \tag{21}$$

  Target is a mid-point threshold defined according to Eq. (20).

- At high traffic load which account for cases where $Target \leq avg < maxTH$, the packet drop function is expressed as:

  $$P_b = \mathrm{e}^{\log(maxP)\frac{2(maxTH - avg)}{maxTH - minTH}} . \tag{22}$$

  Using Eq. (22), a more aggressive drop action will be achieved at high load.

It is worth to mention that *Target* serves the purpose of distinguishing between two traffic scenarios: lower and moderate buffer occupancies, and higher buffer occupancies. A detailed pseudo-code for RED-LE is presented in Algorithm 1.

4. Simulations

In this section, the proposed RED-LE AQM algorithm is implemented using the ns-3 simulator [22]. The effectiveness of RED-LE is evaluated and compared against two algorithms, namely TRED and RED_E, under three different network traffic loads: low, moderate, and high.

The simulation double dumbbell topology (shown in Fig. 2) consists of N TCP connecting sources transmitting to one sink (denoted by D) via two routers R_1 and R_2. These two routers R_1 (with the algorithm implemented) and R_2 are connected together via a bottleneck link with a capacity of 10 Mbps and a propagation delay of 100 ms. Other hosts are connected to the routers via 100 Mbps links characterized by a propagation delay of 5 ms. The N number of

Algorithm 1. Detailed RED-LE's algorithm

```
Initialization:
$avg = 0$
$count = -1$
Upon every packet arrival do
Calculate the average queue size $avg$
**if** router's queue is non-empty **then**
    $avg = (1 - W_q)avg' + (W_q \times q_{cur})$
**else**
    Compute $n$ in which
    $n = f(q\_current\_time - q\_idle\_time)$
    $avg = (1 - W_q)^n \times avg'$
**end if**
**if** $avg < minTH$ **then**
    Accept the packet
    Set $count = count - 1$
**else if** $minTH \leq avg < Target$ **then**
    Set $count = count + 1$
    Based on the linear drop function compute the final
    drop probability $P_a$:
    $P_b = 2maxP(\frac{avg - minTH}{maxTH - minTH})$
    $P_a = P_b/(1 - count \times P_b)$
    Drop arriving packet according to $P_a$
    Set $count = 0$
**else if** $Target \leq avg < maxTH$ **then**
    Set $count = count + 1$
    Based on the exponential drop function compute the
    final drop probability $P_a$:
    $P_b = e^{\log(maxP)\frac{2(maxTH - avg)}{maxTH - minTH}}$
    $P_a = P_b/(1 - count \times P_b)$
    Drop arriving packet according to $P_a$
    Set $count = 0$
**else if** $maxTH \leq avg$ **then**
    Drop arriving packet
    Set $count = 0$
**end if**
**if** count= -1 **then**
    When the router's queue becomes empty
    Set $q\_idle\_time = q\_current\_time$
**end if**

**Saved variables:**
$avg$: average queue size
$q\_idle\_time$: beginning of queue idle time
$count$: packets since last dropped packet

**Preset input parameters:**
$minTH$: router's queue minimum threshold
$maxTH$: router's queue maximum threshold
$maxP$: maximum packet drop probability
$W_q$: weighting factor

**Other:**
$P_b$: current packet marking probability
$q_{cur}$: current queue size
$q\_current\_time$: current time
$f(t)$: a linear function of time $t$
```

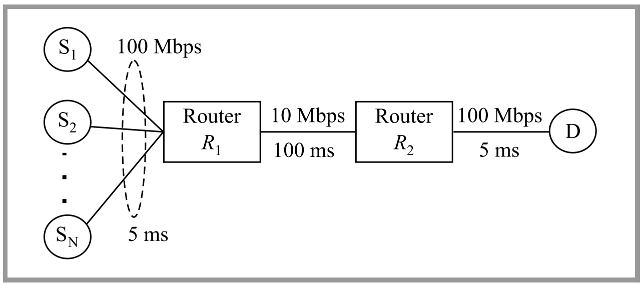


Fig. 2. Network topology.

flows was varied to indicate various levels of traffic loads in the network. The TCP implementation used is New Reno. The buffer size was set to 250 packets, while simulation time was set to 100 s. Other configurations are shown in Table 1.

Table 1
Simulation setup

Input parameter	Algorithm	Value
$minTH$	TRED, RED_E & RED-LE	30 packets
$Target$	RED-LE	60 packets
$maxTH$	TRED, RED_E & RED-LE	90 packets
$maxP$	TRED & RED-LE	0.1
W_q	TRED, RED_E & RED-LE	0.002

4.1. Scenario 1 – Low Load

In this scenario, the number of connecting sources is set to 5. As shown in Fig. 3a, the RED-LE algorithm reduces the average queue size better than both TRED and RED_E algorithms. As shown in Table 2, RED-LE reduced the queue size by 1.9437% and 14.1522% compared with TRED and RED_E, respectively.

Delay performance is shown in Fig. 3b, RED-LE outperformed both TRED and RED_E. As shown in Table 3, delays in RED-LE were by 0.0226% and 0.1140% shorter when compared with TRED and RED_E, respectively.

Figure 3c shows throughput performance. RED_E clearly outperformed both TRED and RED-LE, although the results of RED-LE were better than those of TRED. A detailed analysis is presented in Table 4.

4.2. Scenario 2 – Moderate Load

In this scenario, the number of connecting sources is set to 20. As shown in Fig. 4a, the proposed RED-LE algorithm is clearly more efficient at reducing the average queue size than both TRED and RED_E algorithms. As shown in Table 2, RED-LE reduced it by 5.2565% and 29.7475% percentage decrement when compared with TRED and RED_E, respectively.

Delay performance is shown in Fig. 4b, RED-LE satisfactorily outperformed both TRED and RED_E. As shown

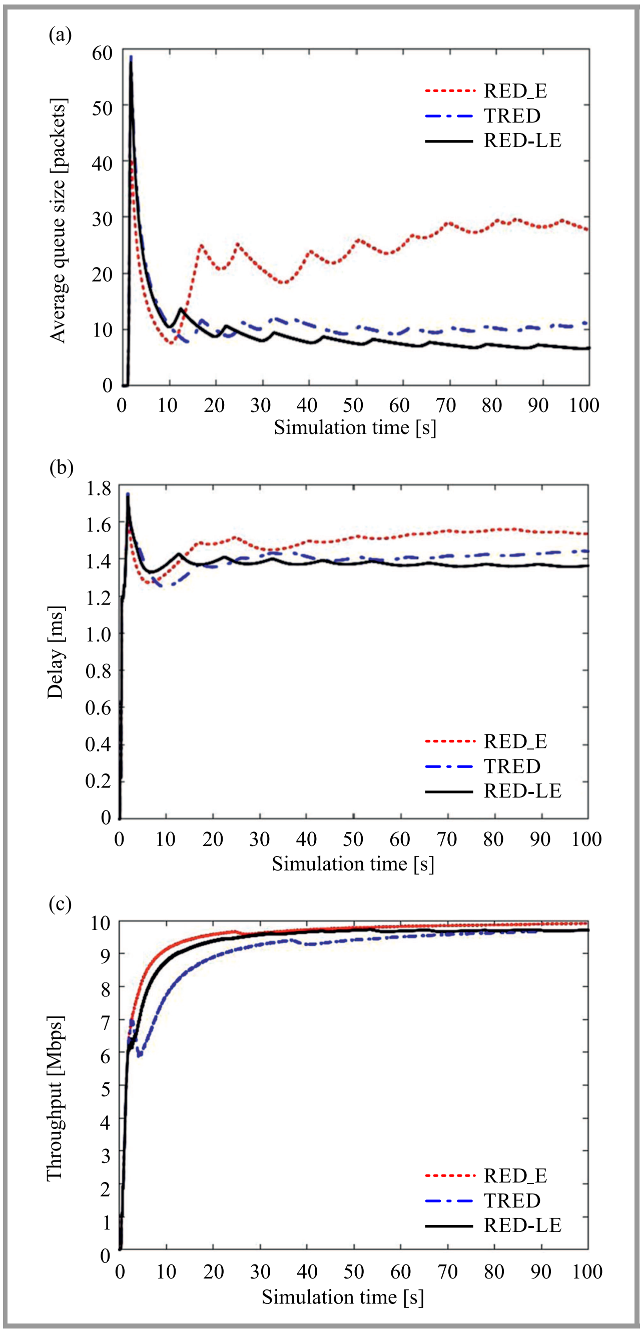


Fig. 3. Low load condition graphs: (a) average queue size, (b) delay, (c) throughput.

Table 2
Performance in terms of average queue size [packets]

Traffic load	AQM algorithm		
	TRED	RED_E	RED-LE
Low	11.1701	23.3786	9.2264
Moderate	16.3599	40.8509	11.1034
High	40.7549	62.0096	27.9556

in Table 3, RED-LE reduced the delay by 0.1777% and 0.8593% compared with TRED and RED_E, respectively.

Table 3
Performance in terms of delay [ms]

Traffic load	AQM algorithm		
	TRED	RED_E	RED-LE
Low	1.3918	1.4832	1.3692
Moderate	5.8040	6.4856	5.6263
High	15.2848	16.8606	14.3817

Figure 4c shows throughput performance. RED_E clearly obtained the highest value when compared with TRED and RED-LE. An analysis is presented in Table 4.

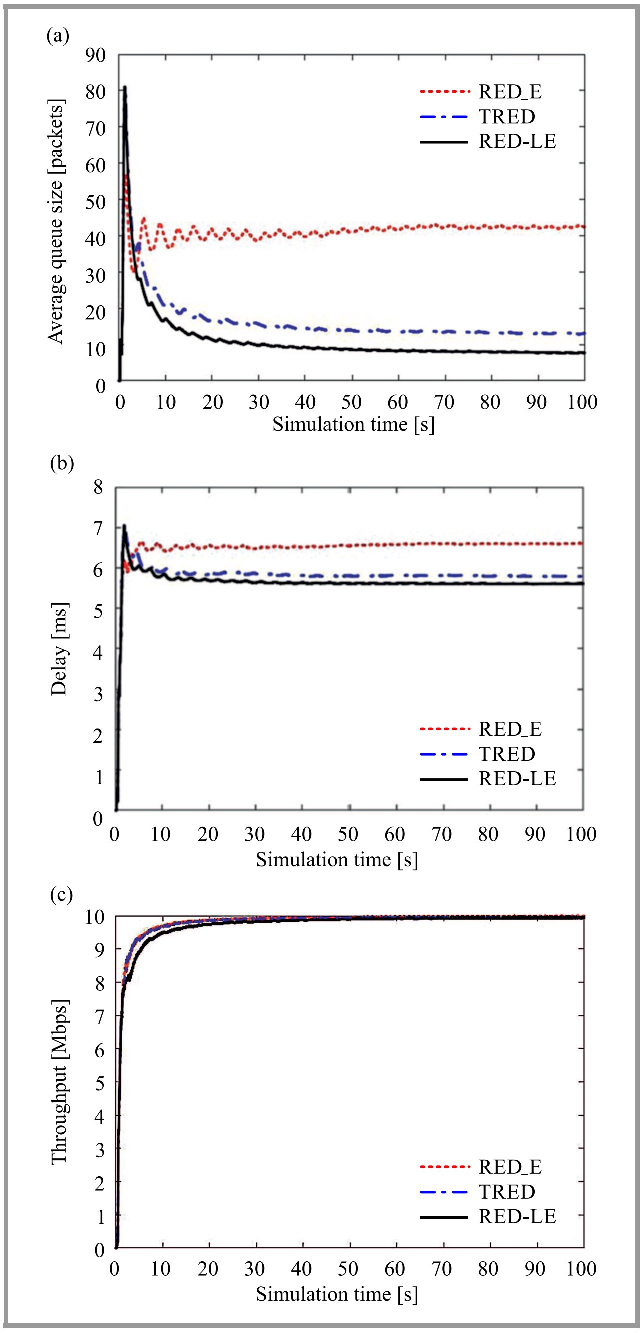


Fig. 4. Moderate load condition graphs: (a) average queue size, (b) delay, (c) throughput.

Table 4
Performance in terms of throughput [Mbps]

Traffic load	AQM algorithm		
	TRED	RED_E	RED-LE
Low	9.0105	9.4899	9.3054
Moderate	9.7709	9.7915	9.6858
High	9.7917	9.8696	9.5417

4.3. Scenario 3 – High Load

In this scenario, the number of connecting sources is set to 50. Again, as shown in Fig. 5a, the proposed RED-LE algorithm clearly outperformed both TRED and RED_E algorithms in terms of maintaining a small average queue size. As presented in Table 2, RED-LE reduced the queue size by 12.7993% and 34.0540% compared with TRED and RED_E, respectively.

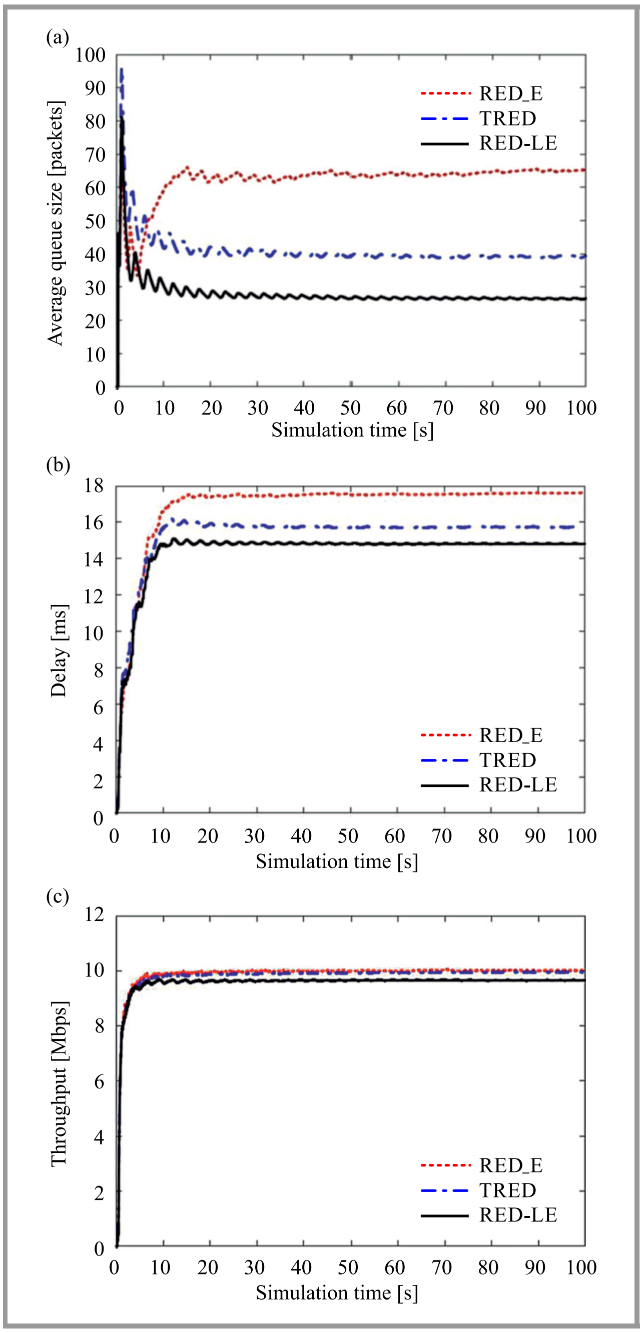


Fig. 5. High load condition graphs: (a) average queue size, (b) delay, (c) throughput.

The delay plot for the three algorithms is shown in Fig. 5b. RED-LE clearly obtained shorter delays than both TRED and RED_E. As presented in Table 3, RED-LE reduced them by 0.9031% and 2.4789% compared with TRED and RED_E, respectively.

Figure 5c shows throughput performance. RED_E obtained the highest value when compared with both TRED and RED-LE. A more detailed analysis is presented in Table 4.

5. Conclusion

In this study, a modest modification is introduced to the packet drop probability function of the RED algorithm. More specifically, a RED-linear exponential (RED-LE) algorithm was suggested for implementation in routers. A comparison was made between the RED-LE algorithm and two other AQM algorithms, namely TRED and RED_E, under various traffic load scenarios in a widely-used network simulator. From experimental results, it can be concluded that RED-LE ensures better and more efficient congestion control by obtaining reduced average queue size and delay values.

References

[1] A. A. Abu-Shareha, "Controlling delay at the router buffer using modified random early detection", *Int. J. of Comp. Netw. and Commun. (IJCNC)*, vol. 11, no. 6, pp. 63–75, 2019 (DOI: 10.5121/ijcnc.2019.11604).

[2] S. B. Danladi and F. U. Ambursa, "DyRED: An enhanced random early detection based on a new adaptive congestion control", in *Proc. of the 15th Int. Conf. on Electron., Comp. and Comput. ICECCO 2019*, Abuja, Nigeria, 2019 (DOI: 10.1109/ICECCO 48375.2019.9043276).

[3] M. Baklizi, H. Abdel-Jaber, S. Ramadass, N. Abdullah, and M. Anbar, "Performance assessment of AGRED, RED and GRED congestion control algorithms", *Inform. Technol. J.*, vol. 11, no. 2, pp. 255–261, 2012 (DOI: 10.3923/itj.2012.255.261).

[4] A. Adamu, Y. Surajo, and M. T. Jafar, "SARED: Self-adaptive active queue management scheme for improving quality of service in network systems", *J. of Comp. Sci.*, vol. 22, no. 12, pp. 253–267, 2021 (DOI: 10.7494/csci.2021.22.2.4020).

[5] N. Kaur and R. Singhai, "Congestion control scheme using network coding with local route assistance in mobile adhoc network", *Int. J. of Comp. Appl. in Technol.*, vol. 60, no. 3, pp. 242–253, 2019 (DOI: 10.1504/ijcat.2019.100298).

[6] L. Pei, F. Wu, and S. Wang, "Periodic, quasi-periodic and chaotic oscillations in two heterogeneous AIMD/RED network congestion models with state-dependent round-trip delays", *Int. J. of Bifurcation and Chaos*, vol. 31, no. 6, 2150124, 2021 (DOI: 10.1142/S0218127421501248).

[7] H. Mohammed, G. Attiya, and S. El-Dolil, "Active queue management for congestion control: Performance evaluation, new approach, and comparative study", *Int. J. of Comput. and Netw. Technol.*, vol. 5, no. 2, pp. 37–49, 2017 (DOI: 10.12785/IJCNT/050201).

[8] A. Ahmed and N. Nasrelden, "New congestion control algorithm to improve computer networks performance", in *Proc. of the 28th Int. Conf. on Innovat. Trends in Comp. Engin. ITCE 2018*, Aswan, Egypt, 2018, pp. 87–93 (DOI: 10.1109/ITCE.2018.8316605).

[9] H. Abdel-Jaber, A. Shehab, M. Barakat, and M. Rashad, "IGRED: An improved gentle random early detection method for management of congested networks", *J. of Intercon. Netw.*, vol. 19, no. 2, 1950004, 2019 (DOI: 10.1142/S021926591950004X).

[10] J. Aweya, M. Ouellette, and D. Y. Montuno, "A control theoretic approach to active queue management", *Comp. Netw.*, vol. 36, no. 2, pp. 203–235, 2001 (DOI: 10.1016/S1389-1286(00)00206-1).

[11] S. Floyd and V. Jacobson, "Random early gateway for congestion avoidance", *IEEE/ACM Trans. on Network.*, vol. 1, no. 4, pp. 397–413, 1993 (DOI: 10.1109/90.251892).

[12] A. Giménez, M. A. Murcia, J. M. Amigó, O. Martínez-Bonastre, and J. Valero, "New RE-type TCP-AQM algorithms based on beta distribution drop functions" [Online]. Available: https://arxiv.org/pdf/2201.01105.pdf

[13] S. Floyd, "Recommendation on using the gentle – variant of RED", 2000 [Online]. Available: http://www.icir.org/floyd/red/gentle.html

[14] H. Abdel-Jaber, "An exponential active queue management method based on random early detection", *J. of Comp. Netw. and Commun.*, vol. 2020, article ID 80904682020, 2020 (DOI: 10.1155/2020/8090468).

[15] Y. Zhang, J. Mab, Y. Wang, and C. Xu, "MRED: An improved nonlinear RED algorithm", in *Int. Proc. of Comp. Science and Inform. Technol.*, vol. 44, no. 2, pp. 6–11, 2012 (DOI: 10.7763/IPCSIT.2012.V44.2).

[16] A. K. Paul, H. Kawakami, A. Tachibana, and T. Hasegawa, "An AQM based congestion control for ENB RLC in 4G/LTE network", in *Proc. of the IEEE Canadian Conf. on Elec. and Comp. Engin. CCECE 2016*, Vancouver, BC, Canada, 2016 (DOI: 10.21109/CCECE.2016.7726792).

[17] C. Suwannapong and C. Khunboa, "Congestion control in CoAP observe group communication", *Sensors*, vol. 19, no. 3433, pp. 1–14, 2019 (DOI: 10.3390/s19153433).

[18] C.-W. Feng, L.-F. Huang, C. Xu, and Y.-C. Chang, "Congestion control scheme performance analysis based on nonlinear RED", *IEEE Systems J.*, vol. 11, no. 4, pp. 2247–2254, 2017 (DOI: 10.1109/JSYST.2014.2375314).

[19] K. Zhou, K. L. Yeung, and V. O. K. Li, "Nonlinear RED: A simple yet efficient active queue management scheme", *Comp. Netw.*, vol. 50, pp. 3784–3794, 2006 (DOI: 10.1016/j.comnet.2006.04.007).

[20] D. Kumhar, A. Kumar, and A. Kewat, "QRED: An enhancement approach for congestion control in network communications", *Int. J. of Inform. Technol.*, vol. 13, pp. 221–227, 2021 (DOI: 10.1007/s41870-020-00538-1).

[21] A. Adamu, V. Shorgin, S. Melnikov, and Y. Gaidamaka, "Flexible random early detection algorithm for queue management in routers", in *Distributed Computer and Communication Networks. 23rd International Conference, DCCN 2020, Moscow, Russia, September 14-18, 2020, Revised Selected Papers*, V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev, Eds. *LNCS*, vol. 12563, pp. 196–208. Springer, 2020 (DOI: 10.1007/978-3-030-66471-8_16).

[22] "The Network Simulator ns-3" [Online]. Available: http://www.nsnam.org

Samuel O. Hassan received his M.Sc. and Ph.D. degrees in Computer Science from Obafemi Awolowo University, Ile-Ife, Nigeria. Currently, he is a Lecturer at the Department of Mathematical Sciences (Computer Science Unit), Olabisi Onabanjo University, Ago-Iwoye, Nigeria. He is a Certified Information Technology Practitioner (*C.itp*). His research interests spans computational mathematics, computer networks and communications, mathematical modeling and simulation, and Internet congestion control.

E-mail: samuel.hassan@oouagoiwoye.edu.ng
Department of Mathematical Sciences
Olabisi Onabanjo University
Ago-Iwoye, Nigeria

Paper

Electrically Small Microstrip Antenna Based on Magnetodielectric Materials

Ararat Stepanyan[1], Hovhannes Haroyan[2], and Arsen Hakhoumian[2,3]

[1] *Institute of Information and Telecommunication Technologies and Electronics,*
National Polytechnic University of Armenia, Yerevan, Armenia
[2] *Chair of Telecommunication and Signal Processing, Yerevan State University, Yerevan, Armenia*
[3] *Institute of Radiophysics and Electronics, Ashtarak, Armenia*

https://doi.org/10.26636/jtit.2022.158821

Abstract—An electrically small microstrip patch antenna based on high permittivity dielectric and magnetodielectric materials (MDM) is investigated in this paper. The basic parameters of microstrip patch antennas based on high dielectric and magnetodielectric materials are compared with other solutions. The analysis shows that an MDM-based patch surface is 7.14 times smaller when compared with a suspended plate antenna. The use of MDM improves bandwidth and offers perfect impedance matching between the material and free space, over a much wider bandwidth.

Keywords*—*electrically small antenna, magnetodielectric material, microstrip antenna.

1. Introduction

Electrically small antennas (ESAs) have become popular these days due to their reduced size and ability of being integrated on a chip. Such antennas are mainly used in wireless, mobile communications, for instance in detection and radio identification applications, as well as in medical instruments and video equipment [1]. The key advantage of ESAs is that the dimensions of the antenna are much smaller than their operating wavelength. As it is the case for any aerial (and the same applies to ESAs), efficient radiation occurs when the operating frequency matches the antenna's resonance.

Because ESAs are small in general, their bandwidth is limited [2]–[4]. Their radiation resistance, meanwhile, is usually much smaller than the ohmic loss on the radiating elements. Thus, their radiation efficiency is suppressed. Traditional on-chip ESAs often rely on spiral-shaped metallic structures, a solution which usually makes the bandwidth and the loss issues even worse. To improve the performance of ESAs, many efforts have been made to identify better approaches to their design and fabrication.

Currently, high dielectric constant and low-loss materials are usually used as substrates in the fabrication of miniaturized antennas. Although a high degree of miniaturization may be achieved using high permittivity dielectric materials, some problems are encountered here:

- existence of a highly confined field around the substrate results in narrow band and lower efficiency of the RF components, such as the antenna,
- impedance in a high permittivity medium is low, which makes it difficult to design proper impedance matching between the source and the antenna.

Magnetodielectric materials (MDM) offer a promising way forward in terms of miniaturization of the antenna and designing ESAs that are capable of overcoming the above mentioned issues [5]–[7].

2. Problem Statement

Microstrip patch antennas (MPA) are popular, low-cost, low-profile aerial structures used when the specific application requires a broadside radiation pattern with a high

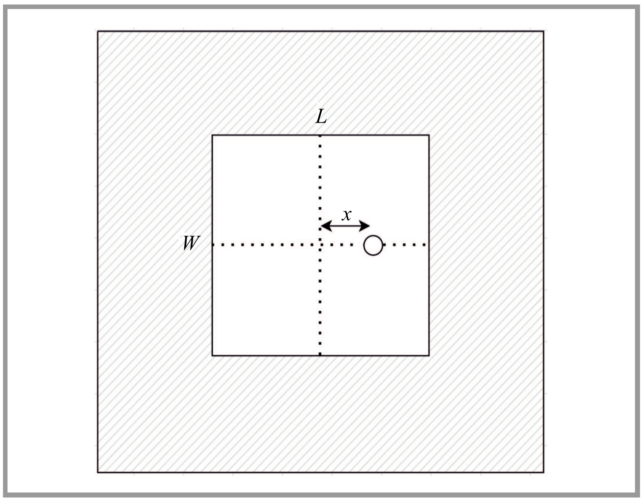


Fig. 1. Schematic view of a coax-fed microstrip patch antenna.

front-to-back ratio. In this paper, miniaturization of an MDM-based patch antenna is presented and a comparative analysis with its high dielectric permittivity counterparts is performed. Figure 1 shows the patch fed below from a coax along the resonant length, where L and W represent patch dimensions, while x denoting the point of coaxial feed [8].

In general, wavelength in the patch substrate material and impedance are interrelated, as shown in Eqs. (1) and (2) [9], where:

$$\lambda = \frac{\lambda_0}{\sqrt{\varepsilon_r \mu_r}} \ , \tag{1}$$

$$Z = Z_0 \sqrt{\frac{\mu_r}{\varepsilon_r}} \ . \tag{2}$$

The MPA radiates from the currents induced on the patch. Equivalently, the magnetic currents around the periphery of the patch and from surface waves induced in the dielectric slab. The surface waves radiate when they reach the edges of the substrate and their emission contributes to the normal patch radiation. The fringing fields from the patch to the ground plane readily excite the lowest-order surface-wave TM_0 mode that has no low frequency cutoff. A dielectric slab of any thickness supports this mode. Surface-wave radiation is controlled by limiting the substrate area or by adding etched photonic bandgap patterns to the open areas of the substrate. However, generally, surface waves are undesirable. As the substrate thickness or dielectric constant increase, the power ratio of surface waves increases as well. The antenna's impedance bandwidth includes directly radiated power and surface-wave power [7].

The antenna quality factor Q_T is a combination of the space-wave radiation Q_R, surface-wave radiation Q_{SW}, as well as dielectric $Q_d = \frac{1}{\text{tg}\delta}$, magnetic Q_m, and conductor $Q_c = h\sqrt{\pi f \mu_0 \sigma}$ factors, such as:

$$\frac{1}{Q_T} = \frac{1}{Q_R} + \frac{1}{Q_{SW}} + \frac{1}{Q_d} + \frac{1}{Q_c} + \frac{1}{Q_m} \ , \tag{3}$$

where tgδ is the dielectric loss tangent, σ is the patch conductivity.

The antenna's bandwidth is related to the quality factor:

$$BW = \frac{1}{\sqrt{2 Q_T}} \ . \tag{4}$$

The MPA bandwidth can be calculated depending on substrate parameters (h height, ε_r dielectric constant and μ_r magnetic constant) and λ_0 free space wavelength, according to [9]:

$$BW = \frac{96\sqrt{\frac{\mu h}{\sqrt{\varepsilon \lambda_0}}}}{\sqrt{2}(4 + 17\sqrt{\varepsilon \mu})} \ . \tag{5}$$

MDM allows to miniaturize the antenna by the same factor as a high permittivity dielectric material, however using moderate values of ε_r and μ_r. As seen from Eqs. (1), (2) and (5), by adding a magnetic material into dielectric substrate and making the values of ε_r and μ_r nearly equal, an improvement in substrate properties may be expected. The effect of adding magnetic material into dielectric material reduces strong field confinement and the medium becomes far less capacitive, but only in terms of its dielectric features. Thus, the use of MDM allows to miniaturize the antenna, improve its bandwidth and achieve perfect impedance matching between the material and the free space, over a much wider frequency range. The MPA size may be calculated using Eqs. (6) and (7) [8], [9]:

Antenna width:

$$W = \frac{c}{2f\sqrt{\varepsilon_r \mu_r}} \ . \tag{6}$$

Antenna length:

$$L = \frac{c}{2f\sqrt{\varepsilon_{eff} \mu_{eff}}} - 2\Delta L \ , \tag{7}$$

where

$$L = 0.412h \frac{\varepsilon_{eff}\mu_{eff} + 0.3}{\varepsilon_{eff}\mu_{eff} - 0.258} \cdot \frac{\frac{w}{h} + 0.262}{\frac{w}{h} + 0.813} \ ,$$

h is substrate height, $\varepsilon_e f f$ is effective dielectric constant.

Feed point:

$$x = \frac{L}{\pi} \sin^{-1} \sqrt{\frac{R_i}{R_e}} \ , \tag{8}$$

where R_i is input resistance, $R_e = \frac{\eta \lambda_0}{\pi W \left[1 - \frac{(kh)^2}{24}\right]}$ resistance at the edge, η is radiating efficiency.

Substrate sizes:

$$W_s = W + 12h \ , \quad L_s = L + 12h \ . \tag{9}$$

3. Results and Discussions

Microstrip antennas are researched under free space (suspended plate antenna) conditions, with a high permittivity dielectric and MDM set at 2.4 GHz.

Here, a full-wave numerical analysis based on FEM/MoM is conducted to investigate MPAs with different substrate materials in order to minimize antenna size. The considered aerials are designed in the FEKO environment.

Suspended plate antenna sizes ($\varepsilon_r = \mu_r = 1$) are calculated using Eqs. (6)–(9), when the distance between the patch and the ground (substrate height) h is 1.6 mm:

- patch size $w = 61$ mm, $L = 58.4$ mm,
- feed position from the center $x = 9$ mm,
- ground plane size $w_g = 89$ mm, $L_g = 91.5$ mm.

VSWR of the antenna in the 2.38 … 2.42 GHz band is lower than 2 (VSWR < 2), $BW = 40$ MHz, relative bandwidth is $\delta = 1.66\%$, and at 2.4 GHz VSWR $= 1.06$ (Fig. 2), gain in 2:1 VSWR band varies within the 9…9.17 dBi range (Fig. 3) and beamwidth is in the θ plane at 2.4 GHz $2\theta_{0.5} = 62.19°$ as in φ plane $2\varphi_{0.5} = 68.64°$ (Fig. 4).

Next, the Rogers RO3210 dielectric was used as a high permittivity substrate, with its parameters equaling: $\varepsilon_{r1} = 10.2$, $h = 1.6$ mm, $\mathrm{tg}\delta = 0.0027$. The antenna's size and feed point are calculated based on Eqs. (6)–(9):

- patch size $w = 25$ mm, $L = 18.3$ mm,
- feed position from the center $x = 3$ mm,
- substrate size $w_s = 47.8$ mm; $L_s = 41.1$ mm.

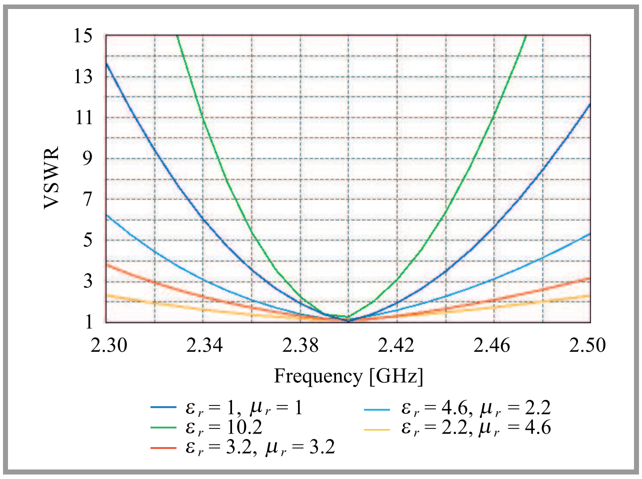


Fig. 2. VSWR of the antenna for different substrates.

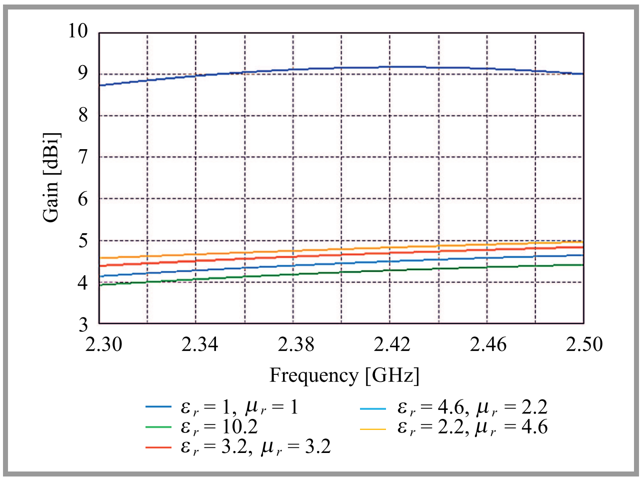


Fig. 3. Gain of the antennas for different substrates.

Using RO3210, the antenna's 2:1 VSWR frequency band is $2.383\ldots2.41$ GHz, $BW = 27$ MHz, relative bandwidth is $\delta = 1.12\%$. At the central 2.4 GHz frequency, VSWR $= 1.02$ (Fig. 2), gain in 2:1 VSWR band varies within the $4.25\ldots4.88$ dBi range (Fig. 3), and beamwidth at 2.4 GHz is, in θ plane, $2\theta_{0.5} = 102.9°$, in φ plane $2\varphi_{0.5} = 99.97°$ (Fig. 4).

For comparison with high dielectric MPA parameters, the characteristics of the MDM material are chosen, thus $\mu_r\varepsilon_r$ is equal to the permittivity of the selected high dielectric material $\mu_r\varepsilon_r = 1\cdot\varepsilon_{r1}$.

At first $\mu_r = \varepsilon_r = 3.2$ are chosen as the applicable parameters, the height and loss tangent are the same: $h = 1.6$ mm, $\mathrm{tg}\delta = 0.0027$. Using formulas (6)–(9), the antenna sizes and feed point are calculated as:

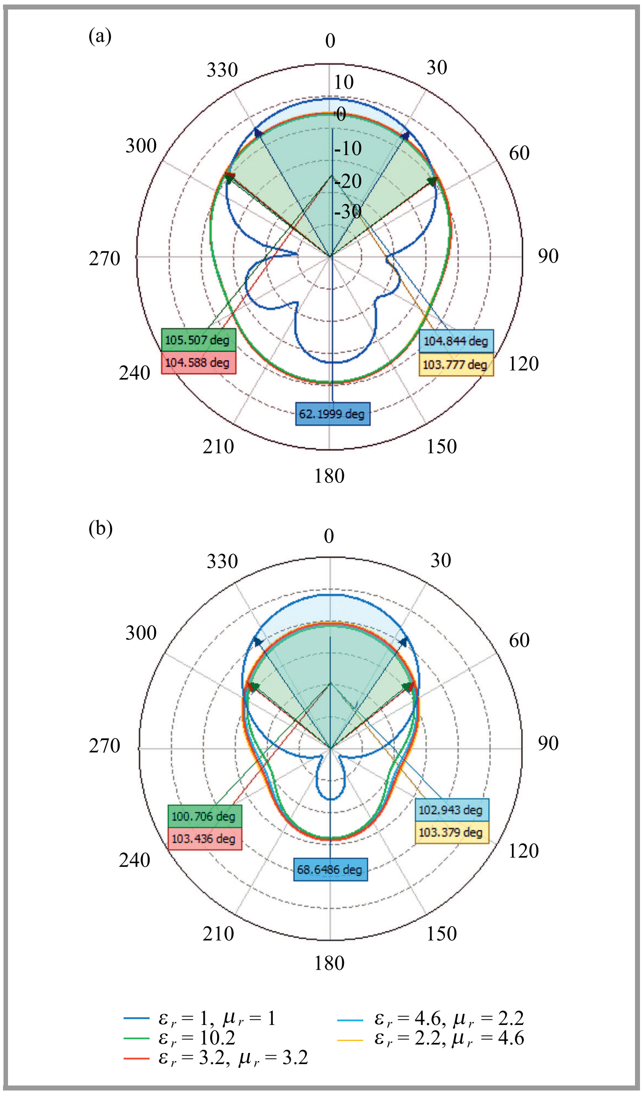


Fig. 4. Antenna radiation pattern at 2.4 GHz in (a) θ and (b) φ plane for different substrates.

- patch size $w = 26.3$ mm; $L = 19$ mm,
- feed position from the center $x = 5$ mm,
- substrate size $w_s = 47.8$ mm; $L_s = 41.1$ mm.

The MDM based antenna's 2:1 VSWR frequency band is $2.349\ldots2.455$ GHz, hence its absolute bandwidth is $BW = 106$ MHz, relative bandwidth is $\delta = 4.41\%$ and at central 2.4 GHz frequency VSWR $= 1.1$. The antenna's gain in the 2:1 VSWR band ($2.349\ldots2.455$ GHz) varies within the $4.9\ldots5.2$ dBi range (Fig. 2), beam width in θ plane at 2.4 GHz $2\theta_{0.5} = 101.9°$, in φ plane $2\varphi_{0.5} = 101.84°$.

In the second stage, the $\mu_r > \varepsilon_r$ case is considered with $\mu_r = 4.6$ and $\varepsilon_r = 2.2$ as substrate parameters. Substrate height is the same at $h = 1.6$ mm. Antenna sizes are identical as in the case with $\mu_r = \varepsilon_r = 3.2$, feed position from the center is $x = 6.5$ mm. Under such conditions, the antenna's 2:1 VSWR frequency band is $2.316\ldots2.479$ GHz (absolute bandwidth is $BW = 153.1$ MHz, while relative bandwidth is $\delta = 6.37\%$). At the central 2.4 GHz frequency VSWR $= 1.1$ (Fig. 2). The antenna's gain in the 2:1 VSWR band

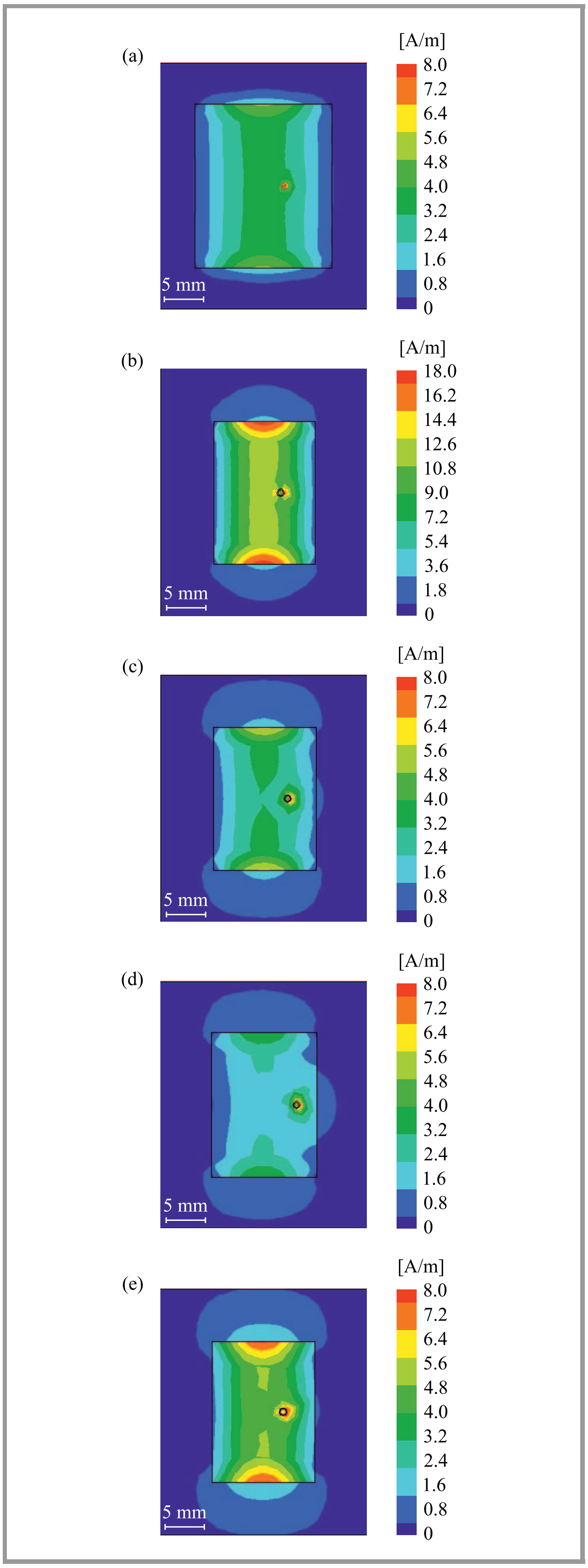


Fig. 5. Antenna surface current distribution for different substrates: (a) $\varepsilon_r = \mu_r = 1$, (b) $\varepsilon_r = 10.2$, $\mu_r = 1$, (c) $\varepsilon_r = \mu_r = 3.2$, (d) $\varepsilon_r = 2.2$, $\mu_r = 4.6$, (e) $\varepsilon_r = 4.6$, $\mu_r = 2.2$.

(2.316…2.479 GHz) varies within the 4.9…5.3 dBi range, beamwidth: in θ plane at 2.4 GHz $2\theta_{0.5} = 101.4°$, in φ plane $2\varphi_{0.5} = 101.7°$ (Fig. 4).

In the third stage, the $\mu_r < \varepsilon_r$ case is studied with $\mu_r = 2.24$ and $\varepsilon_r = 4.6$ as substrate parameters, and the same height $h = 1.6$ mm is considered. Antenna sizes are the identical as those in the case of $\mu_r = \varepsilon_r = 3.2$, feed position from the center $x = 3.8$ mm. The antenna's 2:1 VSWR frequency band is 2.362…2.432 GHz, while absolute bandwidth is $BW = 70$ MHz and relative bandwidth is $\delta = 2.91\%$. At the frequency of 2.4 GHz, VSWR = 1.1 (Fig. 2). The gain in the 2:1 VSWR band (2.329…2.474 GHz) varies within the 4.95…5.1 dBi range (Fig. 3), beamwidth at 2.4 GHz: in θ plane $2\theta_{0.5}$ is 102.3°, in φ plane $2\varphi_{0.5}$ is 101.9° (Fig. 4).

Surface current distribution is shown in Fig. 5. It may be observed that an increase in permittivity creates a highly confined field around the substrate. The increase of permeability results in a decrease in surface current. The existence of a highly confined field around the substrate results in a narrow band.

4. Conclusion

By performing numerical calculations, we have compared conventional microstrip patch antennas based on a high dielectric permittivity substrate with a new aerial based on a magnetodielectric material.

MDM allows to miniaturize the antenna's size by the same factor as high permittivity dielectric material. By using moderate values of ε_r and μ_r and by comparing the results with a suspended plate antenna, we have concluded that the patch surface may be thus reduced 7.14 times.

The antenna's quality factor varies depending on substrate permittivity and permeability. With high permittivity dielectric materials, the existence of a highly confined field around the substrate results in the narrowing of the band. In MDM-based antennas, strong field confinement is reduced, and the medium one becomes far less capacitive, but only in the dielectric part. Thus, the use of MDM results in the miniaturization of the antenna, improvement of its bandwidth and allows to achieve perfect impedance matching between the material and free space over a much wider bandwidth.

Consequently, an electrically small antenna with a wider bandwidth is achieved using MDM in the case of $\mu_r > \varepsilon_r$.

Acknowledgments

This work is supported by a scientific research grant through the Science Committee of MESCS of Armenia (20APP-1C009).

References

[1] G. H. R. Stuart and A. Pilwerbetsky, "Electrically small antenna elements using negative permittivity resonator", *IEEE Trans. Anten. Propag.*, vol. 54, no. 6, pp. 1644–1653, 2006
(DOI: 10.1109/TAP.2006.875498).

[2] H. A. Wheeler, "Fundamental limitations of small antennas", *Proc. of the IRE*, vol. 35, no. 12, pp. 1479–1484, 1947 (DOI: 10.1109/JRPROC.1947.226199).

[3] L. Peng, P. Chen, and A. Wu, "Efficient radiation by electrically small antennas made of coupled split-ring resonators", *Scientific Rep.*, vol. 6, 33501, 2016 (DOI: 10.1038/srep33501).

[4] S. A. Schelkunoff and H. T. Friis, "Antennas and theory", in *Antennas Theory and Practice*. Wiley, 1952 (ISBN: 9780471759003).

[5] E. Andreou, T. Zervos, A. A. Alexandridis, and G. Fikioris, "Magnetodielectric materials in antenna design: Exploring the potentials for reconfigurability", *IEEE Anten. and Propag. Mag.*, vol. 61, no. 1, 2019 (DOI: 10.1109/MAP.2018.2883029).

[6] P. Ikonen, K. Rozanov, A. Osipov, P. Alitalo, and S. Tretyakov, "Magnetodielectric substrates in antenna miniaturization: potential and limitation", *IEEE Trans. Anten. Propag.*, vol. 54, no. 11, pp. 3391–3398, 2006 (DOI: 10.1109/TAP.2006.884303).

[7] R. C. Hansen and M. Burke, "Antennas with magneto-dielectrics", *Microw. and Opt. Tech. Lett.*, vol. 26, no. 2, pp. 75–78, 2000 (DOI: 10.1002/1098-2760(20000720)26:2<75::AID-MOP3>3.0.CO;2-W).

[8] T. A. Milligan, *Modern Antenna Design*. Wiley, 2005 (ISBN: 9780471457763).

[9] L. Huitema (Ed.), *Progress in Compact Antennas*. London, UK: IntechOpen Limited, 2014 (ISBN: 9789535117230).

Ararat Stepanyan graduated from the National Polytechnic University of Armenia (NPUA), Yerevan, Armenia, in 2020. Currently, he is a master's degree student at the NPUA. His current research interests focus on antenna engineering, microwave electronics, electronic warfare and twisted-wave radio technology.

E-mail: araratatepanyan9@gmail.com
Institute of Information and Telecommunication
Technologies and Electronics
National Polytechnic University of Armenia
Yerevan, Armenia

Hovhannes Haroyan is an Associate Professor at Yerevan State University, Chair of Telecommunication and Signal Processing. He is an author of more than 30 scientific articles. His research interests include antenna engineering, wireless communications, spectral efficiency improvement in communications systems, plasmonics and microresonators.

https://orcid.org/0000-0002-2089-8845
E-mail: hharoyan@ysu.am
Chair of Telecommunication and Signal Processing
Yerevan State University
Yerevan, Armenia

Arsen Hakhoumian received his Ph.D. degree in Quantum Radiophysics in 1983. He has served as the director of the Institute of Radiophysics and Electronics since 2006. In 2014, he received a D.Sc. degree. He is a Corresponding Member of the National Academy of Sciences of the Republic of Armenia (NAS RA). His research interests include antenna engineering, low-noise receivers, terahertz physics and technology, microwave radar and telecommunication systems, high temperature superconductivity and microwave radio links.

E-mail: a.hakhoumian@ysu.am
Chair of Telecommunication and Signal Processing
Yerevan State University
Yerevan, Armenia

Institute of Radiophysics and Electronics
Ashtarak, Armenia

Paper

Speeding Up Minimum Distance Randomness Tests

Krzysztof Mańk

Military University of Technology, Warsaw, Poland

https://doi.org/10.26636/jtit.2022.159222

Abstract—Randomness testing is one of the essential and easiest tools for the evaluation of the features and quality of cryptographic primitives. The faster we can test, the greater volumes of data can be checked and evaluated and, hence, more detailed analyses may be conducted. This paper presents a method that significantly reduces the number of distances calculated in the minimum distance, Bickel-Breiman, and *m* nearest points tests. By introducing a probabilistic approach with an arbitrarily low probability of failure, the number of calculated distances proportional to the number of required distances and independent of the number of points was achieved. In the well-known Diehard's minimum distance and 3D spheres tests, the quantity of computations achieved is reduced by the factors of 394 and 771, respectively.

Keywords—Bickel-Breiman test, minimum distance test, m nearest pairs test, randomness test.

1. Introduction

Testing of randomness is important, as a variety of analyses concerned with cryptographic primitives may be reduced to examining appropriately crafted binary sequences. In each such case, the quality of the analysis increases with the volume of data checked, but this leads to increases in time and cost as well. Any improvements in randomness testing translate directly into the quality of evaluation of RNGs, symmetric ciphers, hash functions, and other cryptographic primitives.

First proposed in Diehard as minimum distance and 3D spheres tests, then found in TestU01 as *m* nearest pairs test and, of course, ported to Dieharder as diehard_2dsphere and diehard_3dsphere a family of minimum distance tests raised and took its place in randomness testing. Computationally, the main part of these tests consists in calculating distances between *n* points randomly placed in a *t* dimensional hypertorus. Naive implementation leads to the calculation of $n(n-1)/2$ distances, which, in some cases, may be unacceptable.

Dieharder and NIST's Statistical Test Suite, are the two most commonly used test suites. Both are slightly outdated, but still remain popular due to their availability.

We will distinguish three test cases:

1. minimum distance,
2. *m* nearest pairs,
3. Bickel-Breiman.

The first one is a special case of the second case, but because it relies on a different method to speed up the calculation process, we make a distinction between the two.

The most general formulation is the closest-pair problem, for which deterministic algorithms operating in $O(n \log n)$ are known (described in [1]–[4]), as well as probabilistic algorithms with linear complexity due to [5]–[6].

The best-known deterministic method for performing the minimum distance test is based on Fischler's upper estimation for minimum distance. The algorithm may be found in [7]–[9]. As shown in the following section, this method is characterized by linear complexity.

The contribution of this research consists in introducing a probabilistic approach to minimum distance estimation with an arbitrarily low probability of failure. Even for extremely low probabilities of failure, it is possible to obtain significantly lower estimations of minimum distance, leading to the number of calculated distances being independent of the number of points.

Similar approaches for the *m* nearest pairs and the Bickel-Breiman tests are proposed in the subsequent sections. In all three tests, we obtained the number of calculated distances that was proportional to the number of the needed distances and was independent of the number of points or the number of dimensions.

In the remaining part of this paper, for the sake of clarity, we shall simply refer to a cube and a torus, instead of a *t* dimensional hypercube and a hypertorus.

2. Minimum Distance Test

In this paragraph, we will consider an algorithm deployed to identify two nearest points between *n* points that are randomly placed in a *t* dimensional hypertorus. In 2002, Fischler published a paper [7] on correcting and speeding up minimum distance tests, like the two mentioned from Diehard. Fischler's method is based on dividing the torus into equal cubes and on calculating distance only between points in the same or adjoining cubes. The smaller the cubes, the fewer distances have to be calculated. But below a certain point, no points in the same or in adjoining cubes may be identified, and an error occurs. Fischler gave the upper estimate for the minimum distance in the set of *n* points in a *t* dimensional unit hypertorus:

$$D \leqslant \frac{\sqrt{t}}{\sqrt[t]{n}} ,$$

it is equal to twice the length of the edge of the sub-cube. As it will be shown below, this is unnecessary over the secure approach.

To determine the expected number of calculated distances, let us consider to-the-right-and-down approach illustrated in Fig. 1 for the two-dimensional case.

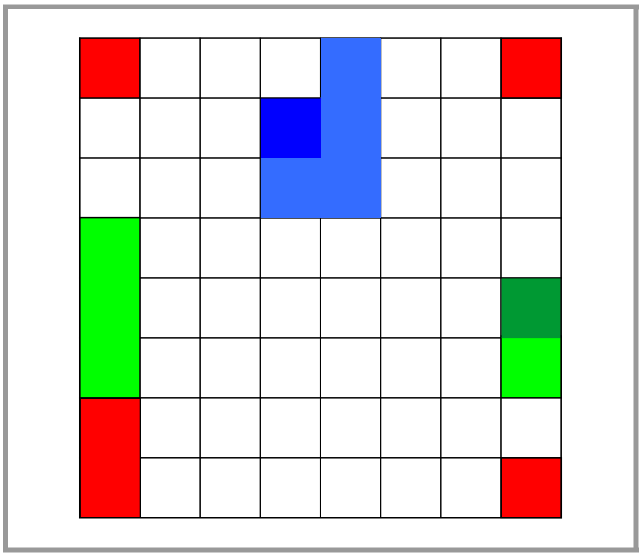


Fig. 1. To-the-right-and-down approach for calculating distances in a 2D torus.

For every cube (fully colored), we need to calculate the distance for every pair of points in that cube and from every point to all points in cubes to the right and down (highlighted by corresponding light color).

For $t > 2$ dimensional case we can for example test all cubes in volume consisting of a series of 3-by-…,-by-3 $t-1$, $t-2\ldots$, 3 cubes, "line" 3-by-1 and single cube. Note that the last two can be treated as one- and zero-dimensional cubes-of-cubes. The total number of cubes in this volume is $(3^t-1)/2$. From Fischler's formula, we can calculate the number of cubes as:

$$\left\lfloor \frac{2\sqrt[t]{n}}{\sqrt{t}} \right\rfloor^t \leqslant \frac{2^t n}{t^{\frac{t}{2}}}.$$

Note that only for t up to 4 expected number of points in every cube will not exceed 1.

Using binomial distribution, the expected number of calculated distances per cube is:

$$E_F = \sum_{i=2}^{n} \binom{i}{2} \binom{n}{i} p^i (1-p)^{n-i}$$
$$+ \frac{3^t-1}{2} \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \sum_{j=1}^{n-i} j \binom{n-j}{j} p^j (1-p)^{n-i-j}$$
$$= \frac{1}{2} n(n-1) p^2 \left(3^t (1-p) + p\right) ,$$

where $p = \left\lfloor \frac{2\sqrt[t]{n}}{\sqrt{t}} \right\rfloor^{-t}$.

If we omit the floor function in p for a large n, we get the approximation:

$$E_F \approx \frac{1}{2} \left(\frac{3t}{4}\right)^t ,$$

and the total number of calculated distances:

$$\frac{1}{2} \left(\frac{3}{2}\right)^t t^{\frac{t}{2}} n.$$

In Diehard's minimum distance test, we have $n = 8000$ and $t = 2$, which gives us the maximum number of 126^2 squares and the expected number of calculated distances slightly exceeds 18,138. For the 3D spheres test, we have $n = 4000$ and $t = 3$, which gives the maximum number of 18^3 cubes and the expected number of calculated distances 37021. For $n = 3000$ and $t = 5$ one would get 4^5 5D cubes and expected number of calculated distances equaling 1,066,477. Those numbers mean, respectively, 1,764, 216, and 4.2 times fewer computations compared to the basic naive algorithm.

From Fig. 1, we can observe that the distance is calculated if and only if two points collide in the same double-cube cube composited from $(3^t-1)/2$ base cubes. To find two nearest points, all we need is at least one such collision, which leads us to a different approach of finding the maximum number of dividing cubes.

We start from the probability of no collisions when placing n points in K boxes being small enough:

$$\frac{K!}{(K-n)!K^n} \leqslant \varepsilon ,$$

where ε can be, for example, 10^{-20}.

For efficiently solve this inequality for K, the well-known relation is used:

$$\mathrm{e}^{-x} \geqslant 1-x ,$$

which leads to:

$$\frac{K!}{(K-n)!K^n} = \prod_{i=1}^{n} \frac{K-i+1}{K} = \prod_{i=1}^{n} \left(1 - \frac{i-1}{K}\right)$$
$$\leqslant \prod_{i=1}^{n} \mathrm{e}^{-\frac{i-1}{K}} = \mathrm{e}^{-\frac{1}{K}\sum_{i=1}^{n}(i-1)} = \mathrm{e}^{-\frac{n(n-1)}{2K}} \leqslant \varepsilon ,$$

and the upper limit for K is:

$$K \leqslant -\frac{n(n-1)}{2\ln\varepsilon} .$$

However, the number of boxes K is not the number of needed cubes, which in fact is:

$$\left\lfloor -\frac{n(n-1)}{2\ln\varepsilon} \right\rfloor 3^t .$$

In the following considerations, the following formula will be used:

$$\left\lfloor \sqrt[t]{-\frac{n(n-1)}{2\ln\varepsilon}} \right\rfloor^t 3^t .$$

Similarly, we can calculate the expected number of calculated distances per cube:

$$E_C = \frac{1}{2} n(n-1) p^2 \left[3^t (1-p) + p\right] ,$$

where $p = \left\lfloor \sqrt[t]{-\frac{n(n-1)}{2\ln\varepsilon}} \right\rfloor^{-t} 3^{-t}$.

Assuming n is large after omitting the floor function, we get:

$$E_C \approx \frac{2\ln^2\varepsilon}{n(n-1)3^t} ,$$

and the total number of calculated distances is:

$$-\ln\varepsilon ,$$

which is independent of the number of points.

According to proposed method for the three examples, we have:

- 2499^2 squares and 46 distances,
- 165^3 cubes and 48 distances,
- 27^5 5D cubes and 76 distances.

Those numbers mean that the number of computations is 394, 771, and 14,033 times smaller compared to Fischler's method, respectively.

If, instead of inequality for e^{-x}, Stirling's approximation for factorial is used, a slightly better formula may be obtained:

$$\left(\frac{K}{K-n}\right)^{K-n+\frac{1}{2}} e^{-n} \leqslant \varepsilon ,$$

which unfortunately cannot be efficiently solved for K and, therefore, was omitted. However, there is not so little flaw in this method.

According to the approach shown in Fig. 1, for any point put into a torus, the distances would be calculated only to the points in the same or neighboring sub-cubes. Unfortunately, in some cases this could lead to a failure in finding the nearest point.

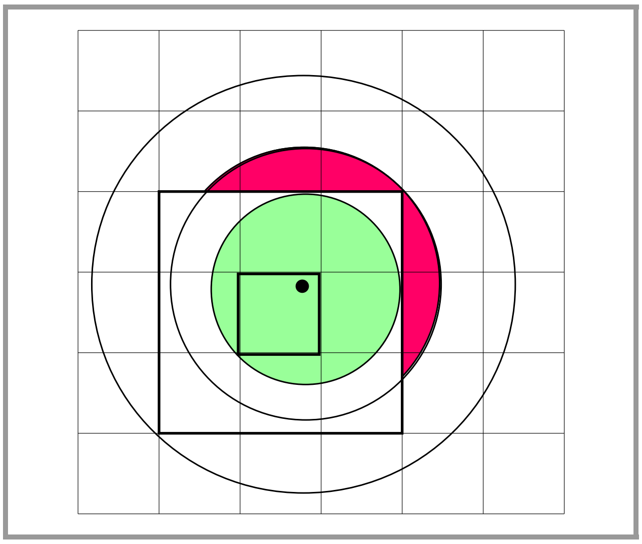


Fig. 2. A flaw in to-the-right-and-down approach.

Let us consider a 2-dimensional case and a point which is located somewhere in the center square. The smaller green circle shown in Fig. 2 covers the entire area inside a 3-by-3 square, such that for any point therein, if there are points closer to the one that is marked, then they are also in that area. This is not true for a larger circle. In this case, for every point outside the green circle, there are points outside the 3-by-3 square which are closer to the one marked, but because they are not situated in the neighboring squares, they will not be included in the search. This error can be fully eliminated only when Fischler's formula is used, because it guarantees that the minimum distance will be lower than the radius of the smaller circle. The largest circle corresponds to all points taken into consideration. In a 2-dimensional case, a 3-by-3 square occupies less than 36% of the maximum size circle, in 3D – less than 16% of the maximal sphere, and in 5D – less than 2.6%, which creates a significant discomfort.

We will use the cumulative distribution function for the minimum distance between n points in a t dimensional hypertorus:

$$\Pr(D \leqslant d) = 1 - e^{\frac{-d^t n(n-1) V_t(1)}{2}} ,$$

where

$$V_t(r) = \begin{cases} \dfrac{\pi^h r^t}{h!} , & t = 2h , \\ \dfrac{2^h \pi^{h-1} r^t}{t!!} , & t = 2h-1 , \end{cases}$$

is the volume of a t dimensional hypersphere of radius r [10].

After skipping the floor function, in the minimum distance test, we get the length of the sub-cube's edge, which is equal to the radius of the smaller circle as:

$$K = \frac{1}{3}\sqrt[t]{\frac{-2\ln\varepsilon}{n(n-1)}} ,$$

and the probability of the minimum distance being greater is:

$$\varepsilon^{\frac{V_t(1)}{3^t}} ,$$

which for $\varepsilon = 10^{-20}$ equals 10^{-7} for $t = 2$, 0.00079 for $t = 3$, and 0.369 for $t = 5$.

By inverting this argumentation, we can easily find such a sub-cube size that assures an arbitrarily low probability of that flaw affecting the computed distances. From:

$$e^{\frac{-d^t n(n-1) V_t(1)}{2}} \leqslant \varepsilon ,$$

we get

$$d \geqslant \sqrt[t]{-\frac{2\ln\varepsilon}{n(n-1)V_t(1)}} ,$$

which returns the number of sub-cubes

$$K = \left\lfloor \sqrt[t]{-\frac{n\,(n-1)\,V_t\,(1)}{2\ln\varepsilon}} \right\rfloor^t .$$

Assuming n is large and omitting the floor function, the total number of calculated distances is:

$$-\ln\varepsilon\frac{3^t}{V_t\,(1)}.$$

According to the proposed method, for the same three examples, we have:

- $1{,}515^2$ squares and 251 distances,
- 91^3 cubes and 573 distances,
- 14^5 5D cubes and 4,065 distances.

Those numbers are substantially worse than those obtained previously, but the expected number of computed distances is still independent of the number of points.
This flaw will not play an important role in the following two tests, because the number of points in each 3-by-... cube will be significantly greater than one.

3. Bickel-Breiman Test

In the Bickel-Breiman test [11], the aim is to find the nearest point for every point out of n points randomly placed in a t dimensional hypertorus. In this case, the analytic approach leads to the lack of ability of dividing into sub-cubes – for a randomly chosen point, the maximum distance to the nearest point is $\frac{\sqrt{t}}{2}$.
To cope with this, we need a collision for every point placed, so we start with the probability of not finding any other point in the 3^t cube:

$$\left(1-\frac{3^t}{K}\right)^{n-1} ,$$

and then the probability that not all the points will find their neighbor is:

$$1-\left[1-\left(1-\frac{3^t}{K}\right)^{n-1}\right]^n \leqslant \varepsilon \ .$$

With simple transformations we obtain:

$$K \leqslant \frac{3^t}{1-\sqrt[n-1]{1-\sqrt[n]{1-\varepsilon}}} \ .$$

As before, the expected number of calculated distances per cube is:

$$E_{BB} = \frac{1}{2}n\,(n-1)\,\frac{1}{K^2}\left[3^t\left(1-\frac{1}{K}\right)+\frac{1}{K}\right] ,$$

and the total number of calculated distances:

$$\frac{1}{2}n\,(n-1)\,\frac{1}{K}\left[3^t\left(1-\frac{1}{K}\right)+\frac{1}{K}\right] ,$$

which is analogically in naive method, "only" with coefficient $\frac{1}{K}\left[3^t\left(1-\frac{1}{K}\right)+\frac{1}{K}\right]$.
This time the improvement is not as spectacular as in the previous case. It could be approximated as:

$$\frac{1}{1-\sqrt[n-1]{1-\sqrt[n]{1-\varepsilon}}}$$

times faster.

For the same three examples, this time we have:

- 36^2 squares and 144 times fewer calculated distances,
- 12^3 cubes and 64 times fewer calculated distances,
- 6^5 5D cubes and 32 times fewer calculated distances,

than in the naive approach, when $\varepsilon = 10^{-20}$.

4. *m* Nearest Points Test

This test can be found in the TestU01 library [12] and is described in paper [10].
Because of the existing recommendation that $n \geqslant 4m^2$, we can assume that the number of nearest points m is much smaller than the total number of points n. In fact, it is so small that in every case from examples presented in Section 3, we get more distances calculated per every subcube than actually needed.
To find m nearest pairs, at least m collisions are needed, having the probability of:

$$1-\sum_{r=0}^{m-1}\frac{K!}{(K-n+r)!K^n}\left\{\begin{matrix} n \\ n-r \end{matrix}\right\} .$$

For given n and m, while applying Stirling's formula for factorial and tabulating values of all needed Stirling's numbers, the number of cubes K can be efficiently computed

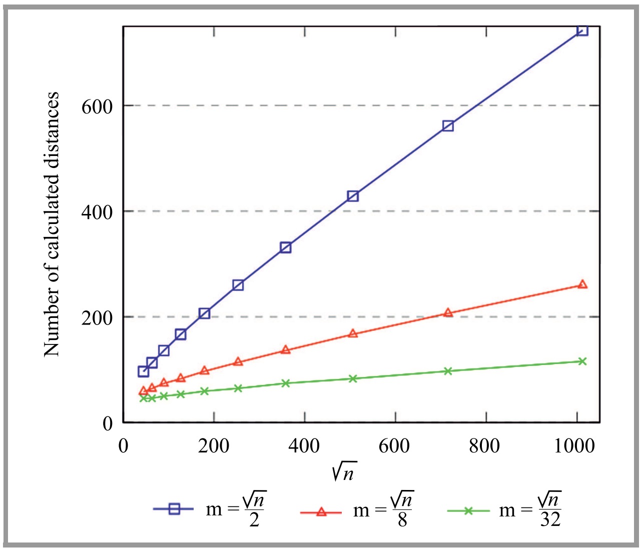


Fig. 3. Expected number of calculated distances as a function of $\sqrt{n}$ in a 2-dimensional case.

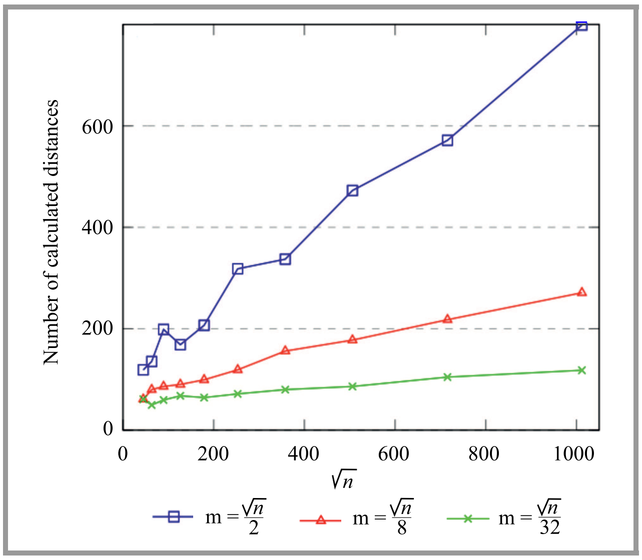


Fig. 4. Expected number of calculated distances as a function of $\sqrt{n}$ in a 5-dimensional case.

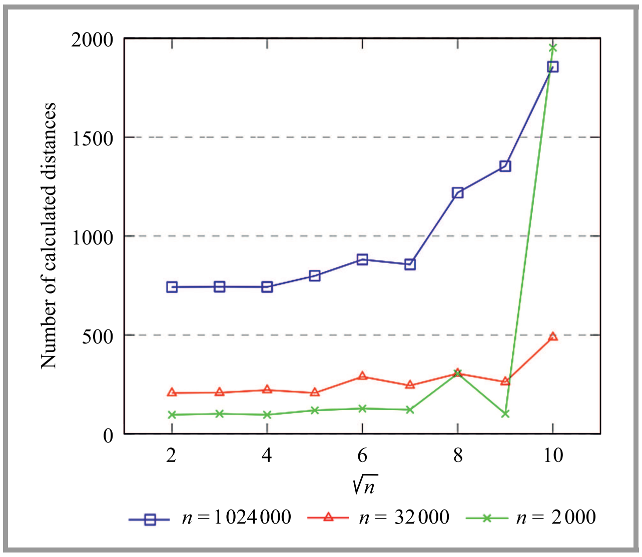


Fig. 5. Expected number of calculated distances as a function of the number of dimensions.

numerically. Like in the minimum distance test, the obtained value of K should be multiplied by 3^t.

On the three examples researched, we have:

- $1,455^2$ squares and 1632 times fewer calculated distances,
- 123^3 cubes and 1,076 times fewer calculated distances,
- 24^5 5D cubes and 1,024 times fewer calculated distances,

than using the formula for the Bickel-Breiman test, so the effort pays off.

The three graphs presented show the relation between the number of points, dimensions, and the expected number of calculated distances. In Figs. 3 and 4, one can observe that in every case the expected number of calculated distances grows linearly according to the $\sqrt{n}$ or, in fact, to the m. The deviations observed arise from the necessary flooring of $\sqrt[t]{K}$.

Figure 5 shows how the number of dimensions affects the expected number of calculated distances for a different number of points. In all cases $m = \frac{\sqrt{n}}{2}$.

Similarly, the gradual increase in the number of calculated distances arises from the flooring of $\sqrt[t]{K}$ – otherwise they would be constant. To understand the cause of that growth, one can calculate the actual number of sub-cubes for a different t. Let $n = 1,024,000$, $m = 505$ and $\varepsilon = 10^{-20}$, for which we obtain $K = 706,379,797$ and from the formula:

$$\left\lfloor \sqrt[t]{K} \right\rfloor^t$$

values from 282,475,249 ($t = 10$) to 706,336,929 ($t = 2$).

5. Implementation

At the beginning, let us consider the minimum distance test and four approaches to implementing it:

- naive implementation using a general-purpose single thread processor (CPU),
- implementation based on sub-cubes with Fischler's result, using a general-purpose single thread processor,
- implementation based on sub-cubes with the presented method, using a general-purpose single thread processor,
- parallel implementation on a GPU engine.

None of the above methods can be compared with regard to the number of computed distances only, due to the fact that:

- naive implementation is based on calculating the distance for all pairs of points,
- naive implementation requires no additional memory,
- the amount of time consumed to review the data will depend on the number of sub-cubes,
- there is no obvious method for implementing the sub-cubes method using a GPU,
- the parallel implementation on a GPU accelerator consists in calculating the distance for all other points for every point,
- the parallel implementation works on a GPU whose compute cores differ significantly from those of CPU cores for which the sub-cubes method is suitable.

Depending on the hardware platform used, it may turn out that the following estimates are inadequate, but they should be easy to adapt and reproduce.

To perform any valid comparisons between the four cases mentioned above, two factors have to be determined:

- the ratio between naive and on sub-cubes based implementations,
- comparison of the speed of a single core of CPU and GPU.

The second factor is easy to obtain. We will consider two top computational devices:

- Tesla V100S PCIe accelerator with 5120 Nvidia CUDA cores and 16.4 TFLOPS on single-precision performance [13],
- AMD Ryzen Threadripper 3990X processor with 64 cores and 128 threads [14] and 3.7 TFLOPS on single-precision performance [15].

It may be assumed that:

- Tesla should be able to calculate 4.43 times more distances than the AMD processor,
- the single core of the AMD processor should be able to calculate 9.02 times more distances than the single core of the Tesla processor.

Based on this assumption, the parallel implementation on the Nvidia Tesla GPU may be estimated as:

- 443 times faster for $n = 8,000$ and $t = 2$ or $n = 4,000$ and $t = 3$,
- 332 times faster for $n = 3000$ and $t = 5$,

than naive implementation on a single core of the AMD CPU. For $n = 8,000$ and $t = 2$, we assumed two runs of calculations due to lack of cores limitations.

As mentioned above, the results will vary depending on the quality of the implementation, so we have limited our two procedures to the evaluation of the square of the distance. Therefore, the only optimization was the limitation of the number of sub-cubes to the power of 2 only.

Table 1
Mean time to complete the calculations for different test cases and methods

Test case method	$n = 8000$ $t = 2$	$n = 4000$ $t = 3$	$n = 3000$ $t = 5$
Naive	180.14	67.53	63.81
Fischler's	1.13	0.95	10.23
Proposed	5.19	1.49	1.72

In Table 1 we present the mean times in milliseconds, required to complete the calculations for different test cases and methods used.

Based on these results, the following conclusions may be formulated:

- time for the naive method strictly depends on the number of calculated distances,
- the sub-cubes method is significantly faster,
- time for sub-cubes depends not only on the number of calculated distances, but also on the number of sub-cubes,
- depending on the test case, the number of sub-cubes based on the Fischler's estimate or on the proposed approach renders better results, suggesting that neither of them is optimal.

The last of the above findings has led to a series of tests, each consisting in the execution performed for a different number of cubes.

For the t-dimensional case, the number of sub-cubes is 2^{dt}, where $d = 2, 3, \ldots$, and 2^d is the number of divisions for every dimension. $d = 2$ is the smallest reasonable case, and d, based on our results, is the maximum case. In Table 2 we show the mean times (in milliseconds) required to complete the calculations for different test cases and numbers of sub-cubes.

Table 2
Mean time to complete the calculations for different test cases and numbers of sub-cubes

Test case d	$n = 8000$ $t = 2$	$n = 4000$ $t = 3$	$n = 3000$ $t = 5$
2	151.67	19.26	10.23
3	24.48	2.21	1.72
4	4.22	0.95	
5	1.53	0.69	
6	1.13	1.49	
7	1.00		
8	1.12		
9	1.81		
10	5.19		

This means that the tuned sub-cubes method is capable of evaluating the minimum distance 180, 98, and 37 times faster in the considered cases, compared to the naive method.

Finally, implementations performed with the use of the CPU and GPU can be compared. A simple division shows that a parallel implementation on a GPU will complete a single evaluation run 2.5 to 9 times faster than a single-core implementation using the sub-cubes method, but the AMD processor may handle up to 128 parallel runs, so the total throughput is greatly in favor of the tuned sub-cubes method.

For the m nearest pairs test and for a small m, the results are similar to those presented above. For bigger values, as well as for the Bickel-Breiman test, the GPU-based implementation will be faster.

6. Conclusion

The Fischler's approach, a solution that has so far been the best known deterministic method for speeding up the minimum distance test performed on a single thread CPU, was significantly improved. Furthermore, means for similar improvements of two other tests have been presented. The methods presented, although of the probabilistic variety, offer the ability to practically mitigate error probability. Thanks to the significant reduction in the number of the computed distances, the tests could be performed on longer sequences, thus increasing their detection capabilities. After precise tuning, minimum distance and *m* nearest pairs tests may be run on modern multicore CPU processors and are capable of outperforming GPUs.

Because minimum distance and 3D spheres tests are only some of many tests in Dieheard's and Dieharder's portfolios, total time complexity gains seem not impressive. On the other hand, the results obtained allow to extend the existing packages by tests covering much longer sequences and thus offering a greater detection capacity.

Further work may be concerned with generalizing the presented methods in order to apply these to more general problems, e.g. the closest bichromatic pair problem.

References

[1] M. I. Shamos and D. Hoey, "Closest-point problems", in *Proc. of the 16th Ann. Symp. on Found. of Comp. Sci.*, Berkeley, CA, USA, 1975 pp. 151–162 (DOI: 10.1109/SFCS.1975.8).

[2] J. L. Bentley and M. I. Shamos, "Divide and conquer in multidimensional space", in *Proc. of 8th Ann. ACM Symp. on the Theory of Comput.*, Hershey, PA, USA, 1976. pp. 220–230 (DOI: 10.1145/800113.803652).

[3] J. L. Bentley, "Multidimensional divide-and-conquer", *Comm. of the Assoc. for Comput. Machinery*, vol. 23, no. 4, pp. 214–229, 1980 (DOI: 10.1145/358841.358850).

[4] K. Hinrichs, J. Nievergelt, and P. Schorn, "Plane-sweep solves the closest pair problem elegantly", *Inform. Process. Lett.*, vol. 26, no. 5, pp. 255–261, 1988 (DOI: 10.1016/0020-0190(88)90150-0).

[5] S. Khuller and Y. Matias, "A simple randomized sieve algorithm for the closest-pair problem", *Inform. and Comput.*, vol. 118, no. 1, pp. 34–37, 1995 (DOI: 10.1006/inco.1995.1049).

[6] A. Andoni, P. Indyk, and I. Razenshteyn, "Approximate nearest neighbor search in high dimensions", 2018. [Online]. Available: https://arxiv.org/pdf/1806.09823

[7] M. Fischler, "Distribution of minimum distance among N random points in d dimensions", Technical Rep. no. FERMILAB-TM-2170, Fermi National Accelerator Lab., Batavia, IL, USA, 2002 (DOI: 10.2172/794005).

[8] M. Rütti, "A random number generator test suite for the C++ standard", Diploma Thesis, Institute for Theoretical Physics, ETH Zürich, 2004 [Online]. Available: http://comp-phys.org/software/rngts/doc/main.pdf

[9] E. Luengo and L. García Villalba, "Recommendations on statistical randomness test batteries for cryptographic purposes", *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–34, Article no. 80, 2021 (DOI: 10.1145/3447773).

[10] P. L'Ecuyer, J. F. Cordeau, and R. Simard, "Close-point spatial tests and their application to random number generators", *Operations Res.*, vol. 48, no. 2, pp. 189–350, 2000 (DOI: 10.1287/opre.48.2.308.12385).

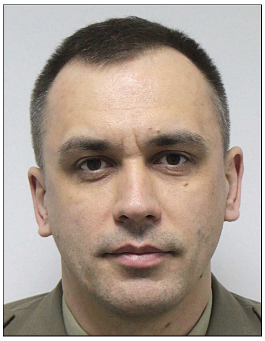
[11] P. J. Bickel and L. Breiman, "Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test", *Ann. of Probab.*, vol. 11, no. 1, pp. 185–214, 1983 (DOI: 10.1214/aop/1176993668).

[12] P. L'Ecuyer and R. Simard, "TestU01: A C library for empirical testing of random number generators", *ACM Trans. on Math. Softw.*, vol. 33, no. 4, Article no. 33, pp. 1-44, 2007 (DOI: 10.1145/1268776.1268777). 33.

[13] Nvidia V100 Tensor Core GPU, Nvidia Corporation V100 Datasheet, 2020 [Online]. Available: https://images.nvidia.com/content/technologies/volta/pdf/volta-v100-datasheet-update-us-1165301-r5.pdf

[14] AMD Ryzen Threadripper 3990X Processor [Online]. Available: https://www.amd.com/en/products/cpu/amd-ryzen-threadripper-3990x

[15] AMD Ryzen Threadripper PRO 3995WX GFLOPS performance [Online]. Available: https://gadgetversus.com/processor/amd-ryzen-threadripper-pro-3995wx-gflops-performance/



Krzysztof Mańk received his M.Sc. degree from the Military University of Technology in 1999. He has been working at the Institute of Mathematics and Cryptology of MUT ever since. His main research interests focus on stream ciphers and test of randomness.

https://orcid.org/0000-0002-5048-9049
E–mail: krzysztof.mank@wat.edu.pl
Institute of Mathematics and Cryptology
Faculty of Cybernetics
Military University of Technology
Warsaw, 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.itl.waw.pl/en/jtit-for-authors

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@itl.waw.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.itl.waw.pl/en/jtit-for-authors

Journal of Telecommunications and Information Technology has entered into an electronic licencing relationship with EBSCO Publishing, the worlds most prolific aggregator of full text journals, magazines and other sources. The text of *Journal of Telecommunications and Information Technology* can be found on EBSCO Publishings databases. For more information on EBSCO Publishing, please visit www.epnet.com.

(*Contents Continued from Front Cover*)

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

National Institute
of Telecommunications