Discover Applied Sciences

Research

Signalling overhead diminution in M-MIMO using NOMA transmission

M. Imran[1] · O. Hayat[2] · R. Ngah[3]

Received: 14 September 2023 / Accepted: 6 December 2023
Published online: 20 January 2024

Abstract
With the increase in demand for mobile and Internet of Things devices, Orthogonal Multiple Access (OMA) cannot manage the maximum number of users. In this technique, the number of radio frequencies must be equivalent to the number of users. It causes to increase in signalling overhead. Therefore, it requires special attention to reduce the signalling overhead to escalate the spectral and energy efficiency. This paper investigates different Non Orthogonal Multiple Access (NOMA) techniques, and the results are compared with OMA techniques. A novel NOMA technique Multi User Shared Access (MUSA) is applied. It adjusts maximum users and has good spectral and energy efficiency compared to OMA techniques. In this proposed technique, maximum complex spreading codes are generated for the users and each user picks that code and transmits its data at the same radio frequency chain. The proposed scheme MUSA-NOMA has 12.8% more energy efficiency and 6.51% spectral efficiency compared to SCMA-NOMA and 32% more energy efficiency and 18.5% spectral efficiency compared to OMA.

Article highlights

- Imagine you're using your phone in a crowded area where many people are also trying to connect, signalling congestion will occur. In traditional setups, it is difficult to manage everyone's signals.
- Instead of dealing with each person's signal separately, NOMA allows the communication system to treat similar groups together, streamlining the process.
- It's like having group discussions instead of one-on-one talks, making the whole communication setup more efficient and less complicated. This way, your phone and the network can handle many connections more smoothly, providing a better experience for everyone.

Keywords OMA · NOMA · MUSA · M-MIMO · Signalling Overhead

1 Introduction

With the fast growth of the Internet of Things (IoTs) and Device-to-Device (D2D) Communications, it is expected that the global mobile information traffic will grow from 10 to 100 times by 2030. However, the existing Fourth Generation (4G) system is inept in accomplishing the requirements, particularly in spectrum efficiency. To accomplish elevated

✉ O. Hayat, ohayat@numl.edu.pk | [1]International Islamic University, Islamabad, Pakistan. [2]National University of Modern Languages (NUML), Islamabad, Pakistan. [3]Wireless Communication Centre (WCC), Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Bahru, Johor, Malaysia.

spectrum efficiency, a critical control and scheduling procedure is required and that causes substantial signalling [1]. The uplink transmission of each base station requires to be granted and scheduled separately and this is mostly done in orthogonal access. In orthogonal access, massive connections would lead to significant signalling overhead in 4G [2]. This signalling overhead increases the energy consumption of devices, as well as makes the terminal structure more composite and costly [3].

The concept of Massive Multiple Inputs Multiple Outputs (M-MIMO) system in Fifth Generation (5G) is to offer more beam-forming gain and multiplexing gain [4]. Indeed, it has been revealed that M-MIMO can accomplish a large magnitude of increment in the capacity of the system [5]. Though, it is very hard to implement M-MIMO practically because of high transceiver mobility and unpredictability [6]. In the existing system, every antenna in the Multiple Inputs Multiple Outputs (MIMO) system generally needs one committed Radio Frequency (RF) chain [7]. In this way, the use of huge numbers of antennas in M-MIMO networks leads to a similarly large amount of RF chains [8]. It is also observed that RF equipment may devour up to 70% energy of a transceiver [9]. Subsequently, the energy consumption and equipment cost brought about by many RF chains in M-MIMO networks become excessively expensive practically. When the quantity of users increases from the dedicated RF chains then signalling overhead increases accordingly.

A lot of studies have been considered to address this issue to reduce energy uses, hardware complexity, and to maximize spectral efficiency. To address these issues antenna selection technique is considered [10–12], but the performance loss still exists in this system. In recent times the idea of beam-space MIMO is presented in the leading work to primarily lessen the amount of needed RF chains in the Millimetre Wave (mm-Wave) M-MIMO system [13]. By utilizing the lens antenna, which assumes a role in acknowledging spatial discrete Fourier Transform beam space MIMO can change the ordinary spatial channel to the beam space channel to catch the channel sparsity at mm-Wave frequencies [14]. Accordingly, the prevailing beams are chosen as indicated by the sparse beam space channel to decrease the amount of desired RF chains. In addition, by the utilization of a lens antenna, thin beams are protected even with a decreased amount of RF chains. It permits fundamentally diminishes the energy needed on each beam and interferences [15]. Therefore, the antenna assortment procedure and the exhibition of beam-space MIMO with beam assortment is near to ideal. Nonetheless, a crucial constraint of beam-space MIMO is clearly considered in distributed manner and every RF chain can just help one client at parallel time–frequency resources.

Accordingly, the maximum sum of users that are supported can't surpass the sum of RF chains [4]. The key purpose is that independence given by the RF chains needs to be greater than or equivalent to the independence required by the end users. In any case, signals for various users can't be isolated by linear activity [16]. In a conventional Orthogonal Multiple Access (OMA) based MIMO system, the sum of demanded RF chains is up to the sum of base station antennas, which is typically huge for OMA based M-MIMO system. Therefore, the immediate use of M-MIMO at high frequencies is expensive due to bigger hardware detriment and energy utilization consumed by RF chains, for example, 250 mW is expended by an RF chain while 64 W is necessary by a M-MIMO system alongside 256 antenna [17].

With the exponential increase of users, OMA cannot manage the maximum number of users and causes signalling overhead. Therefore, different Non-Orthogonal Multiple Access (NOMA) techniques are investigated and applied, and the results are compared with OMA techniques. A novel NOMA technique labelled Multi User Shared Access (MUSA) is selected and it adjusts maximum users and has good spectral efficiency and energy efficiency as compared to OMA. The key purpose following MUSA is non-orthogonal communication and give open access. Every user's data is expanded with a group of complex spreading sequences and user data overlap on the receiving side where Successive Interference Cancelation (SIC) is applied to abort interference among users. The spreading sequence is explicitly intended to deal with the heavy over-burdening of users and to help basic SIC at the receiving end. With non-orthogonal transmission and independence, MUSA provides a massive number of links while minimizing signalling overhead and power feasting simultaneously. The proposed technique achieves maximum spectral and energy efficiency by adjusting the maximum number of users.

2 System model

A single-cell transceiver M-MIMO-NOMA scheme is considered in this paper, where the base station is equipped with K antennas and K will be RF chains for simultaneous transmission to serve K single-antenna users [18] [19–21]. A conventional OMA system is shown in Fig. 1a, while equivalent to a NOMA system is also presented in Fig. 1b. It can be observed from Fig. 1, that in a conventional OMA system, each antenna involves one reserved RF chain, and thus the sum of RF chains is equivalent to the total number of antennas, resulting in excessive energy consumption and hardware cost. There

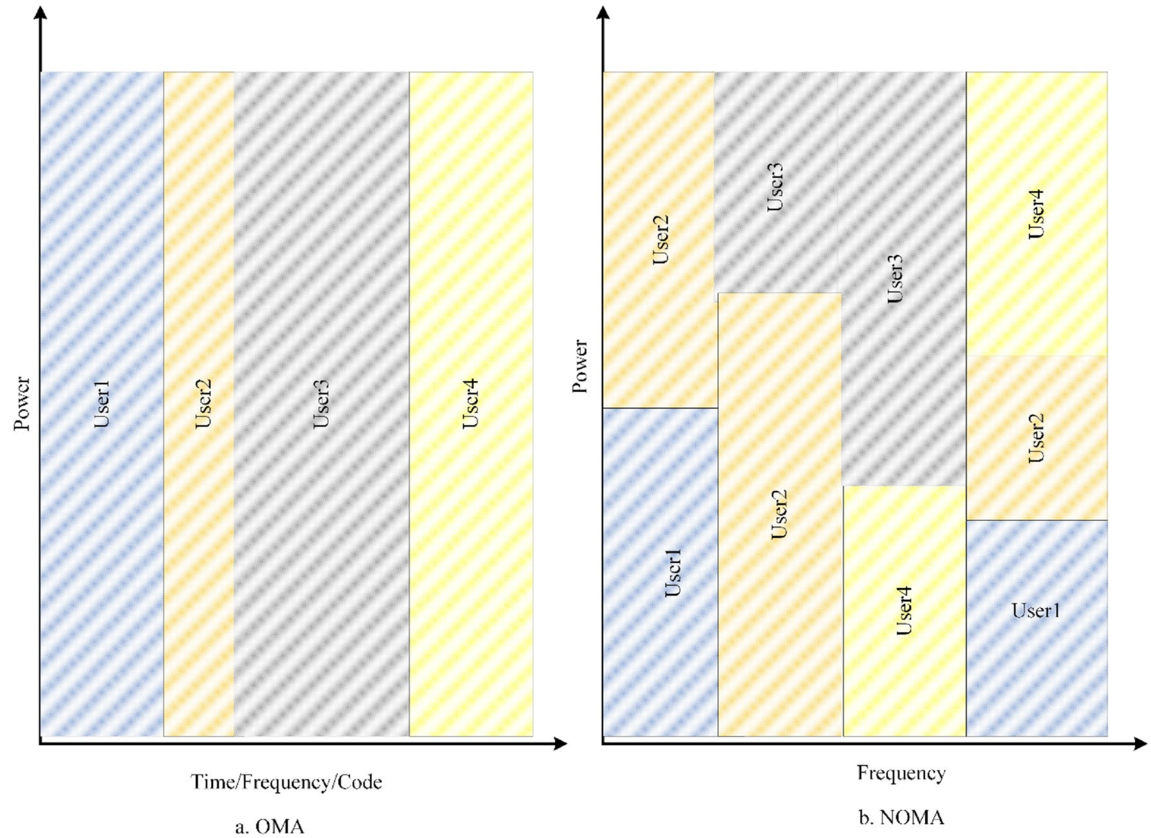


Fig. 1 **a** A conventional OMA **b** NOMA system

are two important types of NOMA systems which are the Sparse Code Multiple Access NOMA (SCMA-NOMA) as presented in Fig. 2 and MUSA-NOMA architecture as shown in Fig. 3. On the contrary, the sum of RF chains in SCMA-NOMA design is less than the sum of antennas, which can be apprehended by a codebook scheme of SCMA-NOMA [22]. In the SCMA-NOMA architecture, multiple codebooks are generated for the users, where each user specifically picks one codebook and correlates its data with codebooks and sends it to the channel. Therefore, multiple users can share data in a single channel at the same time and frequency. For the MUSA-NOMA architecture, multiple families of complex spreading codes are generated for the users, where each user picks one specific code and spreads its modulated data with complex spreading code and multiple users send data at the same channel at the same time and frequency. Consequently, the MUSA-NOMA system is more spectrally efficient and more energy-efficient, and therefore, MUSA-NOMA architecture is considered in this research.

2.1 OMA-based MIMO system

In a traditional OMA-based M- MIMO system, the total users cannot increase the sum of radio frequency chains. In this system the received signal vector $r = \left[r_1, r_2, \ldots \ldots, r_k\right]^T$ can be expressed as

$$r = G^H X P_c + \omega \tag{1}$$

where $c = \left[c_1, c_2, c_3, \ldots \ldots c_k\right]^T$ has $K \times 1$transmitted signal entire K users with normalized power $F\left(cc^H\right) = I_K, P = diag(p)$ involves transmitted power entire K users where $P = [\sqrt{p_1}, \sqrt{p_2}, \sqrt{p_3} \ldots \ldots \sqrt{p_K}]$satisfies $\sum_{k=1}^{K} p_k \leq P$ and this is the largest transmitted power at the base station), $X = [x_1, x_2, x_3, \ldots \ldots x_k]$is the $N \times K$ matrix with$\|x_k\|_2 = 1$for $k = 1, 2, 3, \ldots \ldots, K$, and ω is a noise vector with subsequent distribution $\mathcal{L}\left(\mathcal{M}\left(0, \sigma^2 I_K\right)\right)$. Finally, $G = [g_1, g_2, g_3, \ldots \ldots, g_K]$of size $N \times K$ channel matrix, where g_K of size $N \times 1$designates the spatial channel vector concerning the base station and the Kth user. In this research, "the Saleh Valenzuela model" is considered [23, 24], so g_K can be represented as

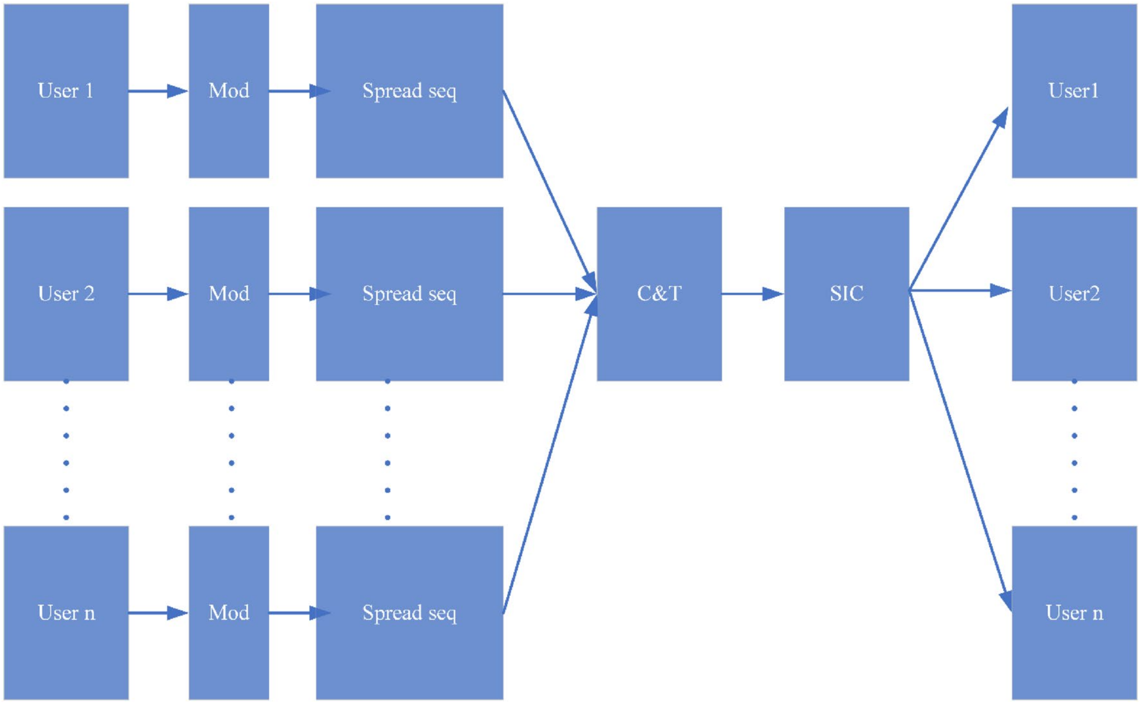


Fig. 2 Sparse code multiple access NOMA (SCMA-NOMA) architecture

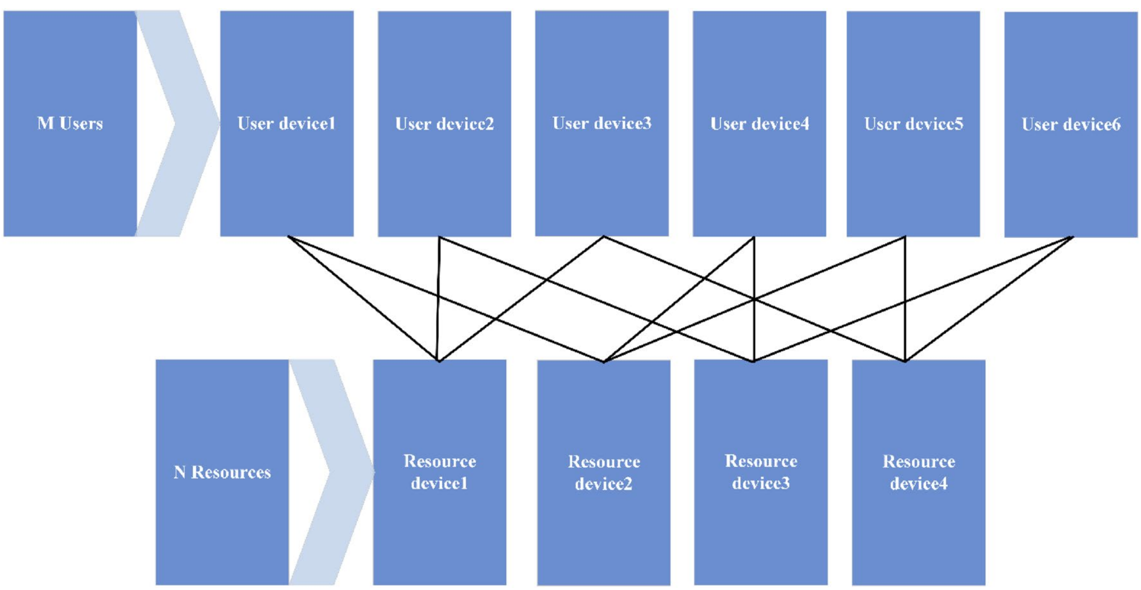


Fig. 3 MUSA-NOMA architecture

$$g_K = \gamma_K^{(0)} b\left(\vartheta_K^{(0)}\right) + \sum_{t=1}^{T} \gamma_K^{(t)} b\left(\vartheta_K^{(t)}\right) \tag{2}$$

where $\gamma_K^{(0)} b\left(\vartheta_K^{(0)}\right)$ is the line of sight (LoS) users of the k^{th} user, in which $\gamma_K^{(0)}$ indicates the complex gain and $b\left(\vartheta_K^{(0)}\right)$ denotes the spatial path.$\gamma_K^{(t)} b\left(\vartheta_K^{(t)}\right)$ for $1 \le t \le T$ is the tth non-line of site (NLoS) equipment's of Kth user, Where T is the total NLoS equipment's. $b(\vartheta)$ is the $N \times 1$array steering vector. It is noted that at high frequencies, the amplitude $\left\{\left|\gamma_K^{(t)}\right|\right\}_{t=1}^{N}$ of

NLoS equipment's are typically $5to10dB$ weedy than the amplitude $\left|\gamma_K^{(0)}\right|$ of the LoS equipment [25]. For the standard uniform and linear array steering vector $b(\vartheta)$ can be stated as,

$$b(\vartheta) = \frac{1}{\sqrt{M}}[e^{-j2\pi\vartheta n}]_{n\in J(M)} \tag{3}$$

where $J(M) = \{i - \frac{M-1}{2}, i = 0, 1, 2, \ldots\ldots, M-1\}$ is a symmetric set of symbols positioned nearby zero. The spatial path is marked as $\vartheta = \frac{c}{\lambda}\sin\theta$, where θ is the tangible direction fulfilling $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$, λ is the wavelength, and c is antenna spacing.

2.2 SCMA-NOMA based MIMO system

The SCMA-NOMA based M-MIMO system has been newly researched, which uses codebooks to decrease the number of desired RF chains in M-MIMO system substantially without any performance deficit. A SCMA-NOMA encoder is defined as $\mathfrak{F} : \mathfrak{D}^{\log_2(N)} \to \mathbb{W}, \mathbb{W} = \mathcal{E}(\boldsymbol{d})$ where $\mathbb{W} \subset \mathfrak{D}^L$ with cardinality $|\mathbb{W}| = N.L$ shows the dimension of $\sqsupseteq$ code word and sparse vector with $J < L$ non-zero input. Let $\boldsymbol{d}$ represents an N-dimensional complex constellation point described within the constellation set $\mathfrak{R} \subset \mathbb{R}^J$ such that: $\mathfrak{D}^{\log_2(N)} \to \mathfrak{R}, \boldsymbol{d} = h(\boldsymbol{d})$. SCMA encoder is defined as $\mathfrak{F} :\equiv \boldsymbol{U}h$ where the binary mapping matrix $\boldsymbol{U} \in \mathfrak{D}^{L\times J}$ solely covers the J measurements of a constellation point to a L-dimensional SCMA codeword. State that $\boldsymbol{U}$ comprises $L-J$ all-zero rows and hence, all the code words in the codebooks contain 0 in the same $L-J$ dimensions. Eradicating the whole zero rows from $\boldsymbol{U}$, the remaining can be characterized by I_M identity matrix defined the binary mapper, and it does not permute the dimensions during the mapping process of subspace $\mathfrak{R}$. The resulting codebook comprises of N codewords separately containing L complex assessment from which the highest J are nonzero stated by the matrix $\boldsymbol{U}$.

A SCMA encoder contains I seprate layers each defined by $\rho_i(U_i, h_i, O_i, P_i, L)$, i = 1,2, 3,…,K. The constellation sets $\mathcal{B}_i$ is generated by the constellation function h_i with length P_i using O_i alphabets. The P_i constellation points in ith dimension are mapped on the matrix U_i to SCMA code words form the code set X_i. Without forfeiture of generalization, it is supposed that all stages have an equal constellation dimension and extent, for example $O_i = O, P_i = P, \forall_i$. By summarizing, SCMA code can also be exemplified by $\rho\left(\left[U_i\right]_{j=1}^K, \left[h_i\right]_{j=1}^K; K, O, P, L\right)$.

L allocated orthogonal resources are used to multiplex of SCMA code words, for example, MIMO spatial layers and Orthogonal Frequency Division Multiple Access (OFDMA) tones. After the multiplexing of the synchronous layer, the received signal can be stated as

$$\boldsymbol{r} = \sum\nolimits_{\boldsymbol{i}=1}^{\boldsymbol{K}} diag\left(g_i\right)\boldsymbol{v_i} + \boldsymbol{\omega} \tag{4}$$

$$\boldsymbol{r} = \sum\nolimits_{\boldsymbol{i}=1}^{\boldsymbol{K}} diag\left(g_i\right)\boldsymbol{x_i}\boldsymbol{h_i}\left(\boldsymbol{a_i}\right) + \boldsymbol{\omega} \tag{5}$$

where $\boldsymbol{v_i} = \left(\boldsymbol{v}_{1\boldsymbol{i}}, \boldsymbol{v}_{2\boldsymbol{i}}, \boldsymbol{v}_{3\boldsymbol{i}}, \ldots, \boldsymbol{v}_{\boldsymbol{Pi}}\right)^T$ is the SCMA codeword of the layer, $g_i = \left(g_{1i}, g_{2i}, g_{3i}, \ldots\ldots g_{Pi}\right)^T$ is the ith layer channel vector and $\boldsymbol{\omega} \cong \mathcal{LM}\left(0, M_0 I\right)$ is the ambient noise, $\boldsymbol{h_i} = \boldsymbol{h}$, the target receiver channel and hence (4) reduced to $\boldsymbol{r} = diag(g)\sum_{\boldsymbol{i}=1}^{\boldsymbol{K}}\boldsymbol{v_i} + \boldsymbol{\omega}$. Using the received signal $\boldsymbol{r}$ and channel knowledge $[\boldsymbol{h_i}]_{i=1}^K$ the optimised detected layer codewords $W = (w_1, w_2, w_3, \ldots\ldots, w_K)$ can be expressed as

$$\hat{\boldsymbol{X}} = \mathrm{argmax} X \in \left(q_{i=1}^K\right)t_i p\left(\frac{W}{r}\right) \tag{6}$$

where $\left(q_{i=1}^K\right)t_i = qt_1, qt_2, qt_3, \ldots\ldots qt_K$. Now this problem has become Matrix Power Function (MPF) problem, which is much simpler to resolve. The complexity of each source node is proportional to N^{eg} while the complexity of MPF is N^I.

3 Proposed MUSA-NOMA based M-MIMO system

To increase spectrum and energy efficiency, and the degree of connectivity in M-MIMO systems, a MUSA-NOMA is proposed. In contrast to the current OMA based MIMO, two or more users are served simultaneously within each designated RF chain in the MUSA based NOMA structure. Specifically, code selection algorithms for example maximization of the Signal-to-Interference-plus-Noise Ratio (SINR) selection and maximum magnitude selection can be utilized to choose one dedicated code for an individual user, and individual RF chain tallies to one code. Note that unlike users are expected to choose the same code, called "differing users." Mostly, for a standard OMA based M-MIMO system with a number of antennas $K = 256$ and number of users $N = 32$ with spatial direction having uniform distribution and the probability for existing users' selection is 87% on the same code. In existing OMA-MIMO systems, users scheduling is accomplished to choose one user out of differing users. The differing users are served simultaneously using the same RF chain in the suggested MUSA based NOMA system. Although the total selected are K_{RF}, the total served simultaneous users N are greater than K_{RF},$N \geq K_{RF}$. T_m for $m = 1, 2, 3, \ldots \ldots, K_{RF}$ represent the total users in the mth code with $T_i \bigcap T_j = \emptyset$ for $i \neq j$ and $\sum_{m=1}^{K_{RF}} |T_m| = N$. The $K_{RF} \times 1$ coded channel vector subsequently the code selection of BS and pth user in qth code and is indicated by $g_{p,q}$, and v_q of size $K_{RF} \times 1$ indicates the standardized precoding vector for users in qth code. It is assume that $\|g_{1,n}^H v_q\|_2 \geq \|g_{2,n}^H v_q\|_2 \geq \|g_{3,n}^H v_q\|_2 \geq \ldots \|g_{|T_m|,n}^H v_q\|_2$ for $n = 1, 2, 3 \ldots \ldots, K_{RF}$. The received signal $r_{p,q}$ at the pth user in the qth code ($q = 1, 2, 3, \ldots \ldots, K_{RF}$, and $p = 1, 2, 3, \ldots ..|T_m|$) is stated as

$$r_{p,q} = g_{p,q}^H \sum_{j=1}^{K_{RF}} \sum_{i=1}^{|T_j|} v_j \sqrt{q_{i,j} t_{i,j}} + \omega_{p,q} \tag{7}$$

$$= g_{p,q}^H v_q \sqrt{q_{p,q} t_{p,q}} + g_{p,q}^H v_q \sum_{i=1}^{p-1} \sqrt{q_{i,q} t_{i,q}} + g_{p,q}^H v_q \sum_{i=p+1}^{|T_m|} \sqrt{q_{i,q} t_{i,q}} + g_{p,q}^H \sum_{j \neq q} \sum_{i=1}^{|T_j|} v_j \sqrt{q_{i,j} t_{i,j}} + \omega_{p,q} \tag{8}$$

Here $g_{p,q}^H v_q \sqrt{q_{p,q} t_{p,q}}$ is desired signal which we want to retrieve while $g_{p,q}^H v_q \sum_{i=1}^{p-1} \sqrt{q_{i,q} t_{i,q}} + g_{p,q}^H v_q \sum_{i=p+1}^{|T_m|} \sqrt{q_{i,q} t_{i,q}}$ is a intra code interference signal and $g_{p,q}^H \sum_{j \neq q} \sum_{i=1}^{|T_j|} v_j \sqrt{q_{i,j} t_{i,j}}$ is an inter code interference signal plus $\omega_{p,q}$ is a noise signal. Where $t_{p,q}$ and $q_{p,q}$ denotes the transmitted signal and power for pth user in qth code, and $\omega_{p,q}$ denotes the noise succeeding the distribution $\mathcal{L}\left(\mathcal{M}\left(0, \sigma^2\right)\right)$. It is noted that, in the existing OMA based MIMO system, only a user is supported in each chosen code, for example $|T_m| = 1\left(m = 1, 2, 3, \ldots, K_{RF}\right)$, while $|T_m| = \left(m = 1, 2, 3, ..., \ldots, K_{RF}\right)$ is larger than the proposed MUSA-NOMA system.

In (8), the first part of the equation is the required signal, the remaining terms are interferences and noise. In (8), the term $\{v_q\}_{q=1}^{K_{RF}}$ is the precoding vector is designed carefully to detain inter-code interferences. Intra-code interferences initiated by superposition coding in NOMA are suppressed by the functioning of SIC corresponding to the enhancing order of corresponding channel gains. For example, pth user in qth code removes the interferences from the kth user (for all $k > p$) in the qth code by performing SIC. By using NOMA in beam-space MIMO, two or more users are served simultaneously in each chain, and thus a total number of supported users will be bigger than the number of chains, i.e., $N \geq K_{RF}$. However, in the proposed MUSA-NOMA system, along with intra-code interferences, users agonize over interferences from other chains. Thus, the combination of NOMA and beam space MIMO is not capable to assure reliable accomplishment to decrease interferences by amplifying the attainable sum rate.

4 Main blocks of the proposed model

In this proposed model, the system consists of three main blocks, transmitter side, channel side, and receiver side as explained in Fig. 4.

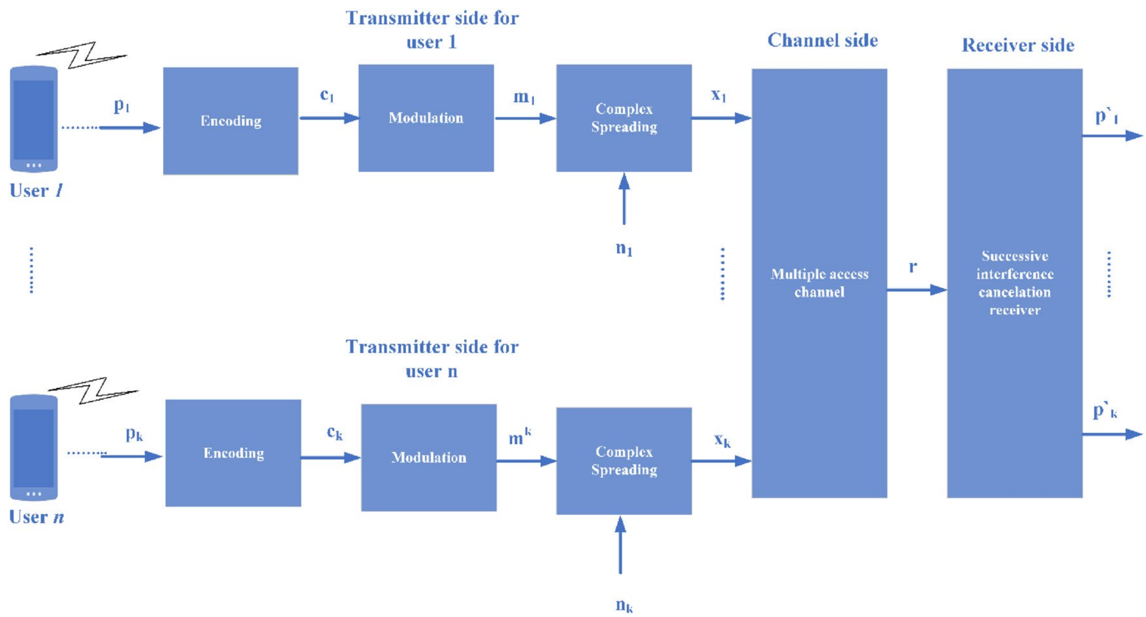


Fig. 4 Main blocks of the proposed system model

4.1 Transmitter side

On the transmitter side, we have multiple main blocks like encoding, modulation, and complex spreading. First, take data from users, there are multiple users like.

$$[U_1, U_2, U_3, U_4, U_5 \dots U_K] \tag{9}$$

Where U shows the number of users. When users from U_1 to U_k send data simultaneously, the transmitter side convert this data into bits. On the transmitter side, the following three main blocks encoding, modulation, and complex spreading are used. The communication procedure starts with the sender who convey a message to the receiver. Encoding is the underlying procedure done by a source in the communication procedure. In encoding, the planned message move into a securable and transferable way that can see distinctly by the receiver. At that point, the encoded message send to the receiver using a channel that is alluded to as a medium. The channel is responsible to convey the message from the source to the receiver. On the transmitter side of the proposed model, k users' data bits p_k that encoded by an encoder; the code rate of the encoder is T, which producing the coded bits

$$c_k = \left[c_k(1), \dots, c_k(N)\right] \tag{10}$$

Here N is the total length of the coded bits c_k. The reason for a communication system is to convey a communication administration from a data source to a data sink. Regularly, the source and receiver are spatially isolated, and a transmission channel is expected to fill in a connection between the imparting elements. Contingent upon the specific transmission medium, the first messages as discharged by the source are mapped onto an alternate configuration or sign portrayal before the real transmission with the end goal that it suits the channel characteristics. This activity is performed at the transmitting end of a communication system and is alluded to as modulation. At the receiver, a switch activity happens, which is called detection or demodulation targeting recouping the first message in our system mode, Quadrature Phase Shift Keying (QPSK) is used. After the modulation of the signal, the data is spreading with a composite spreading code X_K of shortest length, P is passed out to acquire the following spread symbols that are $x_k = n_k * m_k = [X_K(1), \dots, X_K(NL/log2(x))]$ to be transmitted. Here coded bits length is N. We get spread symbols X_K by L. After getting complex spreading symbols we will send these symbols via AWGN channel.

4.2 Receiver side

Now our received signal is

$$r = \sum_{k=1}^{K} x_k h_k + \omega \tag{11}$$

where x_k is the composite transmitted signal, h_k is the channel coefficient matrix,ω is a noise model that has a complex assessment selected from a Gaussian Distribution with zero variance σ^2 and mean. At the side of the receiver, SIC is used to receive the signal then it will cancel the interference of the signal.

5 Working of SIC

If we take a single user signal to retrieve its data, then we assume that

$$r' = xh + \omega \tag{12}$$

A way to deal with exposure of spatially multiplexed signals in SIC expecting that are the spatially multiplexed independently coded signals for example (multi-code words transmission). In a SIC receiver, the signal that is spatially multiplexed is at first demodulated and then decoded that signal. For effective decoding, the information is re-encoded and then compared with the received signals. Another spatially multiplexed signal would be able to be de-modulated and de-coded in the absenteeism of interfering since the 1st received signal. The comparison would be preferably improving the signal to interference. The strategy is repeated up until all spatially multiplexed signals remain de-modulated and de-coded. Here it is noticed that the primary signals could be decoded utilizing that methodology which is dependent upon high interference stages that are compared with those signals that are decoded later simultaneously. Consequently, the ith signal will be de-coded which must be more powerfully coded than the $(i + 1)$th signal. This signal can be accomplished in multiple codeword transmission employing applying changed modulation plus coding arrangements to various spatially multiplexed signals. A system model of SIC is shown in Fig. 5.

To retrieve the signal of every user at the receiver, h^{-1} is multiplied by (12) for the purpose to get the transmitted signals $\acute{x}$

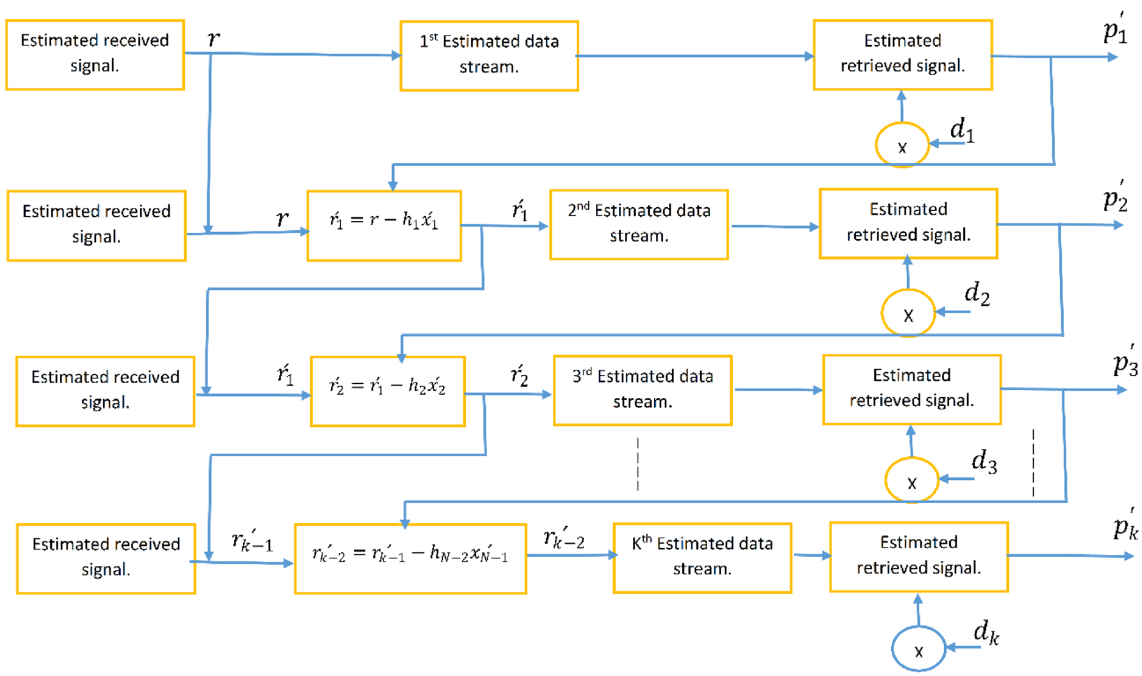


Fig. 5 A conventional OMA system

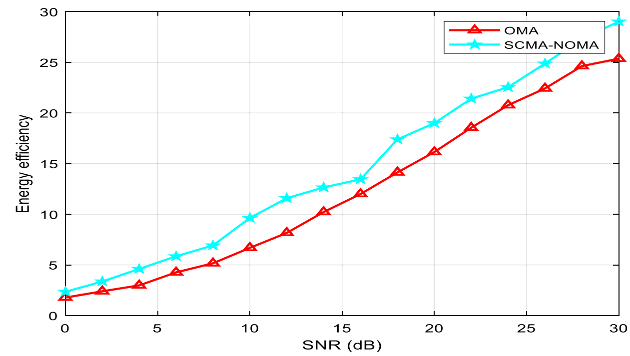


Fig. 6 Energy efficiency comparison of OMA and SCMA-NOMA

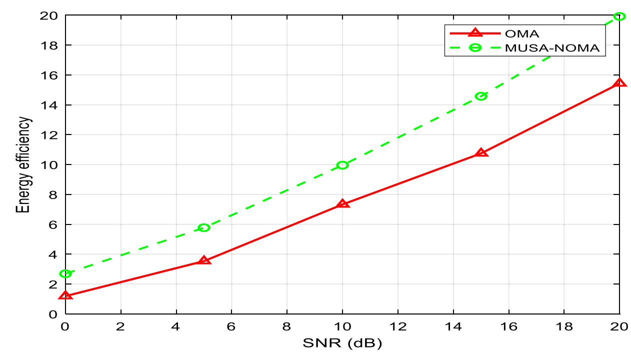


Fig. 7 Energy efficiency comparison of OMA and MUSA-NOMA

$$h^{-1}\,r' = \overset{'}{x} + \overset{'}{\omega} \tag{13}$$

$$x' = h^{-1}\,r' - \overset{'}{\omega} \tag{14}$$

where $\overset{'}{\omega}$ is ω multiplied by h^{-1} it is not easy to compute the term $h^{-1}r'$, because the inverse only exists for a square matrix. h must be a square matrix stand for the same number for both transmitting antennas and receiving antennas. If the matrix is not square, then the inverse is $(h^H h)^{-1} h^H$ as

$$W_{ZF} = \left(h^H h\right)^{-1} h^H \tag{15}$$

where $(.)^H$ Means the Hermitian transpose operation, W_{ZF} is the weight matrix corresponding to the zero-forcing method. For the Minimum Mean Square Error (MMSE) technique, the minimum Mean Square Error (MSE) weight matrix is given as

$$W_{MSE} = (h^H h + \sigma^2 I)^{-1} h^H \tag{16}$$

Here variance of the received antenna is σ^2 and the identity matrix is I. Now (16) will become as.

$$x' = W_{MSE} r' \tag{17}$$

6 Results and discussion

In this section, results are discussed of OMA and different NOMA techniques. The results are also compared in terms of energy efficiency and spectral efficiency.

6.1 Energy efficiency

The results for energy efficiency of two contemporary NOMA techniques, SCMA-NOMA and MUSA-NOMA are compared with the OMA technique. In the first phase, SCMA-NOMA result is compared with the OMA in terms of energy efficiency and

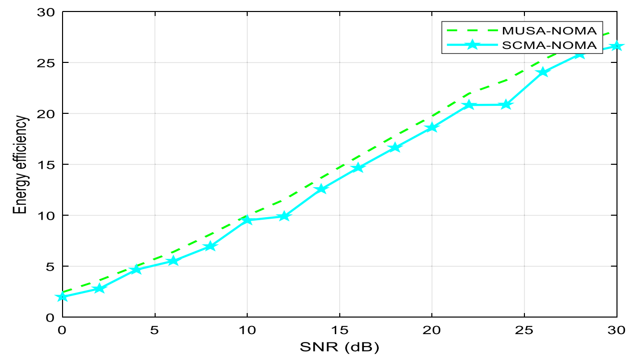


Fig. 8 Energy efficiency comparison of MUSA-NOMA and SCMA-NOMA

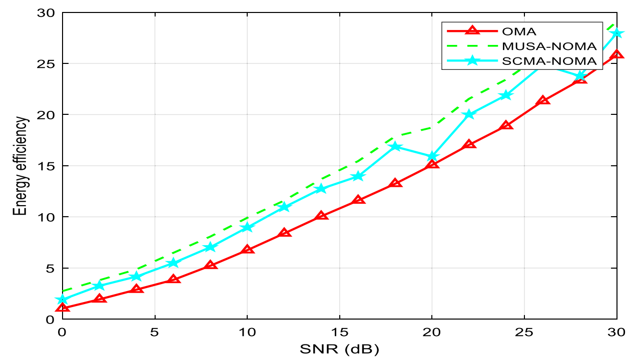


Fig. 9 Energy efficiency comparison of MUSA-NOMA, SCMA-NOMA and OMA

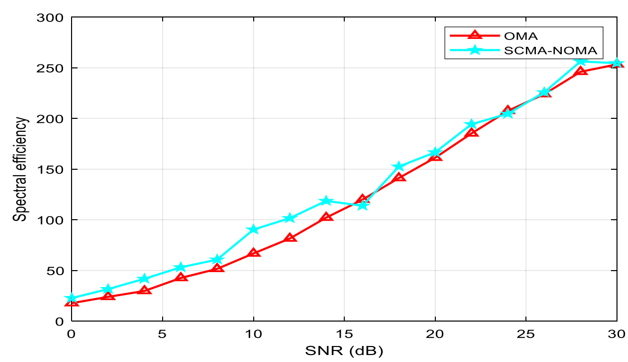


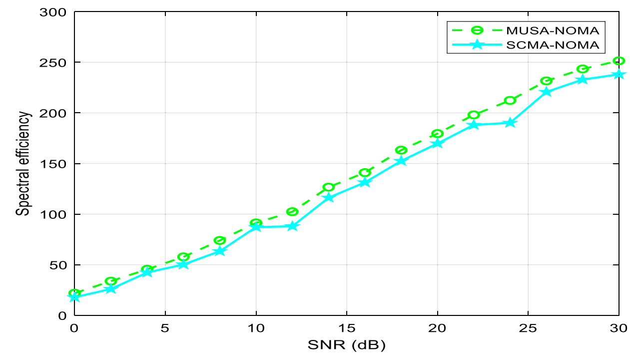
Fig. 10 Spectral efficiency comparison of OMA and SCMA-NOMA

Signal-to-Noise Ratio (SNR) as explained in Fig. 6. The energy efficiency of SCMA is better due to codebooks. Six codebooks' codes are generated and then each user-mapped data stream at each codebook and transmits that data. Along with the comparison of SCMA-NOMA with the OMA, it is very important to compare MUSA-NOMA with the OMA. Therefore, in Fig. 7 results reveal that MUSA-NOMA is much better than OMA. Results are also compared in terms of energy efficiency and SNR. The proposed scheme MUSA-NOMA leads more than OMA. The proposed scheme MUSA-NOMA is also better than SCMA-NOMA as shown in Fig. 8. Overall results are also presented in Fig. 9, in which results of the proposed scheme MUSA-NOMA are compared with SCMA-NOMA and OMA. MUSA-NOMA has 12.8% more performance compared to SCMA-NOMA and 32% compared to OMA.

6.2 Spectral efficiency

The results of the spectral efficiency of SCMA-NOMA and OMA are compared, and results show that the spectral efficiency of SCMA-NOMA is much better than OMA as shown in Fig. 10. The spectral efficiency of SCMA-NOMA is better due to the codebooks, here six codebooks are generated and then each user-mapped data stream at each codebook. The results of NOMA techniques that are SCMA and MUSA techniques are compared in this part. Comparability of the SNR versus spectral efficiency, the energy efficiency of both techniques, and iteration time to convergence of power allocation against spectral efficiency are described in this portion. The results in Fig. 11 show that MUSA always gives better results

Fig. 11 Spectral efficiency comparison SCMA-NOMA and MUSA-NOMA



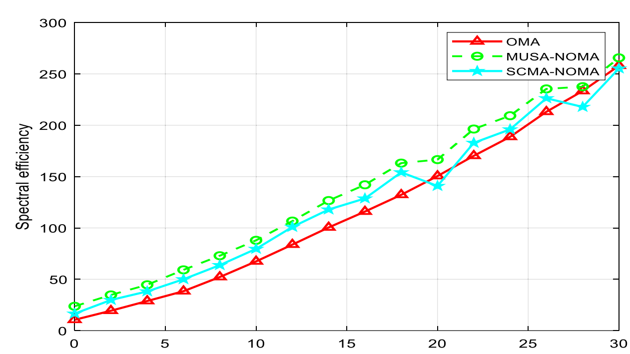
in spectral efficiency and energy efficiency. The spectral efficiency of MUSA is better than the results of SCMA. The spectral efficiency of MUSA is better due to maximum codes, MUSA has more codes as compared to SCMA. That's why MUSA can adjust the maximum user and it has maximum spectral efficiency. MUSA-NOMA are compared with SCMA-NOMA and OMA. MUSA-NOMA has 6.5% more performance compared to SCMA-NOMA and 18.5% compared to OMA.

6.3 Spectral and energy efficiency of MUSA-NOMA, SCMA-NOMA, and OMA

The spectral efficiency of MUSA-NOMA is much better than other techniques are shown in Fig. 12. Comparing the results of NOMA techniques that are SCMA, MUSA technique, and OMA are shown. The SNR versus spectral efficiency, the energy efficiency of all techniques, and iteration time to convergence of power allocation against spectral efficiency are also described. Results show that MUSA always gives better results in spectral efficiency and energy efficiency. In Fig. 9 it is evaluated that the spectral efficiency of MUSA is better than the SCMA technique and SCMA is better than the OMA technique, the spectral efficiency of MUSA is better due to maximum codes, and MUSA has more codes as compared to SCMA and OMA. And spectral efficiency of OMA is less than these two because OMA gets a full RF channel for each user. Therefore, the spectrum is more used than the other two techniques. MUSA can adjust the maximum user in the radio frequency chain, and it uses a minimum RF chain, therefore, MUSA has maximum spectral efficiency.

The description of the energy efficiency of (MUSA-NOMA), (SCMA-NOMA) and (OMA) techniques against SNR are presented In Fig. 12. The energy efficiency of MUSA-NOMA is better than SCMA-NOMA and the energy efficiency of SCMA-NOMA is better than the OMA. The energy efficiency of MUSA is better due to grant free access technique and maximum complex spreading code because multiple users can transmit data at the same radio frequency channel by using these codes without any performance loss and any permission for a channel. By using these techniques, energy can be saved for the devices also maximum users adjusted in the minimum spectrum; in this way, MUSA spectrum efficiency and energy efficiency are much better than OMA. The MUSA-NOMA has 12.8% more energy efficiency and 6.51% spectral efficiency compared to SCMA-NOMA and 32% more energy efficiency and 18.5% spectral efficiency compared to OMA.

Fig. 12 Spectral efficiency comparison SCMA-NOMA, MUSA-NOMA and OMA



7 Conclusion

In OMA, due to large number of users' transmission signaling overhead increases causing an increase in energy consumption by the devices and also consuming maximum spectral resources. A novel NOMA technique applied in this research to adjust the maximum number of users. The NOMA-based techniques are SCMA and MUSA and the results are compared in terms of spectral efficiency and energy efficiency. The proposed model and resulting discussion describe that better spectral efficiency and energy efficiency have been achieved. Results show that the maximum number of users can be adjusted in the minimum radio frequency chain. MUSA has better spectral and energy efficiency than all other techniques which are compared. The capability of MUSA-NOMA to adjust users in a single radio frequency chain is more than other existing techniques. The proposed scheme MUSA-NOMA has 12.8% more energy efficiency and 6.51% spectral efficiency compared to SCMA-NOMA and 32% more energy efficiency and 18.5% spectral efficiency compared to OMA.

Acknowledgements The authors would like to express their gratitude to the Ministry of Higher Education (MOHE) in Malaysia and Universiti Teknologi Malaysia (UTM) for providing financial support for this research through the HICoe Research Grant Scheme (R. J130000.7823.4J620).

Author contributions This research article is derived from the Master Thesis of M. Imran under the supervision of Omar Hayat and Razali Ngah.

Data availability All the data related to the paper is available on our local PC. It will be available on the University after acceptance by the esteem Journal.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Hayat O, Kaleem Z, Zafarullah M, Ngah R, Hashim SZM. Signaling overhead reduction techniques in device-to-device communications: paradigm for 5g and beyond. IEEE Access. 2021;9:11037–50.
2. Li J, Li X, Wang A, Ye N. Beamspace MIMO-NOMA for millimeter-wave broadcasting via full-duplex D2D communications. IEEE Trans Broadcast. 2020;66(2):545–54.
3. Majeed AA, Hburi I. Energy-efficient optimization of mm-wave communication using a novel approach of beamspace MIMO-NOMA. Wasit J Eng Sci. 2022;10(2):223–39.
4. Hayat O, Ngah R, Hashim SZM. Multi-user shared access (MUSA) procedure for device discovery in D2D communication. Telecommun Syst. 2021;76:291–7.
5. Liu P, Li Y, Cheng W, Zhang W, Gao X. Energy-efficient power allocation for millimeter wave beamspace MIMO-NOMA systems. IEEE Access. 2019;7:114582–92.
6. Mchangama A, Ayadi J, Jiménez VPG, Consoli A. MmWave massive MIMO small cells for 5G and beyond mobile networks: An overview. In: 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP). 2020. IEEE, pp. 1–6.
7. Liu Y, Song K, Luo Y, Han D, Zhang Y, Jin M. An enhanced beamspace channel estimation algorithm for wideband millimeter-wave massive MIMO systems. EURASIP J Adv Signal Proc. 2022;2022(1):115.
8. Akbar A, Jangsher S, Bhatti FA. NOMA and 5G emerging technologies: a survey on issues and solution techniques. Comput Netw. 2021;190: 107950.
9. Yuan Y, et al. NOMA for next-generation massive IoT: performance potential and technology directions. IEEE Commun Mag. 2021;59(7):115–21.
10. Asheer S, Kumar S. A comprehensive review of cooperative MIMO WSN: its challenges and the emerging technologies. Wirel Netw. 2021;27:1129–52.
11. Ghafoor U, Ali M, Khan HZ, Siddiqui AM, Naeem M. NOMA and future 5G & B5G wireless networks: a paradigm. J Netw Comput Appl. 2022;204: 103413.
12. Cetinkaya S. On the performance enhancement of beamspace mimo and non-orthogonal multiple access for future cellular networks. Tampa: University of South Florida; 2023.

13. Li H, Dong Y, Gong C, Wang X, Dai X. Gaussian message passing detection with constant front-haul signaling for cell-free massive MIMO. IEEE Trans Veh Technol. 2022;72(4):5395–400.
14. Kumar A, Ambigapathy S, Masud M, Jaha ES, Chakravarty S, Sengar K. An efficient hybrid PAPR reduction for 5G NOMA-FBMC waveforms. Comput Mater Continua. 2021;69(3):2967–81.
15. Kazeminia M, Mehrjoo M, Tomasin S. A D2D-based solution for MTC connectivity problem in NOMA-based cellular IoT networks: dynamic user grouping and resource allocation. Mobile Netw Appl. 2020;25:1998–2011.
16. Kwon D, Kim DK. Channel estimation overhead reduction scheme and its impact in IRS-assisted systems. ICT Express. 2023.
17. Ghous M, Hassan AK, Abbas ZH, Abbas G, Hussien A, Baker T. Cooperative power-domain NOMA systems: an overview. Sensors. 2022;22(24):9652.
18. El Tanab M, Hamouda W. Efficient resource allocation in fast-uplink grant for machine-type communications with NOMA. IEEE Internet Things J. 2022;9(18):18113–29.
19. Yuan W, Kalokidou V, Armour SM, Doufexi A, Beach MA. Application of non-orthogonal multiplexing to mmWave multi-user systems. In: 2017 IEEE 85th Vehicular Technology Conference (VTC Spring). 2017. IEEE, pp. 1–6.
20. Zhao Y, Xu W, Jin S. An minorization-maximization based hybrid precoding in NOMA-mMIMO. In: 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP). 2017. IEEE. pp. 1–6.
21. Hao W, Zeng M, Chu Z, Yang S. Energy-efficient power allocation in millimeter wave massive MIMO with non-orthogonal multiple access. IEEE Wirel Commun Lett. 2017;6(6):782–5.
22. Sharma S, Deka K, Bhatia V, Gupta A. Joint power-domain and SCMA-based NOMA system for downlink in 5G and beyond. IEEE Commun Lett. 2019;23(6):971–4.
23. Sayeed A, Brady J. Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies. In: 2013 IEEE global communications conference (GLOBECOM). 2013. IEEE. pp. 3679–3684.
24. Zeng Y, Zhang R. Millimeter wave MIMO with lens antenna array: a new path division multiplexing paradigm. IEEE Trans Commun. 2016;64(4):1557–71.
25. Rappaport TS, Ben-Dor E, Murdock JN, Qiao Y. 38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications. In: 2012 IEEE international conference on communications (ICC). 2012. IEEE. pp. 4568–4573.