MULTIPLE ACCESS TECHNIQUES FOR 5G NETWORKS
Mrs. Rinkoo Bhatia *1
*1 Assistant Professor, Amity University Madhya Pradesh, Gwalior, India

Abstract:
Fifth generation (5G) wireless networks face various challenges in order to support large-scale heterogeneous traffic and users, therefore new modulation and multiple access (MA) schemes are being developed to meet the changing demands. As this research space is ever increasing, it becomes more important to analyze the various approaches, therefore, in this article we present a comprehensive overview of the most promising Multiple Access schemes for 5G networks. Our article focuses on various types of non-orthogonal multiple access (NOMA) techniques. Specifically, we first introduce different types of modulation schemes, potential for OMA. We then pay close attention to various types of NOMA candidates, including power-domain NOMA, code-domain NOMA, and NOMA multiplexing in multiple domains. From this exploration, we can identify the opportunities and challenges that will have the most significant impacts on modulation and MA designs for 5G networks.

Keywords: 5G; Modulation; Non-Orthogonal Multiple Access.


1. Introduction

In past few years, extensive research has been carried out in fifth generation (5G) wireless networks. The three major categories of applications which should be supported by 5G networks according to the 3rd generation partnership project (3GPP) [1], [2], are enhanced mobile broadband (eMBB) [1], [2]; massive machine type communications (mMTC) [1], [2]; and ultra-reliable and low-latency communications (URLLC) [1], [2]. Also, an important service that is enhanced vehicle-to-everything (eV2X) communications should be provided by 5G networks [1]. All these applications need massive connectivity with high system throughput and high spectral efficiency (SE) and hence create significant challenges to the design of general 5G networks. To cope up with these requirements novel modulation and multiple access (MA) schemes are being explored. In Orthogonal frequency division multiplexing (OFDM) [3]–[5] being used in fourth generation (4G) networks, an appropriate cyclic prefix (CP), enables it to combat the delay spread of wireless channels with simple detection methods due to which makes it a popular solution for current broadband transmission. However, many new demands required for 5G networks cannot be met by traditional OFDM. For example, in the mMTC scenario [1], [2], different types of data are transmitted by sensor nodes asynchronously in narrow bands while OFDM needs different users to be fully synchronized, otherwise large interference will be there among adjacent subbands. To address the new challenges in 5G
networks, different types of modulation have been proposed, such as filtering, pulse shaping, and precoding to reduce the out-of-band (OOB) leakage of OFDM signals. Filtering [6]–[9] is considered to be the foremost method to reduce the OOB leakage and with an adequately designed filter, the leakage over the stop-band can be largely reduced. Pulse shaping can be considered as a type of subcarrier-based filtering that subsidises overlaps between subcarriers even inside the band of a single user, however, it usually has a long tail in time domain. Leakage can also be reduced by introducing precoding [4] to transmit data before OFDM modulation. For leakage reduction in OFDM signals in addition to the aforementioned approaches some new types of modulations have also been proposed specifically for 5G networks. For example, to counter high Doppler spread in eV2X scenarios, transmit data is modulated in the delay-Doppler domain [19]. The above modulations can be used with orthogonal multiple access (OMA) in 5G networks. OMA is core to all previous and current wireless networks; time-division multiple access (TDMA) and frequency-division multiple access (FDMA) are used in the second generation (2G) systems, code-division multiple access (CDMA) in the third generation (3G) systems, and orthogonal frequency division multiple access (OFDMA) in the 4G systems. In such systems, resource blocks are orthogonally divided in time, frequency, or code domains, there by minimalising interference among adjacent blocks and making signal detection relatively simpler. But, OMA supports limited numbers of users since there is limitations to the number of orthogonal resources blocks, which limits the SE and the capacity of contemporary networks. Different NOMA schemes have been proposed to support a large number of and different classes of users and applications in 5G networks. As an alternative to OMA, NOMA introduces a new dimension by performing multiplexing within one of the classic time/frequency/code domains. In other words, NOMA can be regarded as an “addon”, which has the potential to be harmoniously integrated with existing MA techniques. The essence of NOMA is to use power and/or code domains in multiplexing for supporting more users in the same resource block. NOMA is of majorly three types: power-domain NOMA, code-domain NOMA, and NOMA multiplexing in multiple domains. The capacity of 5G networks can be improved significantly with NOMA as the limited spectrum resources can be fully utilized to support more users, even though extra interference and additional complexity is introduced at the receiver.

This article provides a comprehensive overview of multiplexing techniques, including modulation techniques in OMA and various types of NOMA schemes.

The rest of this article is organized as follows. Section II, discusses various NOMA schemes. Section III concludes this article.

2. NOMA

To support higher throughput and massive and heterogeneous connectivity for 5G networks, novel modulations can be adopted or directly NOMA can be used with effective interference mitigation and signal detection methods. Summary of the key features of NOMA are as follows:

1) Improved SE: NOMA exhibits a high SE, which is attributed to the fact that it allows each resource block (e.g., time/frequency/code) to be exploited by multiple users.

2) Ultra high connectivity: With the capability to support multiple users within one resource block, NOMA can potentially support massive connectivity for billions of smart devices.
This feature is quite essential for IoT scenarios with users that only require very low data rates but with massive number of users.

3) Relaxed channel feedback: In NOMA, perfect uplink CSI is not required at the base station (BS). Instead, only the received signal strength needs to be included in the channel feedback.

4) Low transmission latency: In the uplink of NOMA, there is no need to schedule requests from users to the BS, which is normally required in OMA schemes. As a result, a grant-free uplink transmission can be established in NOMA, which reduces the transmission latency drastically.

Existing NOMA schemes can be classified into three categories: power-domain NOMA, code-domain NOMA, and NOMA multiplexing in multiple domains. We will introduce them subsequently with emphasis on power-domain NOMA.

2.1. Power-Domain NOMA

Power-domain NOMA is considered as a promising MA scheme for 5G networks [7-9]. Specifically, a downlink version of NOMA, named multiuser superposition transmission (MUST), has been proposed for the 3GPP long-term evolution advanced (3GPP-LTE-A) networks [40] and has been shown that system capacity and user experiences can be improved by it. The basic principles of various power-domain NOMA related techniques, including power allocation in NOMA, multiple antenna based NOMA, and cooperative NOMA are explained in the following.

2.2. Basic Power-Domain NOMA

By differentiating them with different power levels multiple users within the same time/frequency/code resource block are supported by Power-domain NOMA. Unlike multiuser detection in CDMA or MIMO systems that have multiple observations at the receiver, power-domain NOMA usually only has one observation. Specifically, in the uplink transmission of NOMA, the signal received at the BS can be expressed as

\[ Y = \sum_{a=1}^{A} h_a \sqrt{p_a} x_a + n, \]

(1)

Where \( p_a \) and \( x_a \) are the transmit power and transmit symbols from the \( a \)th user, respectively, \( n \) refers to AWGN with variance \( \sigma^2 \), and the number of users sharing the same resource block is \( A \). The transmit power \( p_a \) for each individual user is carefully adjusted to facilitate SIC at the receiver, that is, to make sure users with stronger powers to be detected with high accuracy. At the receiver (the BS), first the user with the best CSI is decoded with invoking of SIC. Then the corresponding signal component is removed from the received signal. The SIC receiver works in the descending order of the signal strengths. Since users experience different channel conditions the transmit power levels of different NOMA users are usually different. If the first detected symbols are all correct, the received signal-to-interference-plus-noise ratio (SINR) of the \( a \)th NOMA user can be given by
The downlink transmission of NOMA for the two-user case is shown in Fig. 1. Here the users which share the same resource block are differentiated by different power levels with a total power constraint. Typically, the BS sends a superimposed signal containing the two signals for the two users. NOMA allocates less power for the users with better downlink CSI, to guarantee overall fairness and to utilize diversity in time/frequency/code domains. SIC is used for signal detection at the receiver. The user with higher transmit power, that is, the one with smaller downlink channel gain, is first decoded while treating another user’s signal as noise. Once the signal corresponding to the user with higher transmit power is detected and decoded, its signal component will be subtracted from the received signal to facilitate the detection of subsequent users. It should be noted that the first detected user suffers from the highest inter-user interference and also the detection error in the first user will pass to the other users, which is why we have to allocate sufficient power to the first user to be detected.

2.3. Power Allocation in NOMA

NOMA supports unequal transmission rates for users experiencing varying channel conditions by assigning them different transmit powers. Hence, the power allocation technique for different users is critical to power-domain NOMA. As mentioned, the SIC receiver works according to descending order of the signal strengths. Here more powers are allocated to the users with poor CSI. By doing so, interference from the users with good CSI is reduced significantly as less power are allocated to them and hence the detection accuracy at users with poor CSI can also be improved. As the power allocation in NOMA is based on the order of CSI, the cases with perfect and imperfect CSI are different and should be investigated separately, as in [47]. When perfect CSI is available, the optimization problem can be formulated to maximize the individual/sum rate while considering the fairness among different users. While with average CSI, the optimization problem can be formulated to minimize the maximum outage probability.

![Figure 1: Downlink power-domain non-orthogonal multiple access (NOMA)](image-url)
2.4. Multiple Antennas based NOMA

Multiple antenna techniques can provide an additional degree of freedom on the spatial domain, and bring further performance improvements to NOMA. Recently, multiple antenna based NOMA has attracted lots of attention [13], [15], [16]–[18]. Different from single-input-single-output (SISO) based NOMA, where the channels are normally represented by scalars, one of the research challenges in multiple antenna based NOMA comes from user ordering; as the channels are generally in form of vectors or matrices. Currently, the possible designs of multiple antenna based NOMA fall into two categories where one or multiple users are served by a single beam forming vector.

By allocating different users with different beams in the same resource block, the quality of service (QoS) of each user can be guaranteed in multiple antenna based NOMA systems forcing the beams to satisfy a predefined order.

2.5. Cooperative NOMA

In cellular networks, a cell-edge user usually experiences a weaker received signal power and lower data rates compared to those near the BS. Relaying and coordinated multipoint (CoMP) transmission (and reception) techniques have been widely employed to increase the transmission rates for cell-edge users [11]. The scenario with users transmitting at different rates naturally matches the application scenarios typical of NOMA.

The basic idea of relay-assisted NOMA is to use the users with the better CSI as the decode-and-forward (DF) or amplify-and-forward (AF) as relays to improve the transmission rates of the users with poor CSI. A cooperative NOMA model supporting M users with M time slots has been proposed in [19]. In the first time slot, the traditional non-cooperative NOMA scheme is conducted. In the second time slot, the user with the best CSI acts as the DF relay for the user with the second best CSI. In the following time slots, the user with the m-th best CSI works as the relay for the user with the subsequent worse CSI to improve the transmission rates.

CoMP transmission, where multiple BSs support cell-edge users together, is capable of improving the performance of cell-edge users.

2.6. Cognitive Radio Inspired NOMA

To address the spectrum scarcity issue in wireless communications, cognitive ratio (CR) has been proposed which allows unlicensed users to use licensed spectrum as long as they generate no (intolerable) interference to the licensed users. Spectrum sensing is an enabling technique for CR, and how to balance the sensing overhead and CR throughput has been investigated in [20].

If an unlicensed user is at a distance away from the licensed user that shares the same licensed spectrum, then the interference between them will be small in general. Therefore, CR in this case can be regarded as special NOMA in the location domain while the NOMA introduced in the previous sections is in power, code, or multiple domains. Device-to-device (D2D) communications, a popular research area recently, turn to a special CR if the two users in D2D
communications are regarded as a simple CR network. In addition to the location domain, CR can also exploit the spatial domain and the frequency-spatial domain.

**Code-Domain NOMA**

Code-domain NOMA can support multiple transmissions within the same time-frequency resource block by assigning different codes to different users. It has certain spreading gain and shaping gain with the cost of extra signal bandwidth in comparison with power-domain NOMA. Existing solutions to code-domain NOMA mainly include low-density spreading CDMA (LDS-CDMA), low-density spreading OFDM (LDS-OFDM) [23], and sparse code multiple access (SCMA), which are introduced as follows.

1) LDS-CDMA: LDS-CDMA is a novel type of CDMA. Its key feature is that a low-density signature, which has a similar form of the low-density parity-check (LDPC) matrix, is employed for the codebook construction. When the number of users is larger than that of samples per symbol period in the conventional CDMA, MAI is inevitable and optimal multiuser detection is extremely complex. However, due to the sparse structure of the signature in LDS-CDMA, a low-complexity near-optimal multiuser detection scheme, based on a message passing algorithm (MPA), can be applied in the detection of LDS-CDMA, which significantly improves performance.

2) LDS-OFDM: LDS-OFDM [22] has similar properties to LDS-CDMA, except that the output of the signature is mapped into the subcarriers of OFDM rather than the time samples in CDMA. Therefore, a low-complexity MPA detector can be adopted. Compared to LDS-CDMA, LDS-OFDM utilizes multicarrier transmission, which makes it fit for wideband channels. Further, the strong compatibility with OFDM makes it flexible in resource allocation [22].

3) SCMA: In SCMA, by applying a sparse code book similar to the signature matrix in LDS, a certain number of resource blocks can support more users through spreading. Although a part of the users share the same block, another block would be adopted to distinguish different users when collisions occur. Besides the sparse spreading, SCMA utilizes multi-dimensional constellations to reduce the receiver complexity and further improve the SE. Attributed to the multi-dimension property, the constellation in one resource block can be projected into its subspace. For example, a four-point QAM constellation can be projected to a three-point constellation. Even when two points collide in one resource block or to say one dimension, they can be distinguished in the other used blocks. Due to fewer constellation points, the receiver complexity can be reduced. Moreover, the constellation design can focus on improving the detection performance. For example, a design based on constellation rotation and interleaving has been proposed in, which is able to achieve better BER performance compared to the simple LDS-OFDM. Due to the sparse structure of the spreading matrix and the large minimum distance of the multi-dimensional constellation, the detection performance of SCMA becomes excellent even when the resource blocks are overloaded. At the receiver, MPA, which is usually adopted in the decoding of LDPC, is applied in the detection. Due to the sparsity, MPA could achieve near-optimal performance with a much lower complexity compared to the optimal maximum likelihood (ML) and the BCJR algorithms. However, the complexity is still relatively high for user devices. Hence, SCMA also considers clustering the users based on the CSI and allocating different powers to different clusters. When the transmit powers among different clusters vary, the
signals of different clusters can be detected by using SIC, which is similar to the power-domain NOMA. Within each cluster, different users can be distinguished by using MPA. As a result, the combination of SIC and MPA can reduce the complexity of the receiver significantly.

**NOMA Multiplexing in Multiple Domains**

Beyond multiplexing signals in power domain or code domain, some of solutions for NOMA have been proposed to multiplex in multiple domains, such as the power domain, the code domain, and the spatial domain, in order to support massive connectivity for 5G networks. In Section II.A.2, we discussed multiple antenna based NOMA, where NOMA multiplexed in the power and spatial domains. We now introduce another three types of typical NOMA schemes multiplexing in multiple domains: pattern division multiple access (PDMA), building block sparse-constellation based orthogonal multiple access (BOMA), and lattice partition multiple access (LPMA).

1) **PDMA:** In PDMA, non-orthogonal patterns are allocated to different users to perform multiplexing. These patterns are carefully designed in the multiple domains of code, power, and space, to gain the SIC-amenable property. In the presence of this property, the low-complexity SIC based MPA multiuser detection method with reliable performance can be designed to run at the receiver side. At the transmitter, similar to SCMA, the users in PDMA are also spread by a sparse signature matrix. The main difference is that the number of resource blocks occupied by each user in PDMA can vary. For example, seven users can be multiplexed within three resource blocks through the following signature matrix

\[
S = \begin{bmatrix}
1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \sqrt{3} & 0 \\
\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} & 0 & \sqrt{3} & 0 \\
1 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \sqrt{3}
\end{bmatrix}
\]  

(4)

By utilizing the sparse signature matrix, PDMA can increase the system capacity through overloading. Moreover, users can also be multiplexed in other domains, such as power and space. In the same resource block, users can be distinguished by different powers as the power-domain NOMA or different precoders if MIMO is applied. At the receiver side, similar to SCMA, MPA can be adopted in detection due to the sparsity of the signature matrix. When different clusters of users are multiplexed in power and space domains, MPA-SIC can be applied. The detection of users that are multiplexed in the same signature matrix is based on the MPA, which can provide excellent performance. Among different clusters in the power and space domains, SIC can be utilized to reduce the complexity. Besides, a turbo structure can be adopted to combine the detector with the decoder to further improve the performance.
2) BOMA: This technique attaches the information from a user with good CSI to the symbols of a user with poor CSI. Thus the capacity of a multiuser system is increased significantly. As shown in Fig. 2, in order to achieve the same BER performance as a user with good CSI, the user with poor CSI should apply a coarse constellation with a large minimum distance. Hence, the small building block that contains the data of the user with good CSI can be tiled in the constellation of the user with poor CSI. For the user with poor CSI, the center of the building block can be regarded as the constellation point and the tiled building block can be regarded as interference. When the size of the building block is much smaller than the minimum distance of the coarse constellation, the degradation of detection performance becomes minimal. Since the user with good CSI can detect the points in its own constellation, it can also detect all points in the tiled building block constellation and decode the bits from itself.

![Figure 2: Building block sparse-constellation based orthogonal multiple access (BOMA)](image)

Constellation of stronger user, 64QAM

The structure of BOMA is simple and similar to that adopted in current 4G systems. Only minor software changes are required so that BOMA can be easily implemented with the compatibility to massive MIMO, high frequency bands, and other requirements of 5G systems. Besides, BOMA needs no complex power allocation and SIC receiver that are necessary for other NOMA schemes.

3) LPMA: In LPMA, the power domain and code domain are combined to multiplex users. Similar to power multiplexing in power-domain NOMA, the code in LPMA implements a multilevel lattice that allocates different code levels to users with different CSI. Several types of codes can be adopted, such as Construction A and Construction D. For users with the poor CSI, the allocated codes have larger minimum distance that can improve...
detection performance. For users with better CSI, the allocated codes are with smaller minimum distance without degrading detection performance. At the receiver, a SIC decoder is adopted, which is similar to that found in power-domain NOMA.

Besides the code domain multiplexing, LPMA also adopts power multiplexing to enhance those users with poor CSI. With the aid of two degrees of freedom in the multiplexing, the design of LPMA becomes more flexible in comparison with power-domain NOMA. Even if a pair of users has similar CSI, they can still be multiplexed by adjusting the allocated code levels and power levels, therefore the complex user clustering mechanisms adopted in the power-domain NOMA schemes are not required in LPMA.

3. Conclusions

In this article, we provide a comprehensive survey covering the major promising candidates for multiple access (MA) in fifth generation (5G) networks. Non-orthogonal MA is a promising approach that marks a deviation from the previous generations of wireless networks. By utilizing non-orthogonality, it has been convincingly shown that 5G networks will be able to provide enhanced throughput and massive connectivity with improved spectral efficiency.

References


*Corresponding author.

E-mail address: rbhatia@gwa.amity.edu