

Design Framework of Unsourced Multiple Access for 6G Massive IoT

Chunlin Yan^{2*}, Siying Lyu¹, Sen Wang¹, Yuhong Huang¹, Xiaodong Xu^{2,3}

¹ China Mobile Communication Research Institute, Beijing 100053, China

² ZGC Institute of Ubiquitous-X Innovation and Applications, Beijing 100080, China

³ State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

* The corresponding author, email: yanchunlin@zgc-xnet.com

Cite as: C. Yan, S. Lyu, *et al.*, “Design framework of unsourced multiple access for 6g massive iot,” *China Communications*, 2024, vol. 21, no. 1, pp. 1-12. **DOI:** 10.23919/JCC.fa.2023-0357.202401

Abstract: In this paper, ambient IoT is used as a typical use case of massive connections for the sixth generation (6G) mobile communications where we derive the performance requirements to facilitate the evaluation of technical solutions. A rather complete design of unsourced multiple access is proposed in which two key parts: a compressed sensing module for active user detection, and a sparse interleaver-division multiple access (SIDMA) module are simulated side by side on a same platform at balanced signal to noise ratio (SNR) operating points. With a proper combination of compressed sensing matrix, a convolutional encoder, receiver algorithms, the simulated performance results appear superior to the state-of-the-art benchmark, yet with relatively less complicated processing.

Keywords: channel coding; compressed sensing; massive Internet-of-Things (IoT); sparse interleaver-division multiple access (SIDMA); the sixth generation (6G) mobile communications; unsourced multiple access

I. INTRODUCTION

Multiple access is one of the core technologies of wireless communications and has enormous impact on

the system capacity and network performance. From the first generation (1G) to the fifth generation (5G) mobile wireless communication systems, orthogonal multiple access (OMA) has been the primary access scheme especially for enhanced mobile broadband (eMBB) in 4G and 5G systems with large package size. For eMBB, control signalling such as scheduling grant can facilitate the link adaptation which is crucial for the data throughput. Control signaling also includes various messages exchanged during the initial access to wake up terminals from idle state to connected state. Due to the small percentage of control information compared to the large packet size of eMBB and the continuous traffic flows, the benefit of control signaling overwhelms its overhead. Hence, OMA works reasonably well for eMBB where the number of connected terminals is not large in general.

For the sixth generation (6G) networks, it is envisioned that the number of connections for massive Internet-of-Things (IoT) will be 10 – 100 times more than that of 5G[1]. The traffics of massive IoT are mainly small data packages sporadically generated. One typical scenario is the Ambient IoT currently being studied in the 3rd Generation Partnership Project (3GPP)[2]. The traditional grant-based transmission of OMA is no longer suitable due to the excessive control overhead, access delays and device power consumption associated with the multi-step ran-

Received: Jul. 10, 2023

Revised: Nov. 16, 2023

Editor: Yongpeng Wu

dom access procedure and link adaptation. Therefore, a lightly-controlled transmission without connection-setup or dynamic link adaptation is desirable.

To reduce signaling overhead, delay and terminal consumption, grant-free transmission was introduced which has two types: 1) contention-free where the data is transmitted using pre-allocated signatures such as spreading sequences, interleaver patterns, scrambling codes, preamble, etc; 2) contention-based where signatures are randomly selected by users. In the study on 5G multiple access[3][4], the contention-free type of grant-free was studied more thoroughly.

Considering the much denser connections of 6G massive IoT, contention-based grant-free seems more suitable as it can more efficiently handle the signature collision problem. Recently there has been some breakthrough in information-theoretic analysis of contention-based transmission which is also called unsourced multiple access [5][6]. Drastically different from the traditional Shannon formula, Polyanskiy's team at Massachusetts Institute of Technology (MIT) proved that the capacity bound of unsourced multiple access is much higher than that of OMA [7]. Their theoretical work also reveals the potential of joint processing across multiple users for short-length blocks, which cannot be achieved by OMA where each user's data is individually processed.

Inspired by the finding in [7], several unsourced multiple access schemes were proposed for massive IoT scenarios [8–10]. The BCH/forward error correction (FEC) cascaded scheme [11] can tackle the collision and noise problems in Gaussian multiple access channel (GMAC) channel by using an BCH outer code and an inner forward error connection (FEC) code, respectively. In tree search-based compressed sensing (CS) approaches [12], the transmission code word is mapped to columns of CS matrix, e.g., direct mapping of information bits to the symbol vector of multiple access codes. In [13] a sparse Kronecker-product coding scheme is proposed for unsourced multiple access. The data of each active user is encoded as the Kronecker product of two component codewords with one being sparse and the other being forward-error-correction coded. Its performance is very close to the random coding bound. In [14] a data only transmission scheme was proposed where a blind equalization performs the channel estimation without relying on reference signals. In CS preamble coding

+ sparse interleaver-division multiple access (IDMA) coding scheme [15], CS is employed for active user detection and channel estimation, while an iterative receiver is used for multi-user detection.

In sparse code multiple access (SCMA) [16], zeros are inserted in the codewords to reduce the cross-user interference and the complexity of maximum likelihood (ML) receiver. In [17], Reed-Muller codes based CS was proposed where multi-slot transmission strategy is used for random access. In [18][19] low-density parity check (LDPC) code is employed as channel coding scheme for sparse-IDMA where the multi-user detector is of belief propagation (BP) type. In addition, multi-slot transmission is used to reduce the computational complexity. Compared to traditional IDMA which typically allows 10 – 20 users to transmit simultaneously, sparse-IDMA can support much larger scale of superposition transmissions, e.g., 200 – 300 users. Also, sparse-IDMA relies on channel code optimization to improve the performance, whereas in 5G IDMA study the channel codes were limited to New Radio (NR) LDPC.

The main contributions are summarized below.

1. Performance requirements are derived from a typical application scenario of ambient IoT which can guide key parameter setting of unsourced multiple access, such as information block length, code rate of channel encoder, number of repetitions, size of CS matrix, etc.
2. A complete design is proposed where the CS and multi-user detection work together, e.g., the results of CS for active user detection and channel estimation would be used for multi-user detection where collision probability is also considered.
3. Elementary signal estimator (ESE), instead of joint detection with belief-propagation, is used for multi-user detection of sparse-IDMA. In addition, smaller size of FFT-type compressed sensing matrix is used, resulting in better estimation of channel and active user detection. Hence, an overall performance closer to the capacity bound for unsourced multiple access is achieved, yet with reduced computational complexity.
4. The impact of channel coding scheme on the performance of sparse-IDMA is assessed via evolution functions. Compared to LDPC, the potential of convolutional codes in sparse-IDMA is clearly shown and verified by the simulation results.

5. Realistic channel estimation and active user detection are carried out for sparse-IDMA and the performance degradation compared to the ideal channel estimation is shown.

II. THE REQUIREMENTS FROM AMBIENT IOT

Originally from Radio frequency identification (RFID) technology, ambient IoT requires no power or very little power, and may fulfill the vision of “100 billion interconnection”. Ambient IoT will adopt a cellular architecture that supports a larger scale expansion of system capabilities. The study on deployment scenarios of ambient IoT is a hot topic in 3GPP R19 System Aspect Working Group 1 (SA1). A total of 32 use cases have been identified, including typical indoor cases such as warehousing, medical, manufacturing retail, home furnishings, and outdoor cases such as power, supply chain distribution, airport, municipal, and animal husbandry. Based on the characteristics and performance requirements, these use cases can be grouped into four categories: inventory, environmental sensing, terminal positioning, and control commands. Among them, the inventory category is the most challenging in terms of connection density and thus will be the focus in the following analysis. Some key performance requirements of inventory category are listed in Table 1.

According to Table 1, the communication range of an indoor reader is 30 m. In a cellular set up, it means that the cell radius is about 30 m. Hence, the coverage area of a cell (or a reader) is about 2800 square meters. Considering 1.5 million devices per square kilometer, there are roughly 4200 devices in an cell. Although not all of these 4200 devices would transmit simultaneously, up to 100 devices of them would access the network per second when the reader should be able to process in one second. In a typical indoor inventory scenario, the floor area would be around 600,000 square meters. This implies that over two hundred readers (cells) would be deployed where significant inter-cell interference is expected. To overcome this issue, cell-free access can be considered in which each cell (reader) would try to decode devices in neighbor cells. A rule of thumb is that contributions of 3 times of devices in a cell would count for most of uplink received power. Hence, each cell (reader) roughly needs

Table 1. Key performance requirements.

Parameter	Value
Max transmission latency	1 second
Communication range	30 m
Data rate	< 1 kbps
Message size	< 256 bits
Device density	1.5 million per km ²
Device speed	< 6 km/h
Peak reading rate	100 tags per second

to decode up to 300 devices’ data per second.

Similar to RFID, ambient IoT needs to have good coverage. While 30 m does not seem long, the requirement can be quite challenging considering the extremely low transmit power to maximize the battery life. That is why the typical deployment of ambient IoT is in low band, e.g., 800 MHz to ensure small pathloss, together with narrow bandwidth, e.g, smaller than 200 kHz to ensure the enough power spectral density. Because of the low data rate (e.g., less than 1 kbps), binary phase shift keying (BPSK) or quadrature phase shift keying (QPSK) would be sufficient, allowing more efficient amplifiers and better coverage.

Low-band operation also means limited number of antennas on a reader, e.g., 2 or 4 may be typical. Hence, antenna technology would play minor role in improving the performance of ambient IoT, as opposed to eMBB operating in higher bands. It is also seen from Table 1 that the mobility speed of devices in inventory category is less than 6 km/h. This implies that at 800 MHz operating band, the small-scale fading would be slow-varying.

III. AN UNSOURCED SPARSE SUPERPOSITION SCHEME FOR AMBIENT IOT

3.1 Basic Building Blocks of Unsourced Multiple Access

Figure 1 and Figure 2 show the transmitter (of a user) and the multi-user receiver, respectively, for unsourced multiple access using CS and sparse-IDMA. Let K_{tot} and K_a be the total number of users and the number of active users, respectively, where $K_a \ll K_{tot}$. In light of the performance requirements discussed in previous section, $K_{tot} \leq 4200 \times 3 = 12600$, and $K_a \leq 300$.

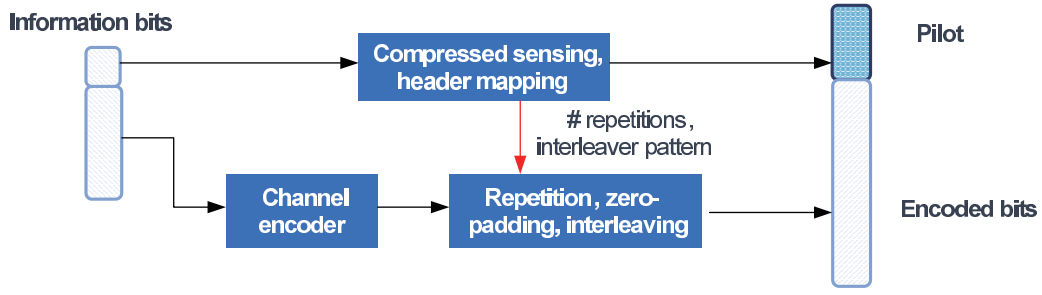


Figure 1. Transmitter for unsourced multiple access.

Each active user has B information bits to send. To facilitate the comparison with results in literature, in this paper we set $B = 100$ which is within the value of 256 bits in Table 1. A user's information bits are first partitioned into two segments e.g. $B = B_p + B_d$. The first segment is to generate preambles for CS of bit length B_p . The second segment is to generate modulation symbols for multi-user (MU) detection of bit length B_d . There are two approaches to transmit the second segment: single-slot vs. multi-slot. In the single-slot approach, all active users transmit simultaneously in a relatively long slot. In the multi-slot approach, each user would pick one of the short slots for transmission.

By using compressed sensing matrix and channel coding, repetition, zero padding, the information bits will be mapped to N_t channel uses (e.g. radio resources) for transmission. The value of N_t depends on the time duration and bandwidth. According to Table 1, the transmission duration is typically no greater than 1 second and the bandwidth is less than 200 kHz. In the simulation evaluation of this paper, $N_t = 30000$ is assumed. In the case of BPSK modulation and 200 kHz system bandwidth, the transmission duration is 0.15 second which meets the requirement. Based on the previous discussion, the channel can be assumed constant within 0.15 second when the mobility speed is below 6 km/h and the operating band is 800 MHz.

It should be emphasized that the encoding process here is independent of a user's identity (which is exactly the meaning of unsourced multiple access). The information vector $\mathbf{u}^{(k)}$ of the k -th active user, $k \in \{1, 2, \dots, K_a\}$ is split into two segments: $\mathbf{u}^{(k)} = [\mathbf{u}_p^{(k)}, \mathbf{u}_d^{(k)}]$. The first segment $\mathbf{u}_p^{(k)}$, whose length is B_p , is used to encode the CS part. The B_d -bit vector that consists the rest of the information is used to modulate the multi-user part, where $B_p \ll B_d$, and $B_d = B - B_p$. The lengths of the encoded CS

and MU parts equal N_p and N_d respectively, where $N_p + N_d = N_t$. N_t is the total message length for each user.

For the CS part of the encoding process, the sensing matrix $\mathbf{F}_{\text{partial}}$ of size $M \times N$ is listed below (elements in each column are normalized to meet the power constraint)

$$\mathbf{F}_{\text{partial}} = \begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \dots & \mathbf{a}^N \end{bmatrix}_{M \times N} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \dots & \dots & \dots & \dots \\ a_{M,1} & a_{M,2} & \dots & a_{M,N} \end{bmatrix}_{M \times N}. \quad (1)$$

Usually, $M < N$. The received signal of CS part can be represented as

$$\mathbf{y}_p = \mathbf{F}_{\text{partial}} \mathbf{x}_p + \mathbf{n}, \quad (2)$$

where $\mathbf{x}_p \in \mathbb{C}^{N \times 1}$ is a K_a -sparse signal vector whose non-zero entries correspond to the channels of active users to be estimated. In the case of GMAC, non-zero entries of $\mathbf{x}_p$ are 1. $\mathbf{y}_p \in \mathbb{C}^{M \times 1}$ is the received signal, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ is the complex Gaussian noise. $M = N_p$ which is essentially the spreading length of the preamble. $N = 2^{B_p}$ which is the total number of permutations of B_p bits. $\mathbf{F}_{\text{partial}}$ consists of M randomly selected and reordered rows of the unitary discrete Fourier Transform (DFT) matrix $\mathbf{F} \in \mathbb{C}^{N \times N}$. The entries of the sparse signal $\mathbf{x}_p$ is assumed to be independently identically distributed (i.i.d.), following Bernoulli-Gaussian distribution.

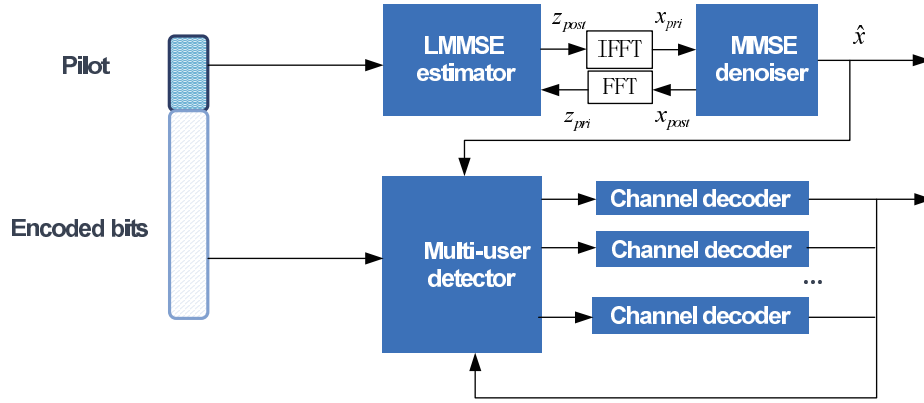


Figure 2. Receiver for unsourced multiple access.

The partial DFT matrix in Eq.(2) can be rewritten as

$$\mathbf{F}_{\text{partial}} = \mathbf{S}\mathbf{F}, \quad (3)$$

where $\mathbf{S}$ is a selection matrix consisting of M randomly selected and reordered rows of the $N \times N$ identity matrix. The above row selection involves a binary to decimal conversion and bit interleaving, e.g., the row of the decimal index corresponding to the interleaved binary sequence $\mathbf{u}_p^{(k)}$ would be selected. By defining

$$\mathbf{z} = \mathbf{F}\mathbf{x}_p, \quad (4)$$

together with Eq.(3), the system model in Eq.(2) can be rewritten as

$$\mathbf{y}_p = \mathbf{S}\mathbf{z} + \mathbf{n}. \quad (5)$$

The data part $\mathbf{x}_{d,k}$ is generated by encoding the second segment $\mathbf{u}_q^{(k)}$, followed by repetition, zero-padding and interleaving. More specifically, $\mathbf{u}_d^{(k)}$ is encoded with code rate of R to obtain a (B_d/R) -bit vector $\mathbf{v}_{d,k}$. Each bit in the vector $\mathbf{v}_{d,k}$ is then repeated several times, and then zero-padded to obtain a vector $\mathbf{x}_{d,k}$ of length N_d . The repetition and interleaving pattern of k -th user are determined by the first segment $\mathbf{u}_p^{(k)}$. Hence, the receiver can obtain the interleaver and repetition information about the k -th user after decoding the first segment $\mathbf{x}_p^k$. When K_a active users transmit simultaneously, the received signal in the data part can be represented as:

$$\mathbf{r}_d = \sum_{k=1}^{K_a} h_k \mathbf{x}_{d,k} + \mathbf{n}. \quad (6)$$

For CS receiver in Figure 2, turbo CS algorithm is employed for iterative detection. The algorithm contains two key modules: a minimum mean square error (MMSE) estimator and a MMSE denoiser [20–22]. When multiple receiver antennas are employed, signals from different antennas can be combined by using spatial MMSE or maximal ratio combination (MRC) operations. For the multi-user detection part (e.g., the second segment of data), another iterative algorithm is used where the extrinsic information propagates between multi-user signal estimator and channel decoder. More details will be elaborated in the next sub-section.

3.2 Signal Detection for CS

Besides MMSE estimator and MMSE denoiser, auxiliary modules in CS receiver include Fast Fourier Transform (FFT) and Inverse fast Fourier Transform (IFFT) for the signal transformation [20] as shown in the upper part of Figure 2. The Turbo CS algorithm can be described below.

CS matrix has two key parameters: M and N . M is the number of rows which corresponds to the physical resources for the preamble. Larger value of M means lower signal-to-noise ratio (SNR) operating point for CS. N is the number of columns in the sensing matrix, which is essentially the number of CS (or spreading) sequences to be selected by users. For unsourced multiple access, larger value of N means lower probability of signature collisions. From hardware implementation perspective, N cannot be too large. Otherwise, the memory requirement and the processing complexity would be infeasible in practical engineering. On

Algorithm 1. Turbo CS Algorithm for CS part.

1: Input: $\mathbf{S}, \mathbf{y}_p, M, N, \sigma^2$
2: Output: $\mathbf{x}^{\text{post}}$
3: Initialization: $\mathbf{z}^{\text{pri}} = 0, \mathbf{v}^{\text{pri}} = 1$
Module A: MMSE estimator
4: $\mathbf{z}^{\text{post}} = \mathbf{z}^{\text{pri}} + \frac{\mathbf{v}^{\text{pri}}}{\mathbf{v}^{\text{pri}} + \sigma^2} \mathbf{S}^H (\mathbf{y}_p - \mathbf{S} \mathbf{z}^{\text{pri}})$
5: $\mathbf{v}^{\text{post}} = \mathbf{v}^{\text{pri}} - \frac{M}{N} \frac{(\mathbf{v}^{\text{pri}})^2}{\mathbf{v}^{\text{pri}} + \sigma^2}$
6: $\mathbf{v}^{\text{ext}} = \left(\frac{1}{\mathbf{v}^{\text{post}}} - \frac{1}{\mathbf{v}^{\text{pri}}} \right)^{-1}$
7: $\mathbf{z}^{\text{ext}} = \mathbf{v}^{\text{ext}} \left(\frac{\mathbf{z}^{\text{post}}}{\mathbf{v}^{\text{post}}} - \frac{\mathbf{z}^{\text{pri}}}{\mathbf{v}^{\text{pri}}} \right)$
Module B: MMSE denoiser
8: $x_j^{\text{post}} = E[x_j | \mathbf{x}^{\text{pri}}], \forall j$
9: $v_j^{\text{post}} = \text{var}[x_j | \mathbf{x}^{\text{pri}}], \forall j$
10: $\mathbf{v}^{\text{pri}} = \mathbf{v}^{\text{ext}}$
11: $\mathbf{v}^{\text{ext}} = \left(\frac{1}{\mathbf{v}^{\text{post}}} - \frac{1}{\mathbf{v}^{\text{pri}}} \right)^{-1}$
12: $\mathbf{x}^{\text{ext}} = \mathbf{v}^{\text{ext}} \left(\frac{\mathbf{x}^{\text{post}}}{\mathbf{v}^{\text{post}}} - \frac{\mathbf{x}^{\text{pri}}}{\mathbf{v}^{\text{pri}}} \right)$
IFFT and FFT transformation
13: $\mathbf{x}^{\text{pri}} = \mathbf{F}^H \mathbf{z}^{\text{ext}}$
14: $\mathbf{z}^{\text{pri}} = \mathbf{F} \mathbf{x}^{\text{ext}}$
15: $\mathbf{v}^{\text{pri}} = \mathbf{v}^{\text{ext}}$

the other hand, N needs be sufficiently large to ensure low collision probability.

When $K_a = 300$ and $N_p = 2^{12}$, the probability of collision-free for each user can be calculated by $(1 - \frac{1}{N_p})^{K_a - 1} = 0.9294$. This means the collision probability is about 0.07 which cannot meet the requirement of P_e less than 0.05. To reduce the collision probability, larger size of compressed sensing matrix is needed which makes the computational complexity less affordable. Thus new paradigm for compress sensing matrix design and its detection is an important research topic.

3.3 Multi-User Detection for Unsourced Multiple Access

Iterative detection is typically used for unsourced multiple access. Here ESE can be applied for multi-user detection in unsourced multiple access [23–25]. Figure 3 shows the transmitter and receiver for IDMA using ESE.

For simplicity, the subscript d is omitted in this subsection. Here $\mathbf{c}_k = \{c_k(j)\}$ and $\mathbf{x}_k = \{x_k(j)\}$ denote the bits before and after the interleaving. The elements of $c_k(j)$ and $x_k(j)$ are either 1 or -1. The received signal can be represented as $r(j) = \sum_{k=1}^{K_a} h_k x_k(j) +$

$n(j), j = 1, 2, \dots, J$. Here h_k is the channel coefficient of the k -th user, $n(j)$ is the additive white gaussian (AWGN) with zero-mean and variance $\sigma^2 = N_0/2$, and J is the length of the frame, e.g., $J = N_d$ in single-slot approach. For simplicity, GMAC channel is considered. Hence $h_k = 1$.

The received signal can be rewritten as [23]

$$r(j) = x_k(j) + \varsigma_k(j), \quad (7)$$

where

$$\varsigma_k(j) = \sum_{k' \neq k} x_{k'}(j) + n(j), \quad (8)$$

denotes the interference part. The mean and variance of $x_k(j)$, denoted as $E(x_k(j))$ and $\text{Var}(x_k(j))$, respectively, can be calculated according to the central limit theorem.

$$E(\varsigma_k(j)) = \sum_{k' \neq k} E(x_{k'}(j)), \quad (9)$$

$$\text{Var}(\varsigma_k(j)) = \sum_{k' \neq k} \text{Var}(x_{k'}(j)) + \sigma^2. \quad (10)$$

The initialization is done by setting $E(x_k(j)) = 0$ and $\text{Var}(x_k(j)) = 1$. The soft-input information for the channel decoder can be calculated as

$$e_{ESE}(x_k(j)) = \frac{2}{\text{Var}(\varsigma_k(j))} (r(j) - E(\varsigma_k(j))). \quad (11)$$

The extrinsic information can be calculated as

$$l_{in}(j) = e_{ESE}(x_k(j)) - \tilde{l}_{ESE}(x_k(j)), \quad (12)$$

where $\tilde{l}_{ESE}(x_k(j))$ is the soft-output of the channel decoder and can be initialized as $\tilde{l}_{ESE}(x_k(j)) = 0, j = 1, 2, \dots, J$. Based on the soft-output information, the a prior mean and variance of $x_k(j)$ can be written as

$$E(x_k(j)) = \tanh\left(\frac{\tilde{l}_{ESE}(x_k(j))}{2}\right), \forall k, j, \quad (13)$$

$$\text{Var}(x_k(j)) = 1 - (E(x_k(j)))^2. \quad (14)$$

This updated prior information will be used as the input in the next iteration. The above iterative detection

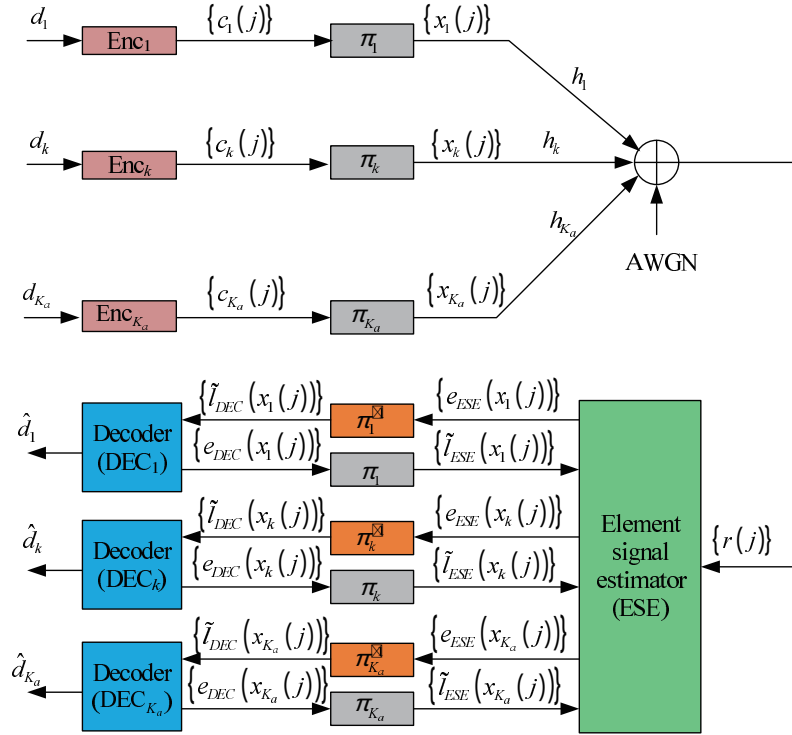


Figure 3. Transmitter and receiver for IDMA.

is summarized as below.

- 1) Calculation of the means and variances of interference by using Eq.(9) and Eq.(10).
- 2) Calculation of the soft information using Eq.(11).
- 3) Carrying out the decoding (DEC) function to generate $\tilde{l}_{ESE}(x_k(j))$.
- 4) Calculation of the means and variances of the user's signal using Eq.(13) and Eq.(14).

For the receiver with ESE + soft decoder, the performance can be analyzed by plotting the transfer functions between $f(\gamma)$ and $g(v)$ which can be calculated as below [26] for GMAC.

$$\gamma^t = \frac{1}{(K-1)v + \sigma^2} \equiv g(v^{t-1}), \quad (15)$$

$$v^t = f(\gamma^t). \quad (16)$$

Here K is the average number of users whose non-zero symbols occupying the same resource. v is the mean of $\text{Var}(x_k(j))$ and t is the index of iterations. γ is the signal-to-interference-plus-noise ratio at the output of the detector and also the post-SNR at the input to the decoder. For sparse-IDMA a large number of zeros are padded, thus the multi-user interference can be diluted sufficiently. For example, the average

number of users multiplexed on the same resource is around 6 for $K_a = 250$.

If the curves of $f(\gamma)$ and $g(v)$ do not cross, the iterative detection can converge. However, when there is crossing, the iterative detection will be stuck at the crossing point.

It is noted that in [15] a belief propagation (BP) type of multi-user detection is used for signal detection of K_a users which tends to have high computation complexity.

3.4 Channel Coding Aspect and Detection Complexity

As a very preliminary study, performances of sparse-IDMA with LDPC codes of 5G NR and convolutional codes (CC) have been evaluated. The rationale of considering convolutional codes is as follows. Unsourced multiple access is featured by short-length packets, e.g., 100 bits where convolutional codes may have advantages over random codes such as turbo or LDPC which typically rely on long interleavers or large parity check matrices to perform well. For unsourced multiple access, decoding and multi-user detection need to cooperate closely, e.g., their transfer functions need to

match each other, so that the iteration can converge to a satisfying point. Powerful channel codes for single-user case may not fare well in this case. Note that Polar codes are not considered in this paper due to the incompatibility to iterative receiver. Although soft-output decoding for Polar codes does exist, significant performance degradation is observed compared with conventional decoding schemes of Polar codes [27].

In [14] a joint detection is applied for the signal detection of sparse-IDMA whose complexity tends to be significantly higher than that of ESE. Taking the example of 28000 samples for data transmission, the complexity of the multi-user detection (compared to ESE) is proportional to the number of inner iterations (within BP) and the mean of the column weight. For $K_a = 250$ the mean of column weight (average number of users occupying the same resource) is around 6, and when the number of iterations is 5, the total complexity of BP for sparse-IDMA is about 30 times of that of ESE. When K_a is a small value, the complexity of sparse-IDMA using BP may be reduced (when the mean of column weight is 2), however, it is still ten times of that of ESE detector.

IV. PRELIMINARY EVALUATIONS

Some key simulation parameters are listed in Table 2. BPSK modulation is used. The channel is AWGN. The channel code rate $R = 1/2$ for both LDPC and convolutional codes. The code length (physical resources) of CS is $N_p = 2000$. The physical resources of multi-user detection is $N_d = 28000$. Hence, the total physical resources is $N_t = 30000$.

In Figure 4 the detection error probability vs. chip level SNR is illustrated for $K_a = 300$ by using Turbo CS algorithm with different values of FFT size. The required SNRs for detection error probability of 0.01 are -16.2, -12.8, and -9.5 dB for FFT size values of 4096, 8192 and 16384, respectively. Thus larger FFT size results in higher probability of detection error. By reducing the FFT size the required SNR can be reduced significantly. To improve the performance of sparse-IDMA, a CS matrix with a small FFT size is preferred.

In Figure 5, the evolution functions of LDPC and convolutional codes for sparse-IDMA are compared at $E_b/N_0 = 3$ dB. The convergence of iterative detection can be characterized by the interplay between

Table 2. Simulation parameters.

Parameters	Values
Channel	AWGN
Modulation	BPSK
Channel coding	LDPC, CC
g for CC	$[133, 171]_{oct}$
Repetition number	2
Number of information bits, $B = B_p + B_d$	$13+87=100$
Code rate, R	$1/2$
Resources for CS, N_p	2000
Resources for MU detection, N_d	28000
Information length for CS, B_p	13
Information length for MU detection, B_d	87
Number of terminals: K_a	$50 \sim 300$

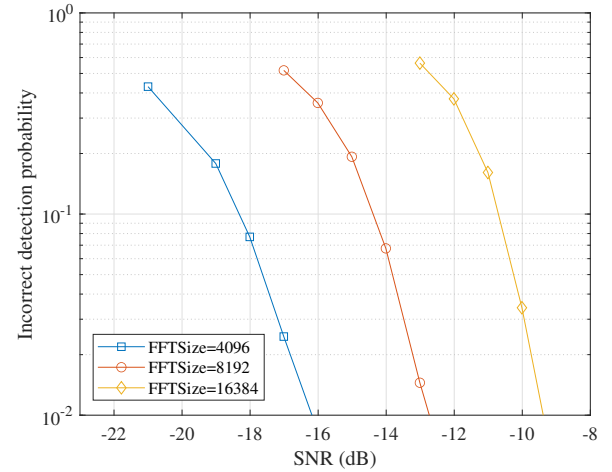


Figure 4. Performance of CS with different FFT sizes.

$f(\gamma)$ and $g(v)$. When K_a is 250 and 200, the value of K can be regarded as 6 and 4, respectively. No crossing is observed between $f(\gamma)$ for CC and $g(v)$ with $K = 6$. This implies that sparse-IDMA using convolutional codes can reach convergence at $\gamma = -1.77$ dB. However, crossing point is observed between $f(\gamma)$ for LDPC and $g(v)$ where v is about 1. This implies that sparse-IDMA by using LDPC cannot converge. By reducing the value of K from 6 to 4, sparse-IDMA using LDPC can reach convergence as Figure 5 shows.

Figure 6 shows the estimated channel for different cases. The left figure is for Case I and the right figure is for Case II. Parameters for these cases are illustrated in Table 3. Here SNR_1 is the SNR defined only by CS signals and SNR_2 is the SNR defined only by sparse-

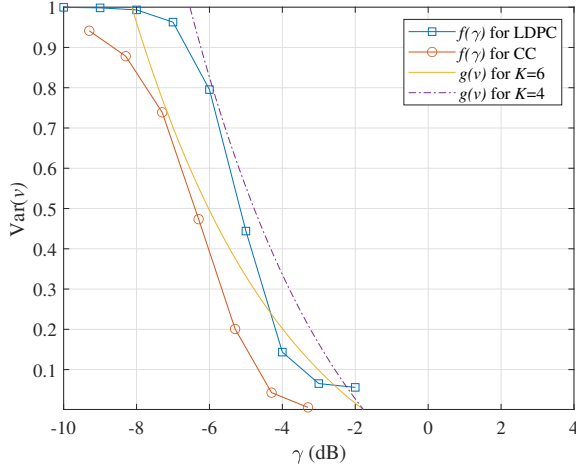


Figure 5. Evolution functions of LDPC codes based and convolutional codes based for sparse-IDMA system. $\frac{1}{\sigma^2}$ is set to -1.77 dB, corresponding to $E_b/N_0 = 3$ dB.

Table 3. Parameters for case I, II and III with FFT size 8192

Case I	Case II	Case III
$SNR_1 = -12.8$ dB, $SNR_2 = -2.3$ dB, realistic active user detection, realistic channel estimation	$SNR_1 = -10.8$ dB, $SNR_2 = -3.3$ dB, ideal active user detection, realistic channel estimation	$SNR_2 = -3.5$ dB, ideal active user detection, ideal channel estimation

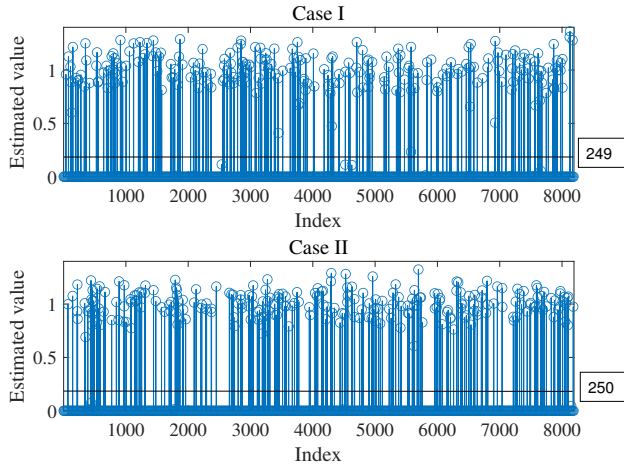


Figure 6. Real channel estimation and real active user detection for $K_a = 250$.

IDMA signals. For Case I, SNR_1 is not high. Thus the active user detection and channel estimation are not very accurate. The number of detected active users

is 249. And the smallest value of estimated channel is about 0.2. SNR_1 in Case II is higher, leading to the improved performance of active user detection and channel estimation. The number of detected active users is 250 and the smallest value of estimated channel is about 0.6.

Figure 7 shows the performance of multi-user interference vs. number of iterations in different cases. It can be seen that the estimated channel is not equal, which somehow facilitates the iterative detection, e.g., smaller number of iterations is needed compared with that of ideal channel estimation case. All cases can achieve convergence after 10 iterations. Compared with ideal channel estimation and ideal active user detection, realistic channel estimation and active user detection shows performance degradation about 1.2 dB in terms of SNR_2 .

Figure 8 shows the required E_b/N_0 for $P_e \leq 0.05$ with different values of K_a . The value of E_b/N_0 is calculated as

$$\frac{E_b}{N_0} = \min_{P_1, P_2, K_b} \frac{N_p P_1 + N'_d P_2}{2B}, \quad (17)$$

such that $P_r(\epsilon | P_1, P_2, K_b) < \epsilon$.

Here P_1 is the power for CS parts with length of N_p , and P_2 is the power for sparse-IDMA part with length of N'_d . Here N'_d is the length of non-zero values among N_d data. And we fix $K_b = \lceil 1.1K_a \rceil$ just as what is done in [15]. For simplicity we choose P_1 which achieves $P_e = 0.01$ for CS part and P_2 which achieves $P_e = 0.04$ for sparse-IDMA part.

It is found in Figure 8 that the performance of sparse-IDMA in [15] is worse than the simulation result of this paper because a smaller FFT size is applied and different channel coding is employed in our scheme. The performance of sparse-IDMA with CC is superior to that with LDPC because $f(snr)$ of CC is better than that of LDPC just as Figure 5 shows. The performance of sparse-IDMA and IDMA is evaluated in [15] where sparse-IDMA is preferred.

V. PRACTICAL ISSUES IN UNCOORDINATED / UNSOURCED TRANSMISSION

In the academic study on unsourced multiple access, hardware impairments are often neglected, such as frequency offset due to the drifting of the crystal oscilla-

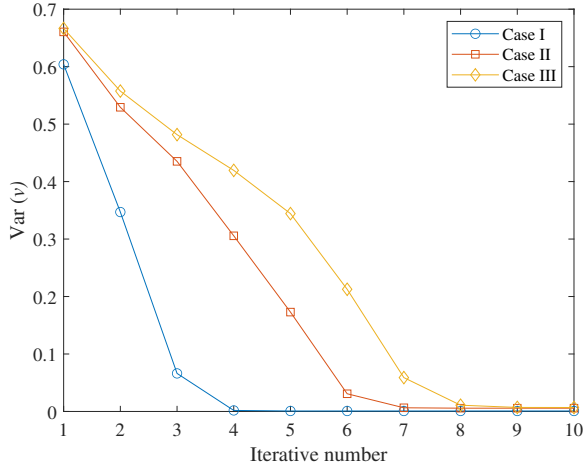


Figure 7. Multi-user interference vs. iterative number by using CC, $K_a = 250$.

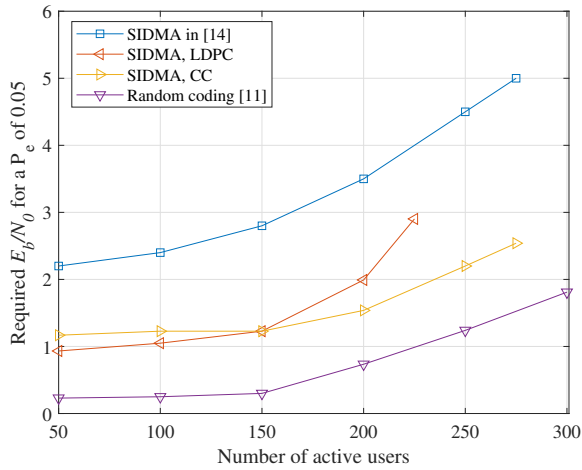


Figure 8. Required SNR vs. Number of active terminals.

tor. In unsourced multiple access the multi-user signals are superposed together. When there is frequency offset which can vary between different terminals, the estimation algorithms designed for single user case cannot be used. The severe multi-user interference makes the signal-to-noise-plus-interference very low, thus estimation of the frequency offset for multi-user is a challenging work. There is also multi-cell interference in practice which may leave some unsourced multiple access schemes not workable or performance significantly degraded. The impact of multi-cell interference on unsourced multiple access should be considered.

Multi-path fading channel should be considered as well. So far AWGN channel has been assumed in

many studies on unsourced multiple access, while in wireless communications multi-path fading channel exists. CS often assumes that the channel is constant within a code block. However, channel fading would make CS detection more challenging. Another difficulty is channel estimation over multi-path fading channel where several path coefficients should be considered.

Multiple-antenna can improve the performance greatly via transmitter-diversity or receiver-diversity. It may be difficult to equip multiple-antenna in transmitter due to the cost constraint and low band frequency applied. However, for base station multiple receiver antennas are possible. Multiple antenna at receiver can improve system performance, but, the computational complexity also increases. Thus how to make a balance is a research topic.

5G LDPC codes are binary codes. Non-binary LDPC is the extension of binary LDPC. Because non-binary LDPC has better performance than that of binary LDPC for small code length [28], and it is verified in this paper that channel coding has significant impact on the performance of unsourced multiple access, non-binary LDPC can be studied for unsourced multiple access in future.

Coverage and energy efficiency are important requirements for 6G applications. To extend the signal coverage or achieve high energy efficiency, low peak-to-average-power-ratio (PAPR) waveform is preferred. Candidate waveforms include single-carrier or constant envelope orthogonal frequency division multiplexing [29]. The later has a PAPR of 0 dB, leading to the wide coverage. Low PAPR waveform is one promising research topic for unsourced multiple access.

VI. CONCLUSION

In this paper the requirements of 6G massive IoT were proposed with ambient IoT as a typical use case. The stringent requirements called for unsourced multiple access where a framework of transmitter and receiver processing was discussed. Two key processes were elaborated and simulated jointly. One is called sparse-IDMA where a large number of zeros are added to reduce the multi-user interference after user-specific interleaver, and the other is CS for active user detection and channel estimation. With suitable combination of CS matrix, convolutional codes, Turbo CS

receiver, ESE algorithm, a superior performance was achieved, compared to the state-of-the-art proposals, yet with relatively less computation complexities. Future work was outlined which includes feature enhancements such as multi-antennas, channel coding and sensing matrix optimization, as well as practical design aspects such as frequency offset estimation, CS for multi-path fading channels, etc.

References

- [1] ITU-R, "Framework and overall objectives of the future development of imt for 2030 and beyond," International Telecommunication Union Radiocommunication Sector, *Tech. Rep.*, 2023.
- [2] 3GPP, "Study on ambient power-enabled internet of things (release 19)," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 22.840, 2023. https://www.3gpp.org/ftp/Specs/archive/22_series/22.840/.
- [3] 3GPP, "Study on non-orthogonal multiple access (noma) for nr," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.812, 2018. https://www.3gpp.org/ftp/Specs/archive/38_series/38.812/.
- [4] Y. Yuan, Z. Yuan, *et al.*, "5g non-orthogonal multiple access study in 3gpp," *IEEE Communications Magazine*, 2020, vol. 58, no. 7, pp. 90–96.
- [5] Y. Yuan, S. Wang, *et al.*, "Noma for next-generation massive iot: Performance potential and technology directions," *IEEE Communications Magazine*, 2021, vol. 59, no. 7, pp. 115–121.
- [6] Y. Li, J. Dai, *et al.*, "Unsourced multiple access for 6g massive machine type communications," *China Communications*, 2022, vol. 19, no. 3, pp. 70–87.
- [7] Y. Polyanskiy, "A perspective on massive random-access," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2523–2527.
- [8] X. Chen, T. Chen, *et al.*, "Capacity of gaussian many-access channels," *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 6, pp. 3516–3539.
- [9] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted aloha," *IEEE Transactions on Communications*, 2011, vol. 59, no. 2, pp. 477–487.
- [10] E. Paolini, C. Stefanovic, *et al.*, "Coded random access: Applying codes on graphs to design random access protocols," *IEEE Communications Magazine*, 2015, vol. 53, no. 6, pp. 144–150.
- [11] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access gaussian channel," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2528–2532.
- [12] V. K. Amalladinne, J. Chamberland, *et al.*, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Transactions on Information Theory*, 2020, vol. 66, no. 10, pp. 6509–6533.
- [13] Z. Han, X. Yuan, *et al.*, "Sparse kronecker-product coding for unsourced multiple access," *IEEE Wireless Communications Letters*, 2021, vol. 10, no. 10, pp. 2274–2278.
- [14] Z. Yuan, C. Yan, *et al.*, "Blind multiple user detection for grant-free msa without reference signal," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.
- [15] A. K. Pradhan, V. K. Amalladinne, *et al.*, "Sparse idma: A joint graph-based coding scheme for unsourced random access," *IEEE Transactions on Communications*, 2022, vol. 70, no. 11, pp. 7124–7133.
- [16] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2013, pp. 332–336.
- [17] S. D. Howard, A. R. Calderbank, *et al.*, "A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes," in *2008 42nd Annual Conference on Information Sciences and Systems*, 2008, pp. 11–15.
- [18] A. Vem, K. R. Narayanan, *et al.*, "A user-independent successive interference cancellation based coding scheme for the unsourced random access gaussian channel," *IEEE Transactions on Communications*, 2019, vol. 67, no. 12, pp. 8258–8272.
- [19] Z. Si, K. Niu, *et al.*, "An encoder and transmission design scheme for random massive access," *IMT2030 NOMA Conference, China, Tech. Rep.*, 2022.
- [20] J. Ma, X. Yuan, *et al.*, "Turbo compressed sensing with partial dft sensing matrix," *IEEE Signal Processing Letters*, 2015, vol. 22, no. 2, pp. 158–161.
- [21] L. Liu, C. Liang, *et al.*, "Capacity optimality of amp in coded systems," *IEEE Transactions on Information Theory*, 2021, vol. 67, no. 7, pp. 4429–4445.
- [22] D. L. Donoho, A. Maleki, *et al.*, "Message passing algorithms for compressed sensing: II. analysis and validation," in *2010 IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo)*, 2010, pp. 1–5.
- [23] L. Liu, J. Tong, *et al.*, "Analysis and optimization of cdma systems with chip-level interleavers," *IEEE J. Sel. Areas Commun.*, 2006, vol. 24, no. 1, pp. 141–150.
- [24] L. Ping, L. Liu, *et al.*, "Interleave division multiple-access," *IEEE Transactions on Wireless Communications*, 2006, vol. 5, no. 4, pp. 938–947.
- [25] L. Ping, P. Wang, *et al.*, "Recent progress in interleave-division multiple-access (idma)," in *Military Communications Conference*, 2007.
- [26] V. Mojtaba, D. Zhiguo, *et al.*, *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Springer, 2019.
- [27] A. Elkelesh, M. Ebada, *et al.*, "Belief propagation list decoding of polar codes," *IEEE Communications Letters*, 2018, vol. 22, no. 8, pp. 1536–1539.
- [28] O. Iscan, D. Lentner, *et al.*, "A comparison of channel coding schemes for 5g short message transmission," in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.
- [29] S. C. Thompson, A. U. Ahmed, *et al.*, "Constant envelope ofdm," *IEEE Transactions on Communications*, 2008, vol. 56, no. 8, pp. 1300–1312.

Biographies



Chunlin Yan received the Ph.D. degree in communication and information system from university of electronic science and technology of China in 2004. He is now with ZGC Institute of Ubiquitous-X Innovation and Applications as senior researcher. He has published more than 30 papers on wireless communication and had applied more than 20 patents on 4G and 5G systems. In 2014 he invented the Gray mapping transmission scheme which is the key technology of 3GPP MUST (Multi-user superposition transmission). From 2016 to 2018 he joined the 3GPP NOMA (Non-orthogonal multiple access) project and proposed a NOMA scheme superior to the international and domestic competitors. His current work focuses on non-binary channel decoder, iterative detection, compressed sensing, 6G new multiple access and 6G new waveform.



Siying Lyu received the B.S. and M.S. degrees from Beijing Jiaotong University in 2017 and 2020, respectively. She is currently with the China Mobile Research Institute. Her research interests include wireless communications, non-orthogonal multiple access (NOMA) and multiple-Input multiple-Output (MIMO).



Sen Wang received the B.S.E. degree from Information Engineering University, Zhengzhou, China, in 2005, and the Ph.D. degree in information and communication engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2013. After graduation in January 2013, he joined the Green Communication Research Center (GCRC), China Mobile Research Institute, as a Project Manager. His research interests include 5G air interface technologies, especially on massive MIMO, new waveform and multiple access, radio resource allocation and performance evaluation for future cellular networks. Together with his colleagues, he has authored 10 IEEE journal and conference papers on CL, ICC, and WCNC in these areas. He also serves as a Principle Member of IMT-2020 requirements and evaluation WG.



Yuhong Huang received the degree from the Beijing University of Post and Telecom. She joined the China Mobile Research Institute in 1996, where she is currently the Deputy General Manager, where she is responsible for the work of wireless technology, terminal technology, test technology, and customer and market research. She has participated in several important projects, such as technical and strategic research, specification, test and trail of GSM900/1800, GPRS/EDGE, CMNet, WLAN, 3G, and LTE. She also leads the research and standardization work of LTE evolution and 5G. She served as the Vice-Chair of the 3GPP SA. She serves as a Steering Committee Member of the NGMN and the Secretary-General of the Global TD-LTE Initiative.



Xiaodong Xu received his B.S degree in Information and Communication Engineering and Master's Degree in Communication and Information System both from Shandong University in 2001 and 2004 separately. He received his Ph.D. degrees of Circuit and System in Beijing University of Posts and Telecommunications (BUPT) in 2007. He is currently a professor of BUPT, a research fellow of the ZGC Institute of Ubiquitous-X Innovation and Applications and a member of IMT-2030 (6G) Experts Panel. He has coauthored nine books/chapters and more than 120 journal and conference papers. He is also the inventor or co-inventor of 51 granted patents. His research interests cover semantic communications, intelligible network, moving networks, multiple access technologies and satellite communications.