White Paper
2015-05

IMT-2020 (5G) Promotion Group

5G Wireless Technology Architecture

Contents

IMT-2020 (5G) Promotion Group was jointly established in February 2013 by three ministries of China (including MIIT, NDRC, and MOST) based on the original IMT-Advanced Promotion Group. The members include the main operators, vendors, universities, and research institutes in China. The Promotion Group is the major platform to promote 5G technology research in China and to facilitate international communication and cooperation.

Introduction

Over the past three decades, mobile communication has experienced an extraordinarily high-speed development, from simple voice systems to mobile broadband multimedia systems. It not only profoundly changes the life style of everybody, but also greatly promotes the advancement of society and economy. Mobile internet and Internet of Thing (IoT) are two major driving forces of future mobile communications, providing a broader development space for the fifth generation mobile communication system (5G). Towards year 2020 and beyond, 1000x mobile traffic growth, 100 billion connected devices, and more diverse service requirements bring great challenges to the system design of 5G. Compared with 4G, 5G needs to support more various scenarios, integrate with other wireless access technologies, and fully utilize both low-frequency and high-frequency bands. Furthermore, 5G should be capable of supporting flexible deployment and high-efficient operation & maintenance, and significantly improving spectrum efficiency, energy efficiency, and cost efficiency, to realize the sustainable development of mobile communication networks.

The traditional mobile communication evolution is based on the revolution of multiple access technologies. The innovation of 5G wireless technologies is more diverse. Several novel multiple access technologies have been proposed, including sparse code multiple access (SCMA), pattern division multiple access (PDMA), and multi-user shared access (MUSA). Further, massive MIMO, ultra-dense network (UDN) and all-spectrum access (ASA) are identified as the key enabling technologies of 5G. Some other technologies, including new multi-carrier waveforms, flexible duplex, advanced modulation and coding, device-to-device (D2D) communication, and full duplex, are also considered to be candidate technologies. 5G will be built on the technical framework enabled by novel multiple access, massive MIMO, UDN, ASA, etc., to fully meet the technical requirements for 5G in 2020 and beyond.

Currently, the global consensus of vision and requirements of 5G is basically reached, and 5G concept and technology roadmap are gradually clarified. As the international standardization work of 5G will start in the near future, it is urgent to elaborate globally agreed 5G technical roadmap and wireless technology architecture integrating various key enabling technologies, to guide the international standardization and industrialization of 5G.

Scenarios and Technical Requirements

Compared with previous generations of mobile communications, 5G needs to meet extremely high performance requirements in more diverse scenarios. Four typical technical scenarios of 5G are derived from the main application scenarios, service requirements, and key challenges of mobile internet and IoT. They are seamless wide-area coverage scenario, high-capacity hot-spot scenario, low-power massive-connection scenario, and low-latency high-reliability scenario.

- Seamless wide-area coverage is the basic scenario of mobile communications. The main challenge for 5G is to provide more than 100 Mbps user experienced data rate with guaranteed service continuity anytime and anywhere, regardless of static state or high-speed moving, and coverage center or coverage edge.
- High-capacity hot-spot scenario mainly targets local indoor and outdoor hot-spot areas where ultra-high data rates should be provided to users and ultra-high traffic volume density needs to be handled. The main challenges include 1 Gbps user experienced data rate, tens of Gbps peak data rate, and tens of Tbps/km^2 traffic volume density.
- Low-latency high-reliability scenario mainly targets special application requirements of IoT and vertical industries such as Internet of Vehicles (IOV) and industrial control. The ms-level end-to-end latency or/and nearly 100% reliability need to be guaranteed in this scenario.
- Low-power massive-connection scenario mainly targets sensoring and data collecting use cases, such as environmental monitoring and intelligent agriculture. This scenario is characterized by small data packets, low power consumption, low cost, and massive connections. Specifically, at least 1 million connections per squared kilometer need to be supported.

To sum up, the major technical challenges of 5G include 0.1 ~ 1 Gbps user experienced data rate, tens of Gbps peak data rate, tens of Tbps/km^2 traffic volume density, 1 million connections per squared kilometer, ms-level end-to-end latency, and 100 times energy efficiency improvement and cost reduction per bit.

5G Wireless Technology Roadmap

In order to meet the technical requirements of 5G scenarios, a suitable wireless technology roadmap is needed to guide the development of 5G standards and industry. With consideration of 5G requirements, technology trends, and smooth network evolution, the wireless technology roadmap can be derived, which includes a new air interface (AI) and 4G evolution AI. The new air interface has two branches: low-frequency and high-frequency new AIs.

As the de-facto unified standard of 4G, LTE/LTE-Advanced has been deployed worldwide. In order to improve user experience continuously and ensure smooth evolution, 4G should be further enhanced. With backward compatibility, the air interface of 4G evolution should adopt enhanced technologies based on LTE/LTE-Advanced framework in the legacy spectrum. The data rate, system capacity, connection capability, and latency can be improved to partly meet the technical demands of 5G.

With the constraint of the current 4G technical framework, the performance of some enhanced technologies, such as massive MIMO and ultra-dense network, is limited, while some other advanced technologies including all-spectrum access and several novel multiple access schemes cannot be employed. As a result, 4G evolution is not capable of meeting the extremely high performance requirements in future. 5G requires a brand new air interface to break the restriction of backward compatibility and make full use of various advanced technologies to meet all the requirements of 5G. The new air interface is the main body of 5G, while 4G evolution is complementary.

A low-frequency new air interface of 5G is required to meet the demands of user experiences and massive connections in wide-coverage and high-mobility scenarios. Moreover, a high-frequency new air interface is needed to achieve ultra-high data rates and system capacity by utilizing abundant high-frequency spectrum resources. According to global spectrum planning and propagation characteristics on different frequency bands, 5G should include a low-frequency new air interface working below 6GHz and a high-frequency new air interface working above 6GHz.

The low-frequency new air interface will be newly designed with the introduction of advanced technologies such as massive MIMO, novel multiple access, and new waveforms. The shorter frame structure, more simplified signaling procedure, and more flexible duplex modes will be supported. The low-frequency new air interface can efficiently fulfill the performance requirements including user experienced data rate, latency, connection density, and energy efficiency in most scenarios such as wide-area coverage scenario, massive-connection scenario, and high-mobility scenario. Although different scenarios have very different requirements, a unified technical solution needs to be designed, where the technical components and parameters can be flexibly configured to adapt to different scenarios.

In the high-frequency new air interface, it is necessary to consider the impacts of high-frequency channel characteristics and radio frequency (RF) components. Correspondingly, the air interface technical solution, including waveform, modulation, coding, and MIMO, should be optimized. Since there are a number of candidate high-frequency bands in a wide range, a unified technical solution should be designed to facilitate standardization, reduce cost, and simplify operation and maintenance. With parameters adjustment, it can adapt to the characteristics of high-frequency channels and RF components.

Due to the poor coverage, the high-frequency air interface is unfit to stand-alone deployment, and needs to be integrated with the low-frequency air interface. In such cases, the low-frequency air interface offers seamless coverage, control plane functions, and basic data transmission, while the high-frequency air interface is supplementary and provides high data rates in hotspot areas when the channel condition is good.

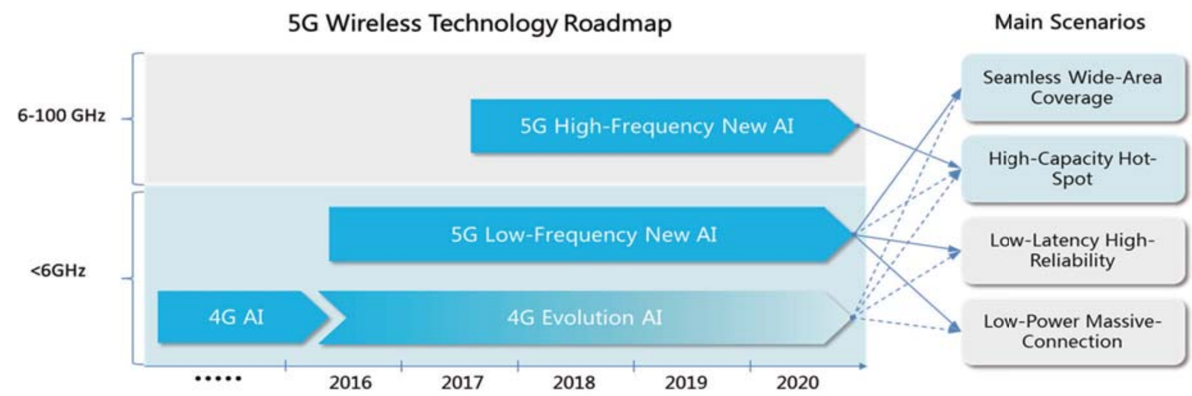


Figure 1 5G wireless technology roadmap

5G Air Interface Technical Framework

1. Design Principle

The technical framework of 5G air interfaces should be unified, flexible, and configurable. Based on such diverse performance requirements, it is preferred to design a specially optimized technical solution for each scenario. However, from the perspective of standardization and industrialization, 5G should have a unified technical framework for the new air interface and 4G evolution. For each scenario, an optimized technical solution can be obtained by flexibly configuring technical components and parameters.

2. Technical Framework

Based on the main functions of mobile communication systems, the technical framework of 5G air interfaces can be divided into the following main modules. They are frame structure, duplex, waveform, multiple access, modulation & coding, antenna, and protocol. Each module should cover the common technical contents as many as possible to reach a good balance between flexibility and simplicity. Different modules can interconnect and work together. According to the technical requirements of each scenario, the technical modules can be optimized and flexibly configured to form the corresponding technical solution. The main technical modules are described as follows.

- Frame structure & channelization: The parameters of 5G frame structure can be flexibly configured for diverse scenarios and services. The bandwidth, subcarrier spacing, cyclic prefix (CP), transmission time interval (TTI), uplink-downlink configuration, etc., can be adapted to different frequency bands, scenarios, and channel environments. The reference signals and control channels can be flexibly configured to support the application of massive MIMO, novel multiple access, and other advanced technologies.
- Duplex: 5G will support traditional FDD, TDD, and their enhancements, and may support new duplex technologies such as flexible duplex and full duplex. Low-frequency air interface will adopt FDD and TDD, and TDD is more suitable for high-frequency air interface. In addition, flexible duplex can allocate uplink and downlink time and frequency resources flexibly to better adapt to non-uniform dynamic service distribution. Full duplex is a potential duplex technology which can support transmitting and receiving at the same time and frequency resource.
- Waveform: In addition to traditional OFDM and single-carrier waveforms, it is possible to adopt new waveforms based on optimized filtering, such as filter bank multicarrier

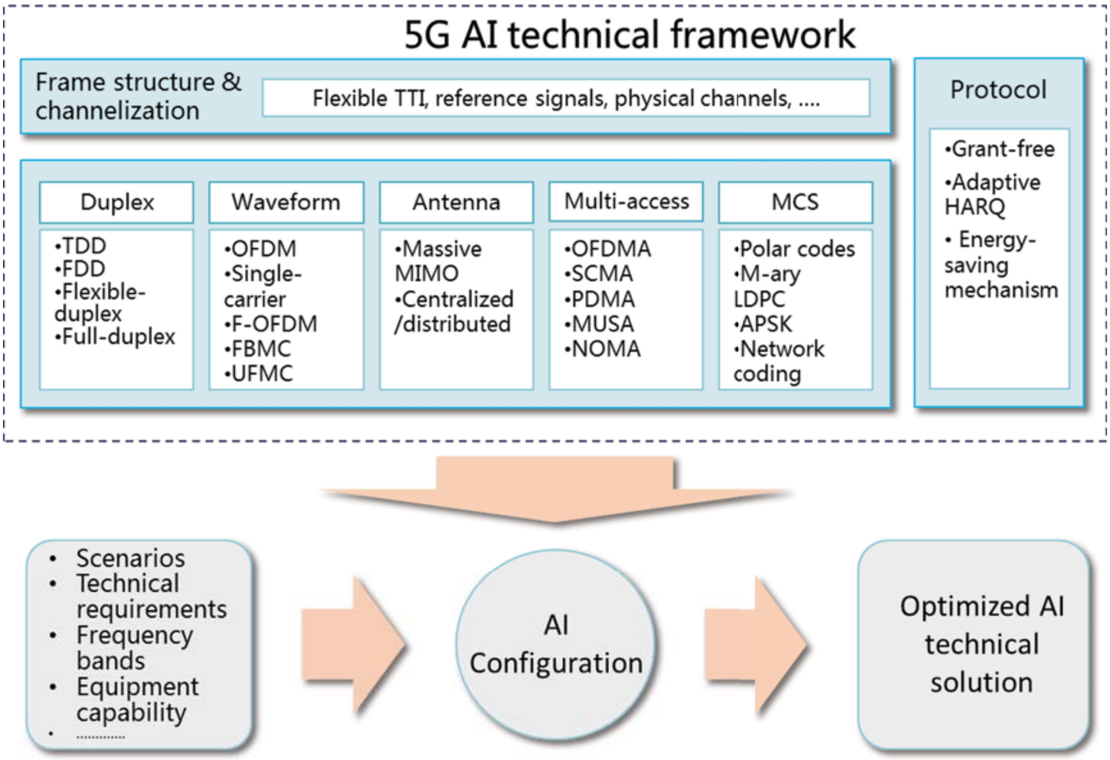


Figure 2 Flexible and configurable AI technical framework of 5G

(FBMC), filtered-OFDM (F-OFDM), and universal filtered multicarrier (UFMC). With very low out-of-band leakage, those waveforms can improve spectrum utilization efficiency, efficiently utilize fragmented spectrum, and coexist with other waveforms. Different waveforms have different out-of-band leakages, overhead, and peak-to-average power ratio (PAPR). Therefore, suitable waveforms can be selected to adapt to the requirements of different scenarios. In some cases, multiple waveforms can coexist.

- Multiple access: 5G will support not only traditional OFDMA, but also some novel multiple access schemes including SCMA, PDMA, and MUSA, which can support more connections and improve spectrum efficiency via multi-user superposition transmission. Moreover, the air interface latency can be significantly reduced by grant-free contention access.
- Modulation & coding: 5G needs to support various services such as high data rate, low data rate, small packet, low latency, and high reliability. For high-data-rate services, M-ary low density parity check (LDPC), polar codes, and new constellation mapping, and faster-than Nyquist (FTN) can further improve link spectrum efficiency, compared with traditional binary Turbo codes and QAM. For low-data-rate and

small-packet services, polar codes and low-code-rate convolutional codes can be used to approach Shannon capacity in the cases of short code length and low SNR. For low-latency services, the coding schemes with fast encoding and decoding algorithms are preferred. For high-reliability services, the "error floor" effect of decoding algorithms needs to be avoided. In addition, there may be a number of wireless backhaul links in dense networks, the system capacity can be increased by network coding.

- Antenna: Massive MIMO will be utilized in 5G systems. Each base station can be equipped with more than one hundred antennas and dozens of antenna ports, which can enable advanced multi-user MIMO to support more users in spatial domain. As a result, the system spectrum efficiency will be improved by several times. Massive MIMO can also be used in high-frequency bands to overcome large path loss by adaptive beamforming. For the application of massive MIMO, the reference signals, channel estimation, channel information feedback, multi-user scheduling mechanism, and receiving algorithms need to be improved and optimized.
- Protocol: The air interface protocol of 5G needs to support various advanced scheduling, link adaption, and multi-connection. The protocol is capable of being configured flexibly to meet the requirements of different scenarios, and efficiently supporting radio access technologies including the new air interface, 4G evolution air interface, and WLAN. To reduce the signaling overhead of massive small-packet services, the grant-free access protocol can be employed to simplify signaling interaction procedure between base stations and terminals. As a result, the access latency will be reduced greatly. The adaptive hybrid automatic repeat request (HARQ) protocol of 5G should match different service requirements in latency and reliability. In addition, 5G should support more efficient energy-saving mechanism to meet low power consumption requirements of IoT services.

Based on the air interface technical framework of 5G, an optimized combination of technologies and parameters can be selected to form a technical solution for specific performance requirements, available frequency bands, equipment capabilities, and implementation cost. The technical framework is capable of dealing with the future new scenarios and service requirements to realize "forward compatibility".

3. Consideration on 5G Low-Frequency New AI

The low-frequency new air interface can be widely used in various scenarios including seamless wide-area coverage, high-capacity hot-spot, low-power massive-connections, and low-latency high-reliability. This technical solution can be optimally designed based on the unified 5G technical framework by integrating key technologies including massive MIMO, novel multiple access, new multi-carrier waveforms, and advanced modulation & coding.

In seamless wide-area coverage scenario, with the good channel propagation at the frequency bands below 6 GHz, the low-frequency new air interface can achieve 100 Mbps user experienced data rate through increasing bandwidth and improving spectrum efficiency. For the frame structure, it is necessary to increase the subcarrier spacing and shorten the frame length to better support wider bandwidth. And the compatibility with LTE frame structure can be taken into account. For example, the frame length can be divided exactly by 1 ms and the subcarrier spacing can be a multiple of 15 kHz. For multi-antenna technology, massive MIMO can be adopted in base stations to improve the system spectrum efficiency. The number of antennas can be up to 128, and the parallel transmission of more than 10 users can be supported. The OFDM waveform can still be adopted, and the same waveform design can be utilized for both uplink and downlink. The new waveform technology like F-OFDM can be used to support the coexistence with technical solutions for other scenarios. The connection capability and spectrum efficiency of OFDMA systems can be improved by introducing novel multiple access technologies based on superposition coding. The reference signals, channel estimation, and multi-user pairing mechanism should be newly designed for the application of massive MIMO and novel multiple access. For duplex technologies, TDD can use channel reciprocity to obtain better performance of massive MIMO. Furthermore, the control plane of macro cells can be further enhanced and the C/U split needs to be supported, in order to realize the efficient control and management for small cells and users.

In high-capacity hot-spot scenario, by increasing cell density, spectrum efficiency and bandwidth, the low-frequency new air interface can meet the requirements of data rates and traffic volume density to some extent. The technical solution should be generally consistent with the solution of seamless wide-area coverage scenario and needs to be further optimized in the following aspects. The parameters of frame structure can be modified according to the channel and service characteristics of high-capacity hot-spot scenario. In some environments without serious interferences, flexible duplex or full-duplex may be introduced. Higher order modulation and coding schemes can be adopted to improve the spectrum efficiency. In order to reduce interferences in densely deployed networks, technologies including adaptive small cell clustering, multi-cell cooperative transmission, and frequency resource coordination can be utilized. Furthermore, the traditional cell boundaries can be broken, so that multiple cells can co-work more efficiently to serve one subscriber to achieve user centric cell virtualization. To provide a flexible backhaul mode for small cells, it should be considered to jointly design access and backhaul. Adaptive frequency allocation between access and backhaul can improve frequency utilization efficiency. In addition, the impacts of wireless networking modes including centralized, distributed, and MESH networks should be taken into account in system design.

In low-power massive-connection scenario, the IoT services featured with small data packet, low power, massive connections and burst transmission do not require large channel bandwidth. Non-contiguous and fragmented low-frequency bands or a few OFDM subcarriers can be used for this scenario. The novel multiple access technologies, such as SCMA, PDMA and MUSA, can support massive connections via multi-user superposition transmission. By supporting grant-free transmission, they can simplify the signaling process and reduce the power consumption. In terms of waveforms, the new efficiently-filtered waveform technologies (e.g., F-OFDM and FBMC) can be used to lower out-of-band interferences, to utilize non-contiguous and fragmented bands efficiently, and to decouple technical solutions between different subband. The coding, modulation, multiple access, and signaling process in different subbands can be configured independently. By narrow-band system design, the system coverage and connection capability can be improved, and the power consumption and cost of terminals can be reduced. Moreover, the energy-saving mechanisms for both connected and idle states need to be improved. In connected state, contention access can simplify the signaling process, reduce user access delay, and shorten turn-on time. In idle state, longer paging intervals can make terminals sleep longer to reduce the power consumption of terminals.

In low-latency high-reliability scenario, to achieve the requirement of latency, the air interface delay should be reduced significantly, and the network forwarding delay should be decreased through minimizing forwarding nodes. To meet the requirement of high reliability, the number of retransmission per unit time should be maximized, and the transmission reliability of single link should be improved. To reduce the air interface delay, the frame length needs to be shortened and the compatibility with the solution for seamless wide-area coverage scenario should be considered. Since the shorter TTI may cause larger CP overhead, the new waveforms supporting non-CP and CP sharing between multiple symbols can be utilized. Multiple access technologies including SCMA, PDMA, and MUSA can enable grant-free transmission, avoiding resource allocation process, to achieve "zero" waiting time for uplink data packet scheduling. To reduce the network forwarding delay, one solution is to shorten the transmission path through sinking some core network functions to access network and localizing service contents. Another solution is to adopt dynamic, clustered access network structure and support dynamic MESH communication link, single-hop and/or multi-hop direct communications among base stations or/and terminals, to further reduce end-to-end latency. To improve data transmission reliability, technologies such as advanced coding and space/time/frequency diversity can be utilized. In terms of protocol, the retransmission performance can be improved via enhanced HARQ mechanism. In addition, technologies such as enhanced coordinated multiple points (CoMP) and dynamic MESH can strengthen the cooperation between base stations and between terminals to further enhance the data transmission reliability.

4. Consideration on 5G High-Frequency New AI

The high-frequency new air interface utilizes extra-large bandwidth to fulfill the requirement of extremely high transmission rate in high-capacity hot-spot scenario. Because of small coverage and strong directivity of high-frequency wireless signals, dense deployment can be considered to achieve very high traffic volume density. Massive MIMO can compensate the large path loss through adaptive beamforming and tracking, and use space multiplexing to support more users and increase system capacity. To support extra-large bandwidth, the subcarrier spacing can be increased over 10 times, and the frame length can be shortened significantly in comparison with LTE. The same waveform design can be adopted in both uplink and downlink, where OFDM is still an important candidate. However, single carrier is also a potential waveform if the challenges caused by high-frequency RF components and propagation characteristics are considered. TDD is preferred to support high-frequency communications and massive MIMO. Considering the requirements of high data rates and system capacity, channel codes that support fast decoding and require less storage should be selected. The high-frequency new air interface requires high-speed backhaul links. By using rich high-frequency spectrum resource, the wireless self-backhaul of base stations can be realized through a jointly designed access and backhaul links. To deal with the poor coverage, the high-frequency air interface needs to be integrated with the low-frequency air interface in support of C/U split. The low-frequency air interface should be responsible for control plane functions, while the high-frequency air interface can be used to achieve high data rates. Dual connectivity and dynamic load balancing can be realized in user planes of the two air interfaces.

5. Consideration on 4G Evolution AI

The air interface of 4G evolution will be based on LTE/LTE-Advanced technical framework. It should be further optimized in several aspects such as frame structure, MIMO, and multiple access technologies. With the guarantee of smooth evolution, it can partly meet the requirements of 5G, such as data rate, latency, traffic volume density, and connection density. The air interface latency can be reduced by decreasing the OFDM symbols in each TTI and introducing the optimized scheduling and feedback mechanisms. The 3D channel information can be utilized to achieve more accurate beamforming, which can support more users and more data streams. Multi-user superposition transmission and enhanced interference cancellation algorithms can be employed to improve the system spectrum efficiency and connection capability. To meet the requirements of IoT applications, the narrow-band scheme can improve coverage, increase connection capability, and reduce power consumption and cost. In addition, the air interface of 4G evolution should be able to cooperate tightly with the new air interface, and serve users via advanced technologies like dual connectivity.

5G Wireless Key Technologies

1. Massive MIMO

MIMO has been widely used in 4G systems. In order to meet the requirements of high data rates and system capacity of 5G, increasing the number of antennas is a major direction of MIMO enhancement. According to the large-scale random matrix theory, when the number of base station antennas greatly exceeds the number of terminal antennas, the random channel vectors between a base station and users tend to be mutually orthogonal, leading to the asymptotic zero interference among users. On the other hand, the significant array gain is able to effectively improve the signal-to-interference-plus-noise ratio (SINR) of individual users. As a result, much more users can be served on the same time-frequency resources.

In practical deployments, the base station with large number of antennas can form extreme sharp beams with high spatial resolution and beamforming gain, which can provide flexible spatial multiplexing capability, improve received power, and suppress multi-user interferences. Thus, the much higher system capacity and spectrum efficiency can be achieved.

The major research areas of massive MIMO include:

(1) Deployment scenario and channel modeling

The potential deployment scenarios of massive MIMO includes macro coverage, high-rise coverage, heterogeneous network, indoor and outdoor hotspots and wireless backhaul, etc. In addition, the massive MIMO system with distributed antenna is one of the possible deployment scenarios as well. The current frequency bands can be utilized in the macro coverage, while high-frequency bands can be adopted in hotspots and wireless backhaul. In the above typical deployment scenarios, field measurements are needed to model the distribution and correlation of a serial of parameters in massive MIMO channel, which can reflect the propagation characteristics in 3D domain.

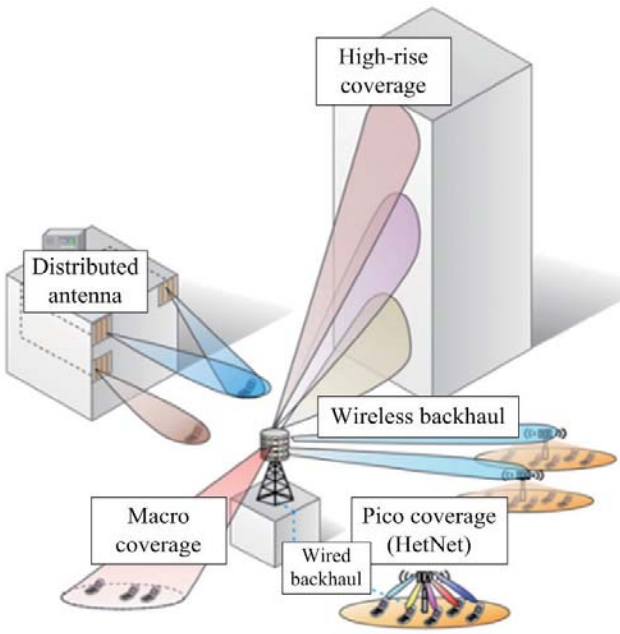


Figure 3 Deployment scenarios of massive MIMO

(2) Transmission and detection technologies

The performance gain of massive MIMO systems is mainly guaranteed by the orthogonality between the channel vectors of the users provided by the large number of antennas. Under real channel conditions with non-ideal factors caused by equipment and propagation environment, for effective suppression of co-channel interference among users and cells, the algorithms of downlink transmission and uplink reception need to be designed carefully. The computational complexity of transmission and detection algorithms are relevant to the scale of antenna arrays and the number of users. In addition, the precoding & beamforming algorithm for massive MIMO will be related with the antenna array structure, equipment cost, power efficiency and system performance. It is expected that the Kronecker product-based separate horizontal and vertical precoding algorithm, hybrid digital-analog beamforming, or staged beamforming will reduce computational complexity of large-scale antenna systems effectively.

(3) Channel state information (CSI) measurement and feedback

CSI measurement, feedback and reference signal design are very important for the application of MIMO technology. To better compromise between the overhead and the accuracy CSI acquisition, in addition to the schemes such as traditional codebook-based implicit feedback and channel reciprocity-based CSI acquisition, new CSI mechanisms including staged CSI measurement and feedback, Kronecker product-based CSI measurement and feedback, compression sensing, and pre-experienced feedback, need to be studied.

(4) Coverage enhancement techniques and high mobility solution

Increasing the number of antennas will offer huge gain on coverage of the traffic channel, However, it will cause some challenges for the broadcast channel which requires uniform coverage for all the users in the cell. In this case, inner-outer dual-ring beam sweeping can be used to cope with the wide coverage problems of the narrow beams. Moreover, the requirements of reliability and high-data-rate transmission under high mobility should be taken into account in the massive MIMO systems. Beam tracking and beam broadening techniques, which are less dependent on CSI acquisition, are able to improve the transmission reliability and data rate by utilizing the array gain efficiently.

(5) Multi-user scheduling and resource management

Due to the high spatial resolution and much more spatial freedom of the wireless access network offered by massive MIMO, considerable performance gain can be obtained by multi-user scheduling, traffic load balancing and resource management.

(6) Large-scale active antenna array

Large-scale active antenna array can be classified into full-digital array and hybrid digital-analog array. Considering the implementation complexity, power consumption and cost, the hybrid digital-analog structure will have great potential when utilized in the high-frequency bands. Factors that have impact on the performance and efficiency of massive MIMO in practical environments include the large-scale active antenna array structure, high-accuracy supervision and calibration, transceiver modules with high efficiency, high reliability, miniaturization, low cost, and modularization, as well as high-accuracy supervision and calibration. These key factors will have great impact on the practical deployment of massive MIMO.

In general, massive MIMO improves system spectrum efficiency, user experience, and transmission reliability, and provides flexible interference management and coordination schemes in the heterogeneous and dense networks. With the breakthrough of key techniques and further development of RF components and antennas, massive MIMO will surely play an important role in the 5G systems.

2. Ultra-Dense Network

Ultra-dense network is a major technology to meet the requirements of ultra-high traffic volume density in year 2020 and beyond. High spectral reuse factor can be achieved in ultra-dense network via densely deployed wireless equipment. This results in hundreds of times capacity improvement in hot-spot areas. Typical scenarios of ultra-dense network include office, dense residential areas, dense urban areas, campus, open-air gathering, stadium, subway, and apartment. The continuing increase of cell density brings many new challenges such as interference, mobility, sites acquisition, transmission resource, and deployment cost. To meet the requirements of typical scenarios and technical challenges, and realize easy network with easy deployment, easy maintenance, and delightful user experience, the joint access and feedback, interference management and suppression, and cell virtualization, are important research areas in ultra-dense network.

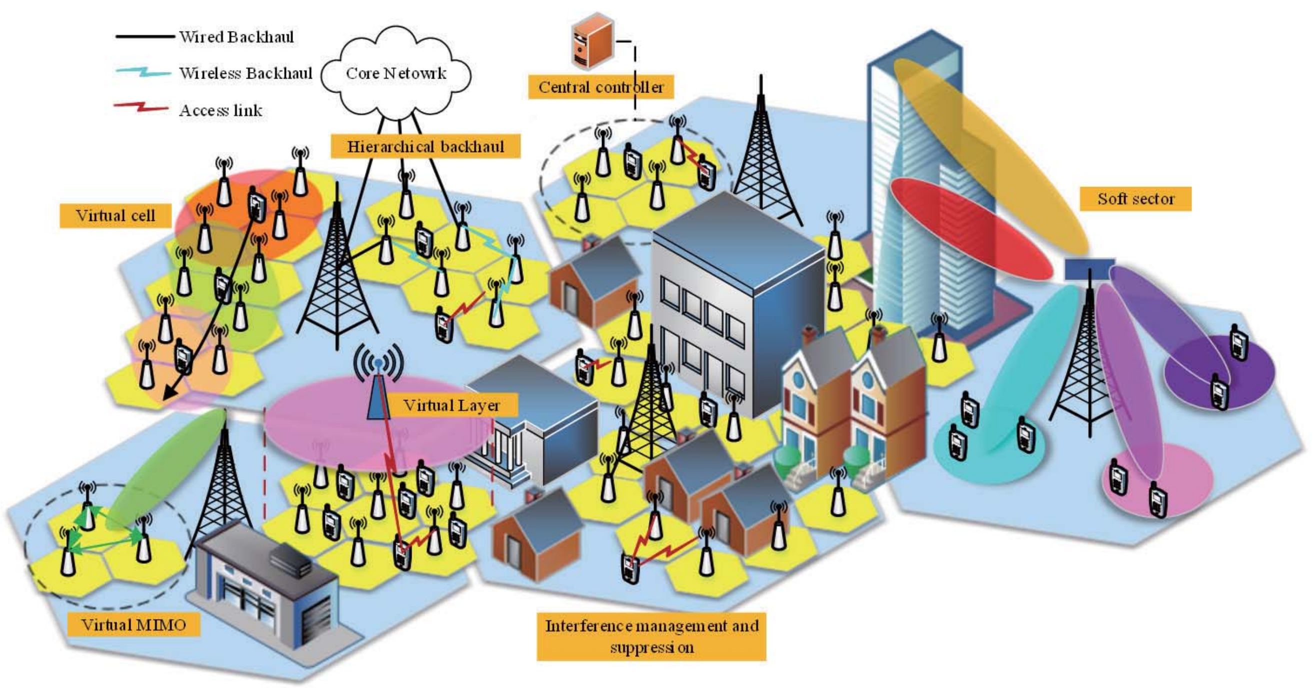


Figure 4 Key Technologies in UDN

(1) Joint design of access and feedback

Joint design of access and backhaul links contains several key techniques, such as hierarchical backhaul, multi-hop, multi-route backhaul, wireless self-backhaul and flexible backhaul.

In the hierarchical backhaul architecture, base stations are categorized by stratum indices. The first stratum includes macro cell layer and small cell layer connected with wired backhaul, the small cells on the second stratum are connected to the base stations via one-hop wireless transmission, and the small cells on the third stratum are connected to small cells on the second stratum via one-hop wireless transmission. Wired and wireless backhaul links are combined together in this architecture, which facilitates a plug-and-play way of network deployment.

Multi-hop and multi-route backhaul technique deals with route selection, establishment and management of multiple routes, dynamic route selection, joint interference management and resource coordination for backhaul and access links, which is expected to provide significant capacity gains.

With wireless self-backhaul, the same radio access technology and frequency bands can be shared between backhaul and access links. The radio resource can be partitioned in time or frequency. There are two research topics in wireless self-backhaul: 1) joint optimization of backhaul and access links, 2) enhanced backhaul link. In joint optimization of backhaul and access links, adaptive resource allocation and dynamic resource sharing between backhaul and access links can improve the overall utilization of resources. In enhanced backhaul link aspect, the broadcast and multiple access nature of wireless communication can be exploited to achieve transmission and reception of independent data streams in spatial sub-channels, thus offering more spatial degrees of freedom and resulting in improved throughput. Multiple nodes or terminals can cooperate to form a virtual MIMO network which can support high rank MIMO and reduce the interference.Flexible backhaul is an efficient and economical solution to improve the backhaul capability of UDN. Through flexibly utilizing available network resource (including both wired and wireless resource), flexibly adjusting network topology and backhaul strategy to match the network resource and service payload, and flexibly allocating the network resource in both backhaul and access link to improve the end-to-end transmission efficiency, flexible backhaul is expected to meet the end-to-end quality-of-service requirement with lower deployment and operating cost.

(2) Interference management and suppression strategy

Ultra-dense network can effectively improve the system capacity. However, as the cell density increases, there would be more overlap between cells, leading to severe interferences. The strategy of interference management and suppression includes adaptive clustering, multi-cell coherent cooperative transmission based on centralized control, and cluster-based multi-cell frequency resource coordination. By turning on/off small cells at subframe level, adaptive clustering can form dynamic cell clusters and terminate cells that are not serving users or in need for extra capacity, thus reducing the interference to neighboring cells. Multi-cell coherent cooperative transmission enables joint transmission over appropriate set of neighboring cells. Therefore, terminals can combine the signals from multiple cells without interference, resulting in system throughput enhancement. Cluster-based multi-cell frequency resource coordination can improve cell edge performance by optimal partitioning of frequency resource between dense clusters, e.g. utilizing the same frequency bands in each cluster and different frequency bands in different clusters, which can improve edge user experienced rates.

(3) Cell virtualization

Cell virtualization includes user centric virtual cell, virtual layer technology, and soft sector technology. User centric virtual cell has the goal of edgeless network. Driven by the coverage and services, virtual cell keeps updating with the movement of users and maintains a good link quality and QoS/QoE with each terminal, regardless the location of the end user. Virtual layer technology is a solution that relies on virtual layer and real layer networks. Wherein, the virtual layer is responsible for broadcasting, paging and mobility management, while data transmission is carried on the real layer. No cell reselection or handover is needed for the mobile users within the same virtual layer, thus the user experience can be delightful. For soft sector technology, multiple soft sectors are formed by means of multiple beams generated by the centralized entity. It can significantly reduce the cost of site rental, equipment and transmission. Soft sector technology can provide a unified management platform between soft sectors and real sectors, and reduce the operation complexity. It is a light solution with easy deployment and easy maintenance.

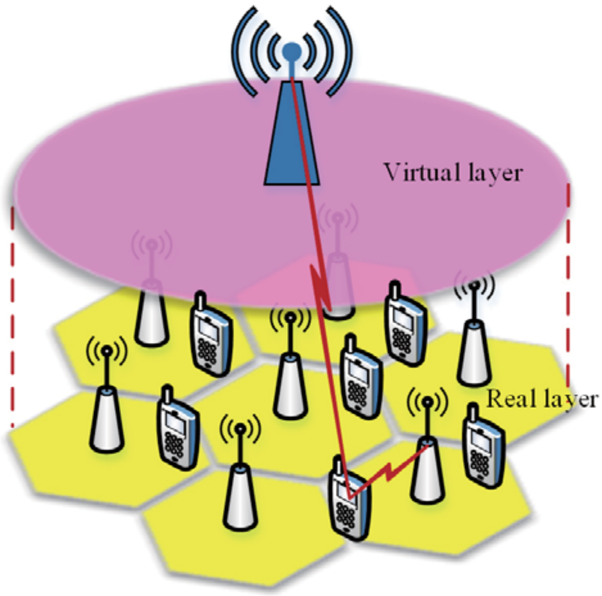


Figure 5 Virtual layer technology

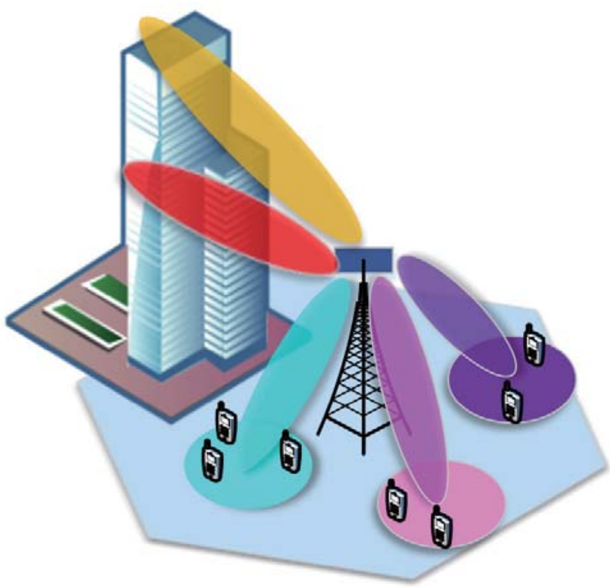


Figure 6 Soft sector

3. All-Spectrum Access

All-spectrum access involves low-frequency (LF) bands below 6 GHz and high-frequency (HF) bands above 6 GHz, where the former is the core bands of 5G that are used for seamless coverage, and the latter is the supplementary bands that are used to achieve high data rates in hotpot areas. To fully exploit the advantages of different frequencies, all-spectrum access supports hybrid network integrating low-frequency and high-frequency bands to simultaneously meet the requirements of seamless coverage, ultra-high data rates, and ultra-high capacity. Considering the big differences between high-frequency bands and traditional low-frequency bands, the research on all-spectrum access mainly focuses on key technologies involving high-frequency mobile communications. Nowadays, it has been widely recognized that the high-frequency range for 5G is from 6 GHz to 100 GHz, which can provide abundant available spectrum to obtain more than 10 Gbps user transmission rates. As a result, the ultra-high capacity and ultra-high data rates required by 5G can be achieved.

High-frequency communication has been implemented in military communications and WLAN, but its application in cellular communication is just in a beginning stage. High-frequency signals are susceptible to environmental effects such as obstruction, reflection, refraction and air absorption in mobility, which makes high-frequency channels have some properties obviously different from conventional cellular channels, for example, strong propagation path loss, fast channel variation, low diffraction ability, etc. Hence, it is necessary to conduct comprehensive research on high-frequency channel measurement and modeling, new air interface design, networking technology, and RF components.

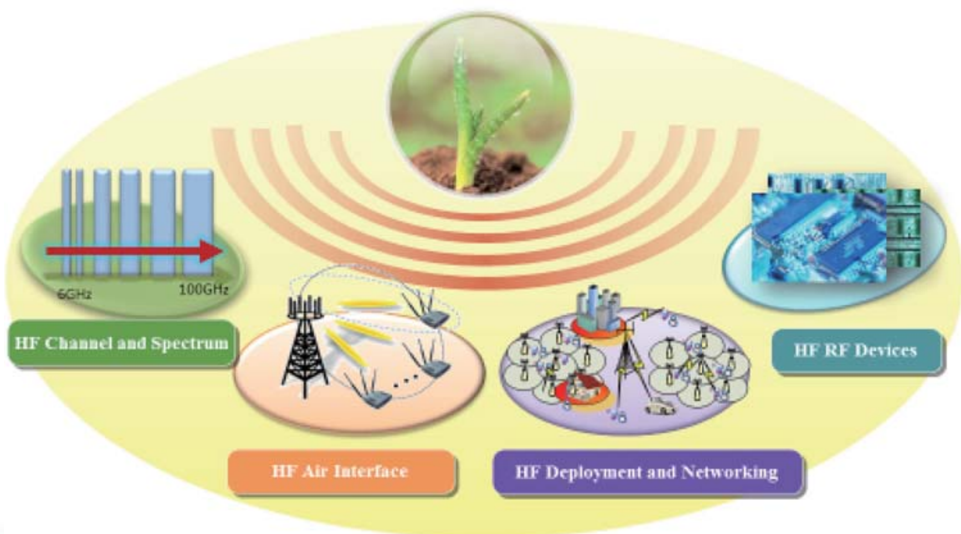


Figure7 Key Technologies of HF Communication

(1) HF candidate spectrums and channel characteristics

Currently, the frequency bands of interest ranges from 6 GHz to 100 GHz, containing authorized and unauthorized bands, symmetric and asymmetric bands, contiguous and non-contiguous bands, etc. On the basis of the potential high-frequency candidates and the existing industrial channel measured results, the research work in this area is to study the channel propagation characteristics, derive the channel models, analyze and evaluate the suitable scenarios, and select the suitable high-frequency bands. The main candidate bands for 5G focused by the industry include 6 GHz, 15 GHz, 18 GHz, 28 GHz, 38 GHz, 45 GHz, 60 GHz, and 72 GHz. The measuring scenarios include indoor and outdoor hotspots.

The preliminary channel measurements indicate higher frequencies come with larger propagation loss. A new feature of high-frequency channels is its dependence on antenna configuration, especially for transmission loss, delay spread, received power and AoA/AoD. Therefore, how to decouple channel measurements with antenna configuration is a research focus of high-frequency channel modeling. The parameters on channel transmission loss can be obtained by two widely accepted path loss models: close-in reference and floating intercept. By comparison, the close-in reference model is more robust if the measured data is insufficient, while the floating intercept model is more reasonable in the case of sufficient measured data.

(2) HF air interface design

Based on the channel characteristics, the high-frequency air interface research focuses on transmitting and receiving beamforming with multi-antennas and antenna arrays related techniques as well as narrow beam training and tracking for larger coverage. The research topics also include the signal waveforms for different high-frequency channel propagation properties, and the unified frame structure and access mechanism for low-frequency and high-frequency hybrid networks. In addition, it is interesting to study the adaptive spectrum sensing techniques supporting multiple frequency usage modes such as authorized and unauthorized frequency bands. Furthermore, modulation and coding schemes, burst transmission, interference management, efficient MAC technologies, and joint access and backhaul are studied to improve the performance of the high-frequency air interface.

(3) LF and HF hybrid networking

As the supplement to the low-frequency air interface, the high-frequency air interface is mainly deployed as indoor and outdoor hotspots to provide high data rates. Due to the propagation characteristics of high frequency signals, hybrid networks integrating low-frequency and high-frequency air interfaces are required. With C/U splitting, ultra-dense networking, and high-frequency adaptive backhaul, the large traffic volume and high data rates in some scenarios like hotspots can be achieved with low deployment cost. With the features of narrow beams and small coverage, the high-frequency communications can be employed in some new wireless application scenarios including D2D and vehicle radar.

(4) HF components

Compared with low and medium frequencies, high frequency devices are more convenient for massive antenna integration and device miniaturization. Benefit from the commercialization of microwave products, 6-100 GHz devices are relatively mature, especially for 14 GHz, 23 GHz, 28 GHz, V-band, and E-band. However, a further breakthrough is required to apply them in cellular communications. Both high-frequency power amplifier and low-noise amplifier need to be further improved in aspects of power efficiency and phase noise. The ADC and DAC devices need to support the sampling frequencies of at least 1GHz channel bandwidth. The new type high-frequency array antenna is required to realize high-gain beamforming and wide-range spatial scanning.

To verify these key technologies, the mobile industry has developed some high-frequency communication prototype systems. Among them, an E-band prototype system can realize the transmission rate up to 115 Gbps, and a 20-40 GHz prototype system can support 10 Gbps transmission rate. The feasibility of using high-frequency communication to support extremely high data rates has been preliminarily verified.

In brief, high-frequency communication can significantly improve data rates and system capacity using abundant high-frequency spectrum resources. It is regarded as a breakthrough technology for conventional cellular communications. To facilitate its standardization and industrialization, further breakthroughs are needed in channel propagation properties, air interface technical solution, high-frequency networking, and RF components.

4. Novel Multiple Access

Towards year 2020 and beyond, mobile internet and IoT will become the main forces driving the evolution of mobile communications. 5G should not only improve the spectrum efficiency significantly, but also support the connections of massive devices. Furthermore, there are high requirements on the simplification of system design and signaling process. All of these requirements make great challenges for the existing orthogonal multiple access technology.

Novel multiple access, represented by SCMA, PDMA, and MUSA, can significantly improve the spectrum efficiency and access capability by superimposing multi-users information in the same transmission resources and utilizing advanced algorithms to recover multi-user information at the receiver. Moreover, grant-free transmission can be enabled to simplify the signaling process and reduce the air interface transmission latency.

(1) Sparse code multiple access

SCMA is a novel non-orthogonal multiple access technology based on the code domain superposition. This technology combines the low-density code and modulation technology, selects the optimal codes through conjugation, substitution, and phase rotation, and makes different users transmit information based on assigned codebooks. At the receiver side, the message passing algorithm (MPA) is utilized to decode the data. As a result of non-orthogonal sparse code superpositioning, SCMA can support more users to connect simultaneously with the similar spectrum resource, and improve the quality of single-user link through multi-dimensional modulation and spread spectrum. In addition, grant-free random access can be achieved through blind detection techniques and the collision insensitivity of SCMA codewords. This can contribute to the reduction of implementation complexity and latency, and therefore is suitable for IoT services with small-packet transmission, low power, and low cost.

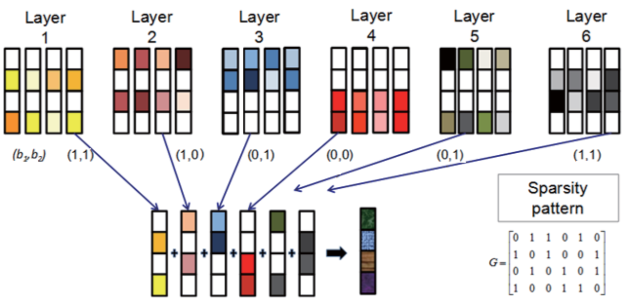


Figure 8 The principle of SCMA

(2) Pattern division multiple access

Based on multi-user information theory, PDMA utilizes pattern segmentation to separate user signals at the transmitter and serial interference cancellation at the receiver. The maximum capacity of multiple access channels can be approached. The user pattern can be designed independently or jointly in space, code, and power domains. Moreover, using user characteristic pattern at the transmitter, pattern segmentation technology can contribute to better distinguish different users, and therefore improve the detection performance of serial interference cancellation at the receiver.

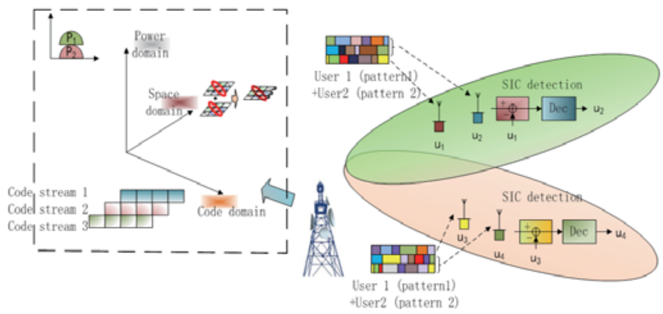


Figure 9 The principle of PDMA

(3) Multi-user shared access

MUSA is a non-orthogonal multiple access scheme base on code domain superposition. For the uplink, the modulated symbols of different users are spreaded by specific spreading sequence, and then are transmitted on the same resources. The SIC algorithm is utilized to decode the user data at the receiver. The design of spreading sequence is the key element that impacts the performance of MUSA scheme. Good cross-correlation property is required when the code length is short (e.g., 4 or 8). For the downlink, based on the conventional power superposition scheme, MUSA utilizes mirror constellation to optimize the symbol mapping of paired users and to improve the downlink performance.

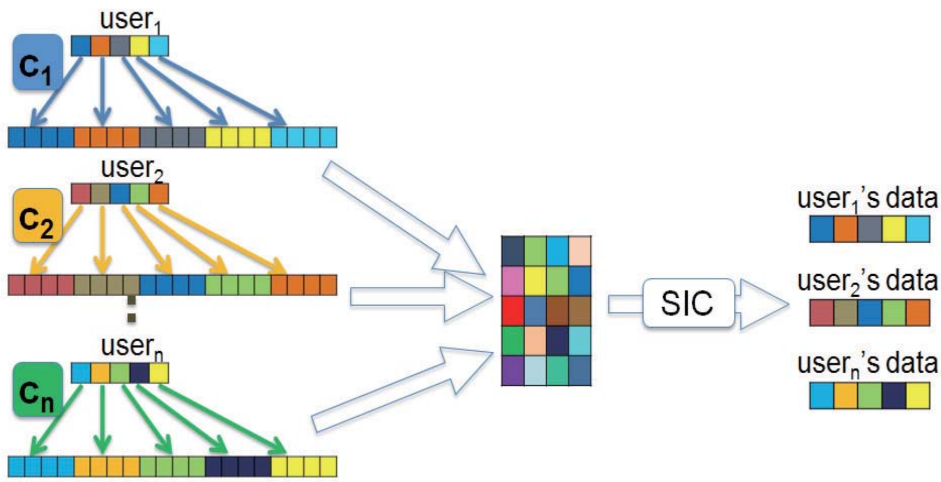


Figure 10 The principle of MUSA

Based on simulation analysis, compared with LTE systems, about 30% improvements on downlink spectrum efficiency and more than 3 times improvements on uplink user connectivity could be achieved by introducing novel multiple access technology. Furthermore, grant-free transmission can be utilized to simplify the signaling process and reduce the data transmission latency.

User experienced data rate, connection density, and latency are the top three key performance indicators of 5G. Compared with OFDM, the novel multiple access technologies above mentioned can not only improve the spectrum efficiency and support more user connections, but also be capable to reduce the latency significantly, and therefore can be one of the fundamental core technologies of 5G.

5. New Multi-Carrier Waveforms

OFDM has been used extensively in 4G, and is still considered as an important candidate waveform for 5G. However, it is insufficient to only use OFDM to deal with the diversified services, higher spectrum efficiency, and massive connections. Some new multi-carrier waveforms can be introduced to assist OFDM to better meet the entire requirements of 5G.

With high robustness against multi-path fading channel, and easy combination with MIMO transmission, as well as flexible frequency-selective scheduling, OFDM can efficiently support mobile broadband (MBB) services. However, OFDM has some drawbacks, e.g., high out-of-band emission (OOBE), sensitivity to time-frequency synchronization error, and the same waveform numerology for the whole bandwidth.

To better support the various application scenarios of 5G, the research of new multi-carrier waveforms should identify the requirements of each scenario. Firstly, the new multi-carrier waveform needs to better support new service. 4G mainly focus on the MBB services, however, 5G will embrace diversified types of services, such as the machine-type communications (MTC) and vehicle-to-vehicle (V2V) communication, which propose new requirements to these basis waveform. In other words, besides the conventional MBB services, the new waveforms will have to provide efficient support for machine-type communications. Secondly, due to new technologies and new services emerging continuously, new multi-carrier waveforms should have good scalability. The new technologies and services can be supported by simply parameter configuration or modification. Finally, in order to meet diverse requirements of 5G, new multi-carrier should achieve good compatibility with other technologies such as new modulation and coding schemes and new multiple access schemes.

To meet these requirements, several new multi-carrier waveforms have been proposed, such as filtered-OFDM (F-OFDM), universal filtered multi-carrier (UFMC) and filter-bank multi-carrier (FBMC). A common feature of these new waveforms is that filters are employed to suppress the OOBE and relax the requirement on time-frequency synchronization, thus to avoid the above-mentioned limitations of OFDM.

The filters in UFMC and F-OFDM are implemented at the granularity of each subband. The main difference is that F-OFDM uses longer filters, and the signal processing procedure is the same as the conventional OFDM in each subband, providing backward compatibility. In contrary, in order to maintain the time-domain inter-symbol orthogonality, UFMC uses shorter filters and the CP of OFDM is replaced with empty guard period. The filter in FBMC is implemented at the granularity of each subcarrier. FBMC gives up the orthogonality in complex domain for better time-frequency localization, which in turn allows for larger freedom in adaptation to channel variations. Moreover, CP is not required by FBMC, and therefore the signaling overhead can be reduced.With these new multi-carrier waveforms, some key requirements on 5G can be satisfied.

(1) Support flexible and configurable new air interface based on new multi-carrier. F-OFDM and UFMC can support independent waveform numerology configuration in different subbands by subband-based filtering. Based on different service types, the system bandwidth can be divided into several subbands, each for a certain service type with the most suitable waveform numerology, such as TTI duration,

subcarrier spacing and CP duration. In this way, a flexible and configurable air interface can be provided, leading to an enhanced capability of supporting different service types, as well as increased flexibility and scalability.

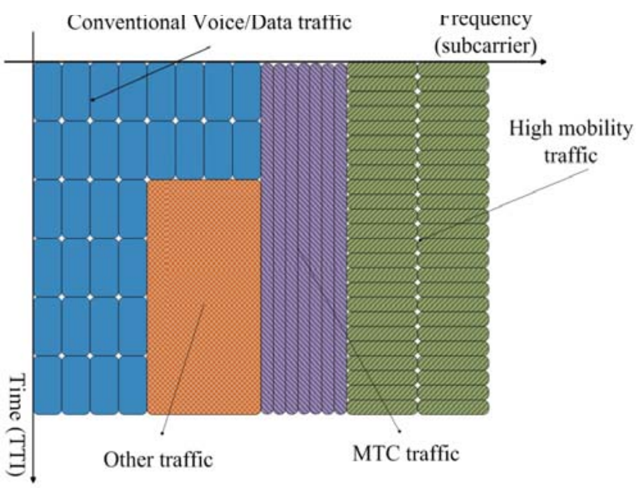


Figure 11 Flexible and adaptive frame structure supported by F-OFDM

(2) Support some specific scenarios and service types: For V2V and high-speed train, the channels may be doubly selective and significantly vary between devices because of high moving speed and complicated scattering environments among vehicles. In such cases, the new multi-carrier schemes are needed. For example, FBMC can better adapt to the characteristics of doubly selective channels by optimizing the prototype filters using real-time channel state information to achieve a better performance.

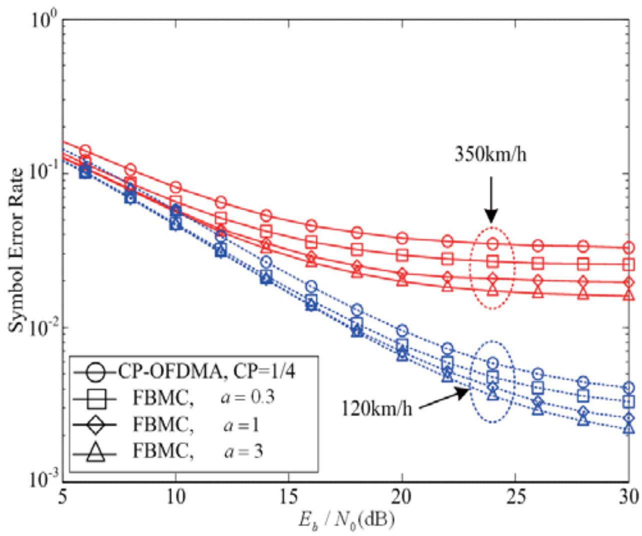


Figure 12 Symbol error rates of OFDM and FBMC in high-mobility scenarios
(a is the scaling factor of the prototype filter)

(3) Support asynchronous transmission with reduced overhead: For the proposed multi-carrier waveforms, i.e., FBMC, UFMC and F-OFDM, the OOBE is suppressed by filtering, which leads to a reduced guard band consumption. With greater separation between adjacent subbands, the requirement on global synchronization is not needed, so that the synchronization overhead can be reduced, and even inter-subband asynchronous transmission can be enabled.

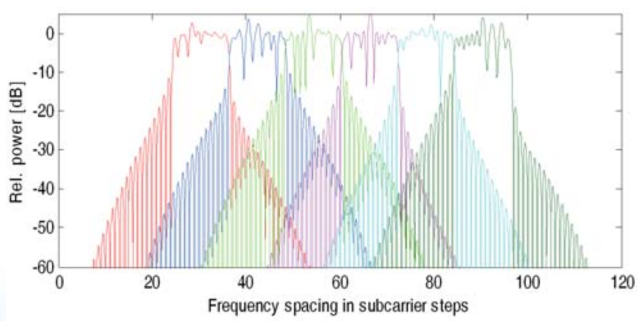


Figure 13 Illustration for the OOBE of UFMC

In summary, it is difficult for a single waveform to fulfill all the requirements of diversified application scenarios and service types of 5G. Multiple waveforms can coexist in 5G systems, each for a specific scenario. The design of new multi-carrier waveforms should start from the requirements of scenarios and services, and then select the most suitable waveform and numerology to enable the best performance for each type of services.

6. Advanced Modulation and Coding

There are a number of scenarios for 5G, each having drastically different performance requirements. For example, in high-capacity hot-spot scenario, the requirement for single user link data rate is extremely high. Hence, the system should be able to support high spectrum efficiency and long code blocks, as long as there is wide bandwidth available and the channel condition is good. In dense deployment scenario, wireless backhaul would be widely used, which drives more advanced channel coding and routing strategies to reduce the inter-node interference, to fully utilize the broadcast nature of over-the-air transmission, to meet the high capacity requirement of the systems.

There are quite a few distinct research directions in advanced modulation and coding. They can roughly be categorized into three groups: link level modulation and coding, network coding and link adaptation. Among link level coding, there are non-binary channel coding, bit mapping and joint modulation & coding. Link adaptation includes rateless codes based, and rate-compatible, and some engineering oriented coding, as illustrated in Fig.14. In non-binary codes, the link performance can get very close to Shannon's limit, via Galorie Field operation and bit interleaving. The diversity gain is also achieved in this process. New bit mapping uses co-centric ring like constellation, amplitude and phase shift key (APSK), to approach Shannon's capacity at high SNR region. Joint modulation and coding is characterized by rotated phase of constellation, which can make the link more robust in fast fading channels. Advanced link adaptation allows the modulation and channel coding to match the fast fading channel more precisely through optimizing code structure and retransmission bit distribution base on rateless codes, rate-compatible codes, and some engineering oriented codes. Network coding utilizes the broadcast nature of air interface and can make use of inter-node communications to collect the relevant bit information, in order to further improve the system capacity.

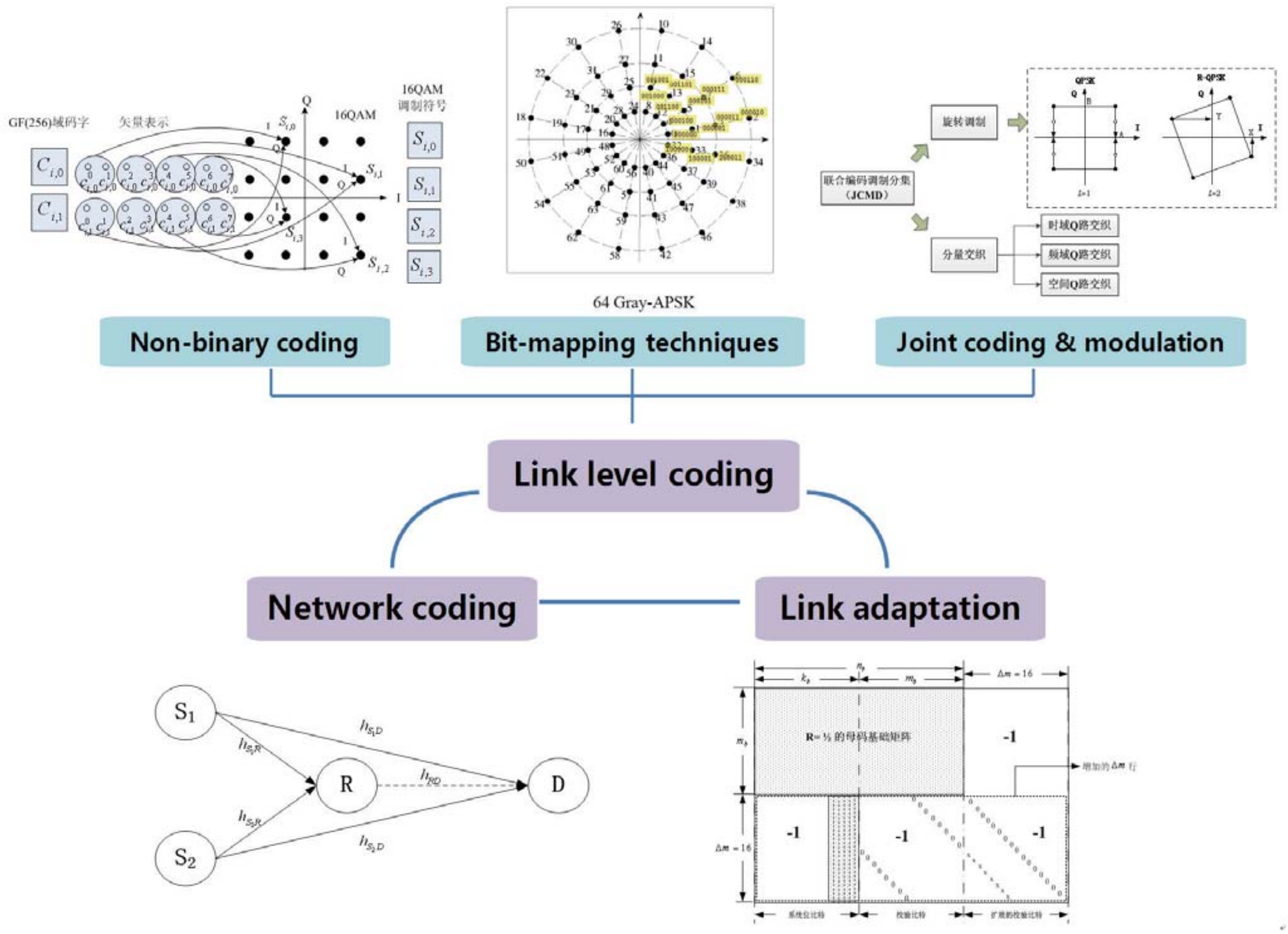


Figure 14 Categorization of advanced modulation and coding

(1) Link level modulation and coding

The design objective of non-binary codes is to get better performance, however with the comparable decoding complexity as binary codes. Currently, M-ary LDPC and non-binary repeat accumulate (RA) codes are considered as promising codes. Apart from the code structure itself, the mapping from non-binary field to constellation is another important aspect and can directly affect the performance. The new bit mapping, for example, APSK, results in the amplitude distribution being very close to Gaussian, compared to the traditional QAM. These new constellations, together with channel coding and advanced demodulation algorithms can be jointly designed. One example is FTN whose performance is very closer to Shannon's capacity. Its design should aim at reducing the complexity of sequence detection and the optimization of filters. A few sub-optimal algorithms have been proposed whose performance is slightly worse than the optimal detector, however with much reduced computational complexity. In link level codes, envelop coding provides a simple and effective way to improve the error correction capability of the channel, via XOR operation of information bits in multiple sub-codeblocks.

(2) Network coding

There are two aspects of study on network coding: codes design and system-level design. The design of network codes has the target of improving the efficiency of cooperative transmission, and closer matching with the channel conditions of each link in the entire chain. System design encompasses user pairing, selection of router, and resource scheduling. Network coding is highly coupled with deployment scenario, and detailed solutions should target certain scenarios, for example, cooperative relay or bi-directional transmission. They would have different impacts on the air interface specifications.

(3) Link adaptation

The main study on link adaptation is to design multiple choices of code rates and code block sizes. In the area of rateless codes, forward stack algorithm can significantly reduce the decoding complexity of spinal codes. In rate-compatible codes, LDPC has the potential to be designed with more flexible code rates and block sizes. In the engineering optimization, soft ACK/NACK can improved the efficiency of transmission, and better match the fast fading in the absence of accurate channel state information feedback.

The link level performance of 16-ary codes and binary LDPC is compared in Fig.15, with the assumption of AWGN channel. At 1% BER, the gain of non-binary codes is around 0.5 dB.

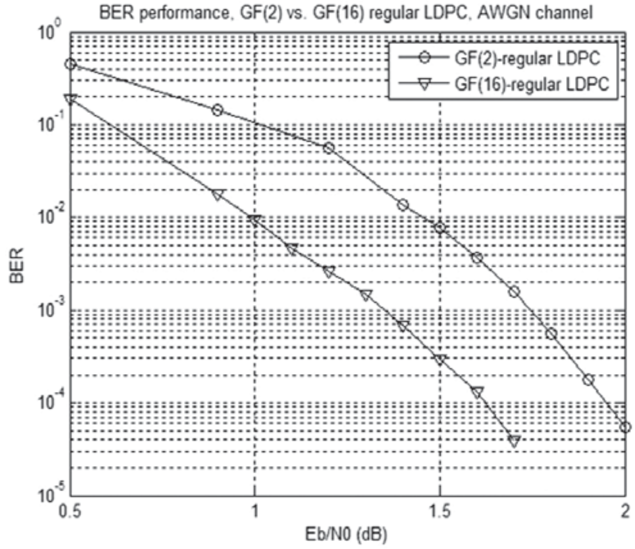


Figure 15 Link simulation comparison between 16-ary codes and binary LDPC

Fig.16 shows the performance comparison between APSK and QAM. At 0.01%, the refined APSK (NE-APSK) can improve the performance by about 0.5~0.9 dB. And the gain is even more pronounced when the modulation order gets higher.

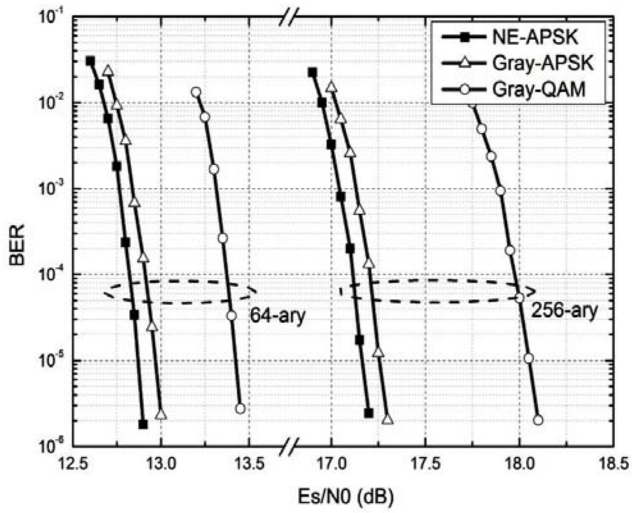


Figure 16 Link simulation comparison between APSK and QAM

The system throughput CDF is illustrated in Fig. 17 where the network coding plus the reasonable routing and user pairing can provide gains.

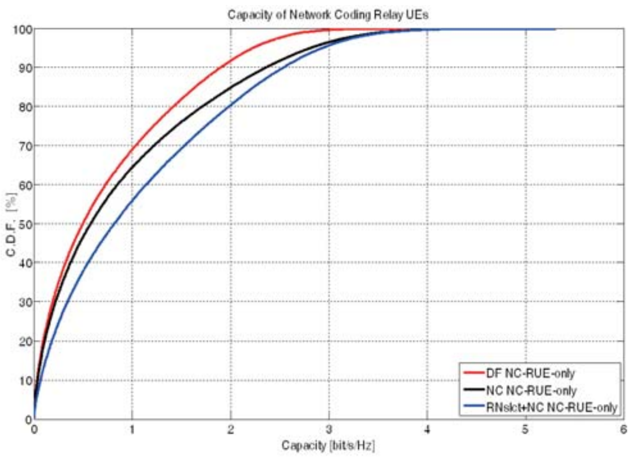


Figure 17 Gain in system throughput via network coding

Many techniques in advanced modulation and coding can be used in various scenarios. M-ary LDPC, new bit mapping and FTN are suitable for wide-band and large data transmission. They show significant gain in high SNR environment, thus also suitable for hot-spot high capacity scenario. Joint modulation and coding can improve the diversity via using multiple transmit antennas and would be used for wide coverage scenario. Network coding can improve the cooperation between multiple nodes. Hence, it would be an important technology in ultra-dense networks.

7. Device to Device (D2D)

Along with the technology development, more and more smart devices come out such as smart phones, tablets, wearable devices, smart meters, and smart vehicles. They have more powerful wireless communication capacity, supporting not only the traditional cellular communications such as 2G/3G/4G, but also Wi-Fi, Bluetooth, LTE-D2D between devices. The cooperation between cellular and direct communications may derive more use cases and improve user experience. For example, fast D2D can be applied to vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications in intelligent transportation system (ITS), multi-user cooperative communications, data sharing network, and low cost D2D for IoT devices.

Vehicle direct communication (VDC): Internet of Vehicles includes not only the remote communication between vehicle and the Internet (telematics), but also frequent V2V, V2I, and vehicle-to-pedestrian (V2P) communications (generally called V2X). Utilizing the large coverage of cellular networks, telematics services can be easily supported. The VDC solution enhanced by D2D can achieve low latency, short range, high reliability V2X services to deal with all the communication requirements of Internet of Vehicles. The VDC solution where D2D is tightly coupled with cellular networks can realize the combination between centralized scheduling and distributed communications, and therefore can meet the stringent requirements of V2X.

Multi-user cooperative communications (MUCC): In future, cooperative communications can be applied at network and terminal sides. By connecting with each other, devices in the proximity can cooperate and relay data. Each device may have several paths towards base stations. If the channel of one path is bad, a better path can always be selected so that the system throughput, communication reliability and user experience can be improved.

Data sharing network: With base station controlling or assisting, devices can establish direct data transmission ad hoc network to share data mutually. By transmitting data between each other, the network load can be decreased and the system throughput can be improved.

Low cost D2D: For latency insensitive and cost sensitive IoT systems, the hierarchical M2M access mode can be adopted. The IoT devices can connect to the cellular network via relaying. The low cost D2D communications between the end devices and relay devices, the cost of end devices can be greatly reduced, which facilitates the large scale deployment of IoT services.

The key issues and solutions include:

(1) In VDC scenarios, V2V, V2I, and V2P need to frequently broadcast in short range to exchange safety messages including location and velocity, in order to improve road safety and transportation efficiency. The V2X safety messages have stringent requirement on latency and reliability. General speaking, the latency for safety messages should be less than 100ms, and high packet delivery ratio (PDR) is needed for high reliability. With enhanced D2D communications, and centralized scheduling and resource allocation, the collision of V2X messages will be minimized and the reliability can be improved. Furthermore, the scheduling interaction process should be optimized to reduce the latency and complexity of data exchange between the network and devices.

(2) In MUCC scenarios, devices will be controlled by the network to complete pairing and establish MUCC relationship. In the process of communications, each device can be controlled, managed, and clearly charged by the network. For the protocol design, a new protocol layer is needed for the network to select the best path towards the end device according to the channel quality of different devices. The data will be sent to the device with the best channel quality, and then will be forwarded to the destination device.

(3) In data sharing network scenario, the power of devices is limited, thus the transmission range is usually short. The SNR can be improved through network coding to combine the received data from multiple links. Since devices may move all the time, the topology of devices may change frequently. Once the topology changes, the related devices need to report the information to the base station. Then, the base station can optimize the data transmission path timely to guarantee the service continuity.

(4) In low cost D2D scenario, the low cost, low power IoT end devices and relay devices need to be designed well. The low cost D2D connection can replace the direct connection between IoT end devices and cellular network, meanwhile, the network should be able to manage, control, and charge the IoT end devices with the guarantee of security. That will help large spreading for the

IoT applications such as metering and wearable devices.

8. Flexible Duplex

With the development of online video and social network service, future mobile traffic will presents varying characteristic: Especially, the ratio of DL and UL traffic may change with time and location. However, the current static allocation of the frequency/time resources for downlink and uplink transmission in LTE is not efficient to support dynamic asymmetry cell traffic. With flexible duplex, the resources could be allocated for DL or UL transmission dynamically, which could improve spectrum utilization efficiency effectively. Flexible duplex could be applied to low power small cells or relay nodes.

Flexible duplex could be realized by time or frequency domain schemes. In FDD system as shown in Fig.18, by setting different UL/DL slot ratio according to traffic demand, the UL bands can be assigned for both UL and DL transmission in the time domain scheme. In the frequency domain scheme, the UL bands are assigned as flexible bands (DL or UL bands) to adapt to the traffic demand. In TDD systems, each cell can allocate the DL/UL slot ratio based on their own service demands, even if they are located very close to each other, which is similar to the scheme applied in FDD UL bands.

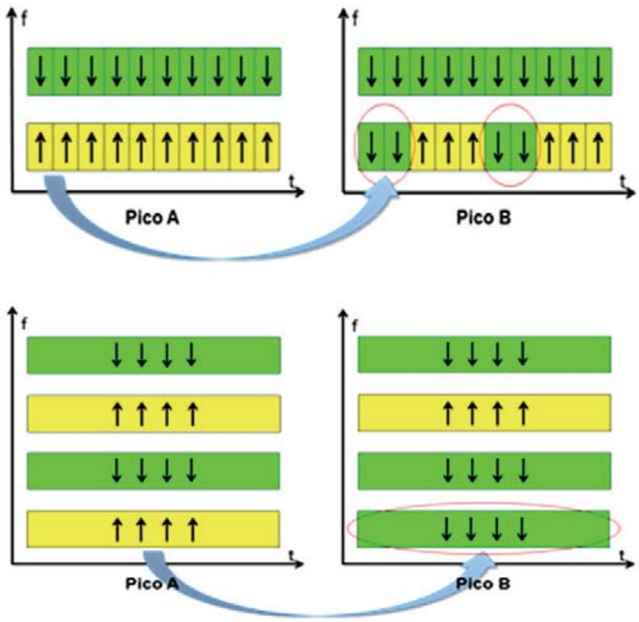


Figure 18 Flexible resource allocation in time and frequency domains

The main technical difficulty of flexible duplex is how to deal with interference signals of different transmission direction, such as DL-UL or UL-DL interference. In LTE system, the design of DL and UL signals have many difference, such as multiple access method, subcarrier mapping, reference signal pattern, which may bring difficulty to interference recognition and cancellation. Therefore, the symmetrical design for DL and UL signals is the essential point to improve system performance. In current LTE system, a uniform signal format could be realized by adjusting UL or DL signals, such as carrier moving, adjust RS pattern or RS muting method. Then, different direction signals could be recognized and cancelled. For the upcoming 5G, new bands and new multiplex access schemes are likely to be applied, which is triggering the new design of UL and DL signals. If the new UL and DL signals have the symmetrical design including design of carrier mapping and orthogonality of RS signals, the interference signals in different direction are transformed into signals in the same direction, and then it could be cancelled by current interference cancellation or interference coordination method dealing with signals in the same direction.

To further limit the interference between base stations, the transmit power level of base stations can be reduced to the level of mobile terminals. The trend of traffic allocation is that macro cells will be mainly responsible for management and control functions, while pico cells will take more mobile traffic. Therefore, flexible duplex may be more suitable for pico cells with low transmission power.

Fig.19 and Fig.20 show the cell average UL&DL throughput in dense cluster scenario. Four pico cells are located within one cluster, two with heavy UL load, and two with heavy DL load. If DL-UL or UL-DL interference signals are transformed into signals with the same direction, then the system throughput could be improved significantly. In the figure, "WO IC" corresponds to the situation before the former mentioned method is applied, and UL and DL signals have different structure, which makes it difficult for receivers to do interference cancellation. "IC-x" corresponds to the situation after the method is applied, so interferences from different direction could be cancelled by the aid of network assistant. The results show that with symmetrical design, different direction interference signals could be cancelled effectively, that's especially helpful for UL throughput improvement. As inter-cell interferences mainly come from the other two pico cells with different signal direction within the same cluster, only the interferences within the same cluster need to be considered in practice, which could help reduce the complexity and simplify the calculation in the receiver.

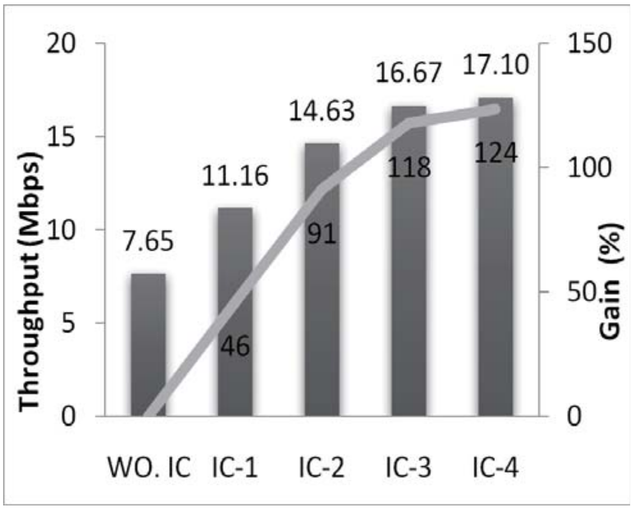


Figure 19 UL cell average throughput

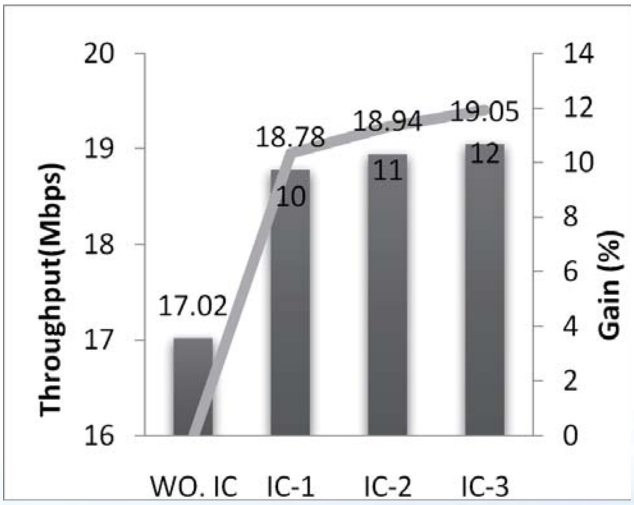


Figure 20 DL cell average throughput

Although the application of flexible duplex may be currently constrained by spectrum regulation, it is still possible to apply for modifying the regulation if the interferences between neighbor systems could be controlled and the gain of flexible duplex is widely approved. Therefore, facing the challenge of improving network performance brought by mobile broadband services, it is essential to further study the key techniques and scenarios to promote technology mature and industry upgrading.

Flexible duplex meets the trend of TDD&FDD convergence, it can enhance network efficiency and flexibility. Flexible duplex can be applied to both 5G and 4G enhancement. Furthermore, flexible duplex has a good forward compatibility and its design methods can be referred to full duplex systems.

9. Full Duplex

Revolutions in wireless communication theories and techniques have been driven by the contradiction between the ever-increasing wireless communication service demands and the limited spectrum resources. For future mobile communications, one of the key objectives is to improve the spectrum efficiency for FDD and TDD by leveraging a unified spectrum resource management strategy. The full duplex, a novel duplex architecture building upon advanced self-interference cancellation techniques, is a promising solution for achieving that objective. In theory, full duplex technology is capable of doubling the spectrum efficiency of existing communication systems.

The main research content of full duplex is divided into two aspects.

(1) Self-interference cancellation

From the physical-layer perspective, the core problem of full duplex is how to efficiently eliminate the self-interference, which is generated from the local transmitter itself, at the receiver side. To this end, there have been considerable researches in academia and industry, and it is recognized that joint self-interference cancellation from spatial domain, RF domain and digital domain is a viable way to enable full duplex. In particular, by applying joint self-interference cancellation, it is already able to achieve more than 115 dB self-interference cancellation over a bandwidth of 20 MHz. The spatial domain self-interference cancellation suppresses self-interference (SI) at the propagation stage through antennas' placement, zero-forcing beamforming and highly isolated transmit-receive antennas, etc., the RF domain self-interference cancellation cancels the SI signal at the analog domain, where an inverted SI signal in phase and amplitude is added at the RF frontend, and the digital domain self-interference cancellation further suppresses the SI by reconstructing and removing its linear and nonlinear residuals.

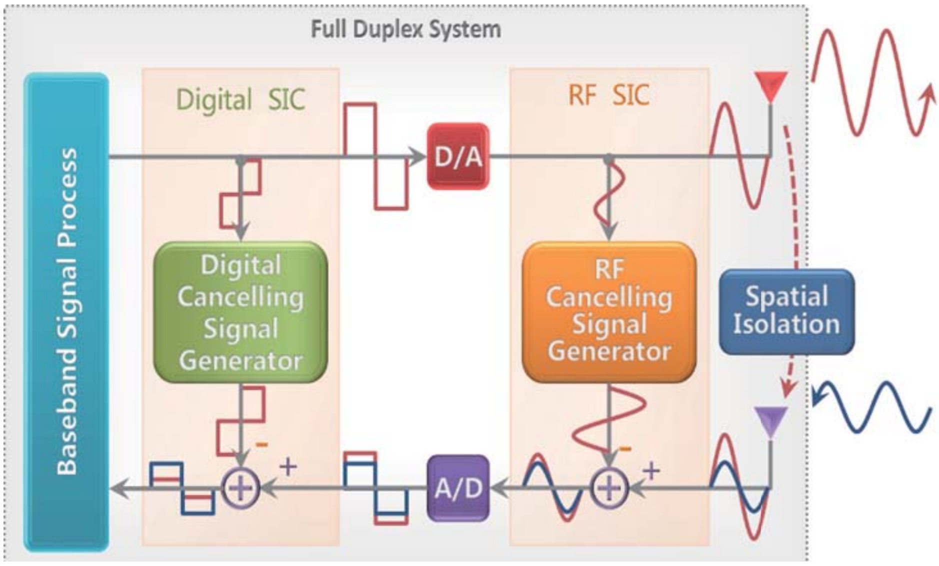
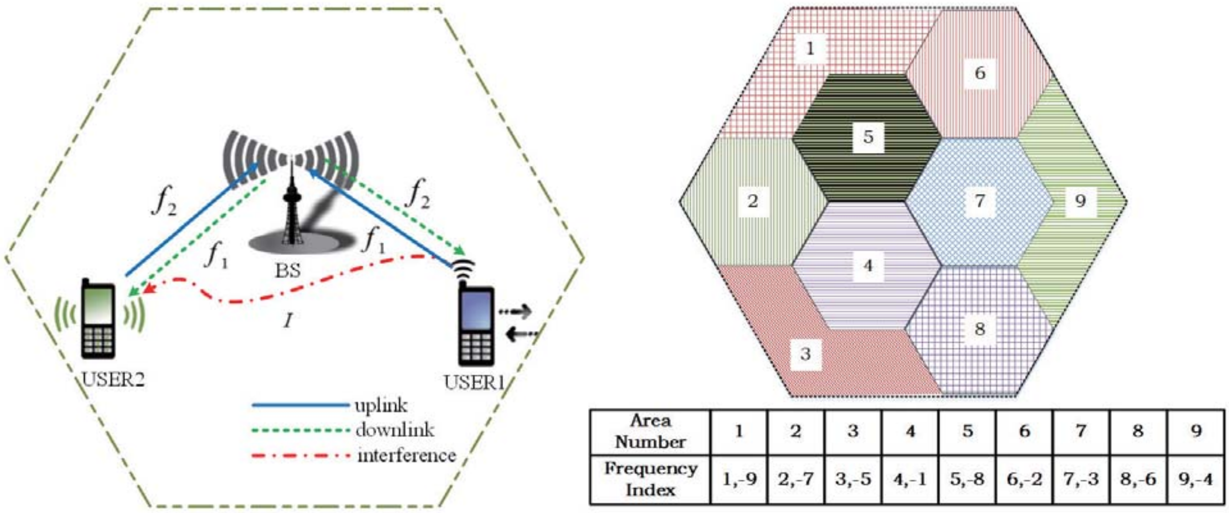


Figure 21 The principle of the full duplex

(2) Network technology in full duplex

From the network perspective, full duplex has more flexibility to control the way of transmission and reception. By changing the way of traditional spectrum usages, it is expected that full duplex will bring about more innovative techniques on access protocols and wireless resource management, etc., which in turn requires more efficiently designed network architectures. To this end, preliminary studies have been conducted by the academia and industry, with an emphasis on the following directions: the network architecture design for hybrid full-duplex/half-duplex networks, the inter-node interference coordination strategies, the resource management of full-duplex networks and the design of LTE frame structure under full-duplex mode etc.

As an illustration, let us consider a network consisting of a full-duplex base station (BS) and multiple FDD terminals. A frequency pair (f1, f2), indexed by i, is scheduled for user 1's uplink and downlink transmissions, respectively; meanwhile, user 2 is allocated on the same frequency pair by reusing the reverse links, i.e., f_2 for the uplink and f_1 for the downlink. Clearly, the frequency utilization factor under this allocation scheme is doubled as long as the inter-user interference can be effectively eliminated. To fulfill this, it is a must for BS to operate in full-duplex mode at frequencies f_1 and f_2, and meanwhile, at the user side, the cell partitioning method can be employed to mitigate the inter-user interference. Fig.22 shows an example of this partitioning method. There are totally 9 pairs of carrier frequencies (indexed by 1~9) to be used. The area is divided into 9 partitions and allocated with specific frequency pairs as depicted in Fig.22. Theoretical analysis and simulation results demonstrate that the probability of frequency reverse reuse increases with the number of users, and that the frequency utilization gain over conventional half-duplex systems is improved significantly. In addition, under the terminals' allowable interference power constraint, it is more favorable to divide the area into the smallest possible partitions (*P* in Fig. 22) to achieve higher spectrum efficiency.



Area Number	1	2	3	4	5	6	7	8	9
Frequency Index	1,-9	2,-7	3,-5	4,-1	5,-8	6,-2	7,-3	8,-6	9,-4

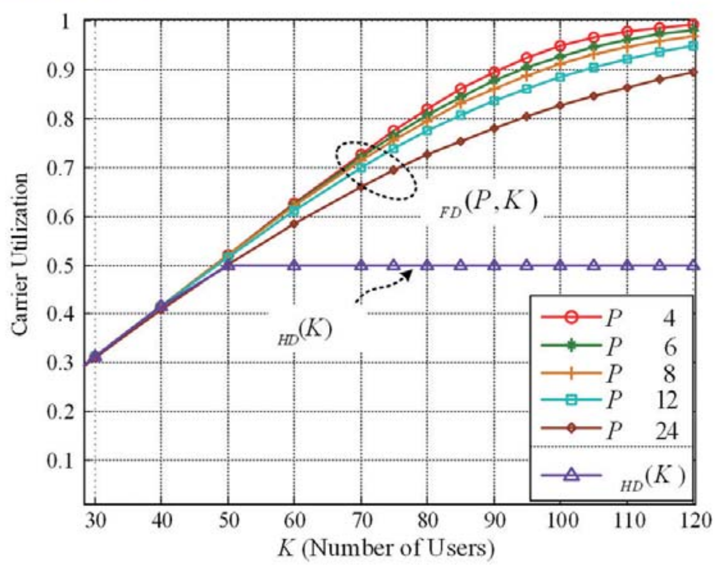


Figure 22 Simulation results

Source: Shao Shihao, Liu Donglin, Deng Kai, Pan Zhengang, Tang Youxi, "Analysis of carrier utilization in full-duplex cellular networks by dividing the co-channel interference region," IEEE Communications Letters, vol.18, no.6, June 2014, pp.1043~1046.

To put full duplex into practice, there are still many technical challenges to be addressed. Specifically, these challenges include the suppression of SI signals with high power and large dynamic range, compact circuit design for multi-antenna RF domain cancellation, novel network architectures and self-interference cancellation methods for full-duplex communications, co-existence, and evolution with the FDD/TDD half-duplex system, etc.

Overall, full duplex technology has more flexibility to transceiver and network designs, capability of eliminating the difference between FDD and TDD, and consequently the potential to improve the system spectrum efficiency. Full duplex technology is suitable for scenarios with limited and fragmented spectrum, such as the indoor, low-mobility, and low-power communications. Other more complex and advanced application scenarios will be taken into account with the maturity of the full duplex technology.

10. Spectrum Sharing Technology

To satisfy the requirements of ultra-high traffic and ultra-high data rates for 5G, the innovative spectrum utilization methods should be further investigated to extend usable frequency of IMT system, in addition to requiring more IMT dedicated spectrum. In 5G system, spectrum sharing technique can enable optimal dynamic spectrum allocation and management among different networks or systems. Furthermore, it has the self-adaption functions of autonomous access network and handover between networks. Spectrum sharing technique could achieve efficient, dynamic and flexible spectrum utilization to improve air interface efficiency and coverage which finally increase integrated frequency utilization efficiency.

For spectrum sharing technology, the prioritized scenarios include intra-operator inter-RAT spectrum sharing, inter-operator spectrum sharing, spectrum sharing in unlicensed band, and spectrum sharing in secondary access.

- Intra-operator inter-RAT spectrum sharing scenario: The target bands could be the new bands for 5G and/or the legacy IMT bands. The different RATs here could be deployed in co-located macro cells, inter-site macro cells, or macro-small cells.The existing carriers can be aggregated with spcetrum sharing carrier.
- Inter-operator spectrum sharing scenario: The target bands could be unreleased IMT planned bands. Different operators can be small cells in the same coverage of hot-spot area or macro-small cells in the same coverage.
- Spectrum sharing in unlicensed band scenario: The main target bands are unlicensed bands including 2.4GHz, 5GHz, etc. The involved station types are small cells and Wi-Fi access points in target bands. These stations can be deployed jointly with the macro cells using the existing low-frequency bands.
- Spectrum sharing in secondary access band scenario: The target bands are licensed bands for other systems such as satellite, broadcast, and radar. IMT system nodes could be managed and licensed by database or occupy frequency based on spectrum sensing solution.

The essential issues and potential solutions for spectrum sharing are as follows:

(1) Network architecture and interfaces

The network architecture needs to be modified. The centralized architecture will be a basis, which can be combined with distributed architecture. The new interfaces between new nodes and between sharing nodes need to be designed. It should be considered to add spectrum management nodes in the higher layer of sharing nodes for maintenance and management of the shared frequency resource pool, acquiring requirement application of sharing nodes and performing spectrum allocation decisions. The sharing nodes operate measurements and requirements calculation, receive the results of spectrum allocation and execute resources reconfiguration within the node. The above architecture may involve database technology, such as spectrum map generation and management, registration and authentication, channel assignment, and learning mechanism.

(2) High layer technology

The high layer technology for spectrum sharing needs to be studied to solve the spectrum dynamic change issue due to sharing and coexistence of networks in different priorities. Based on different system architectures, the study should focus on management of large fragmented frequency resources, spectrum fair negotiation and optimized allocation among multi-sharing nodes and spectrum switching, access control and cross-layer design based on the prediction and cost analysis. The related study should also analyze the impacts on the existing network access, traffic management, mobility and other operations.

(3) Physical layer technology

Physical layer technology for spectrum sharing is also in the scope of study. The possible enhancements include: obtaining usage status of spectrum based on sensing/detection methods, designs for measurement and feedback mechanisms and physical channels and reference signals, supporting results reporting, spectrum allocation and utilization, interference management based on the self-cognitive of interference, adaptation to changing of interference backgrounds, combination with cognitive network technology, and analysis of possible multiple access methods.

(4) RF technology

For RF technology, multi-mode multi-band chips are becoming a mainstream in the market. New radio design supported spectrum sharing should also be studied. This new RF design needs to support broader range of bands/ frequency, suppress interference efficiently under multi-channel simultaneously operation, support flexible bandwidth of RF and modulation when access to different systems at the same spectrum, and support broadband spectrum sensing through multi-channel detection or compressed sensing, etc., to identify the RF reference structure and parameters.

The other technical issues include coexistence issue between shared spectrum systems, network architecture design under each scenario, and security of multi-system integration.

Initial simulation for inter-operator spectrum sharing is evaluated under indoor small cell dense deployment scenario. Some evaluation assumptions include:

- Contiguous 180 MHz bandwidth in 3-6 GHz frequency band for spectrum sharing between operators.
- Evaluation scenario is the indoor hot-spot dual stripe scenario (6 floors) defined in 3GPP TR36.814.
- The traffic model is burst traffic, where the packet size is 8 Mbytes and the packet arrival rate of 0.5-2.5.

Two spectrum allocation methods, including dedicated allocation method and spectrum sharing method, are compared in the evaluation. For the dedicated allocation method, each operators individually occupies 60 MHz bandwidth. For the spectrum sharing method, three operators share the whole 180MHz spectrum, where users select subscribed operator to access.

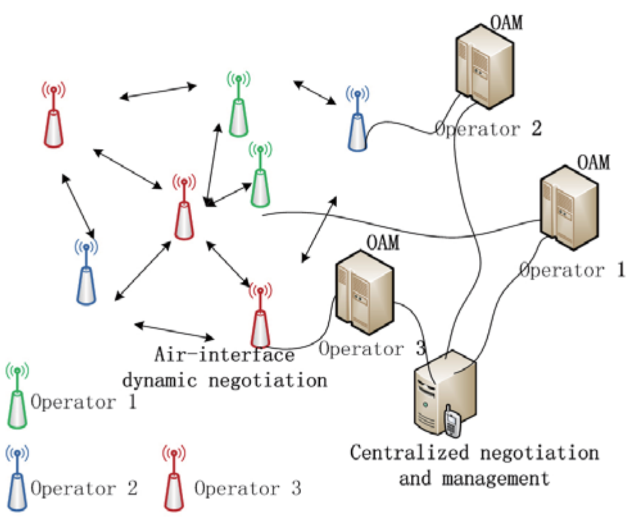


Figure 21 Topology of scenario

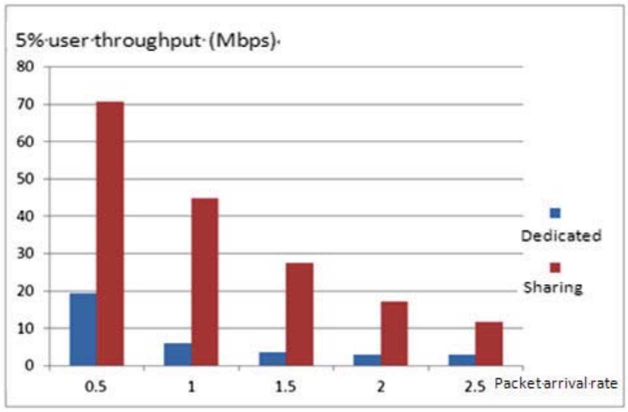


Figure 22 Simulation results

From the simulation results, the throughput of spectrum sharing is much higher than the sum of all the operator's throughputs of dedicated allocation. Especially for edge users, the usable spectrum of these users is increased and the user experienced rate can be improved 2-7 times higher. When sharing 180MHz bandwidth, edge user experience can still reach 12 Mbps under 90% load.

Spectrum sharing technology can promote new spectrum management methods and improve the existing network capacity by addressing some technical issues such as system architectures, interfaces, air interface technology, and interference management under different key scenarios (intra-operator, inter-operator, unlicensed bands accessing and secondary accessing). It is compatible to the existing technologies including carrier aggregation, database, and radio resource management. In the meanwhile, it is able to fully integrate with the other key technologies of 5G, e.g., supporting ultra-dense network based on dynamic spectrum management, jointly utilizing the existing IMT bands and higher frequency bands to meet the high spectrum demand of 5G. Thus, it could enhance the total spectrum efficiency of different operators and solve the load balancing issues of multi-RAT and multi-cell. The user experienced rate could also be improved. The available IMT spectrum can be expanded when the over-load and/or over-idle of operators spectrum may be relieved. Furthermore, spectrum sharing can significantly contribute to the performance of 5G in wide-area and hotspot scenario. However, spectrum sharing techniques require support of spectrum regulatory policy which needs to establish new usage rules, security strategy, and economic models, and also have higher requirements on baseband algorithms and devices capabilities. From the current international research trends worldwide, spectrum sharing is one of key technologies for 5G system with a good application prospect.

Conclusion

Based on a unified air interface technical framework, 5G will develop through the new air interface (including low-frequency and high-frequency branches) and 4G evolution air interface, and will be powered by a group of key technologies such as novel multiple access, massive MIMO, ultra-dense networking, and all-spectrum access. By flexibly configuring technical modules and parameters, the optimized technical solutions can be derived for specific scenarios including seamless wide-area coverage, high-capacity hot-spot, low-power massive-connection, and low-latency high-reliability, to fully meet the requirements of mobile internet and IoT in year 2020 and beyond.

IMT-2020 (5G) Promotion Group is willing to strengthen cooperation with global organizations, enterprises, research institutes, and universities to define a globally recognized 5G technology architecture and promote the development of 5G standards and industry.

Main Contributors

Contacts
Tel：+86-10-62300182
Email：imt2020@catr.cn