

An Overview of Next-Generation Mobile WiMAX Technology

Sassan Ahmadi, Intel Corporation

ABSTRACT

The growing demand for mobile Internet and wireless multimedia applications has motivated the development of broadband wireless-access technologies in recent years. Mobile WiMAX has enabled convergence of mobile and fixed broadband networks through a common wide-area radio-access technology and flexible network architecture. Since January 2007, the IEEE 802.16 Working Group has been developing a new amendment of the IEEE 802.16 standard (i.e., IEEE 802.16m) as an advanced air interface to meet the requirements of ITU-R/IMT-advanced for 4G systems, as well as for the next-generation mobile network operators.

Depending on the available bandwidth and multi-antenna mode, the next-generation mobile WiMAX will be capable of over-the-air data-transfer rates in excess of 1 Gb/s and of supporting a wide range of high-quality and high-capacity IP-based services and applications while maintaining full backward compatibility with the existing mobile WiMAX systems to preserve investments and continuing to support first-generation products. This tutorial describes the prominent technical features of IEEE 802.16m and the potential for successful deployment of the next generation of mobile WiMAX in 2011+.

INTRODUCTION

The growing demand for mobile Internet and wireless multimedia applications has motivated the development of broadband wireless-access technologies in recent years. Mobile WiMAX was the first mobile broadband wireless-access solution based on the IEEE 802.16e-2005 standard [1] that enabled convergence of mobile and fixed broadband networks through a common wide-area radio-access technology and flexible network architecture. The mobile WiMAX air interface utilizes orthogonal frequency division multiple access (OFDMA) as the preferred multiple-access method in the downlink (DL) and uplink (UL) for improved multipath performance and bandwidth scalability. Since January 2007, the IEEE 802.16 Working Group has

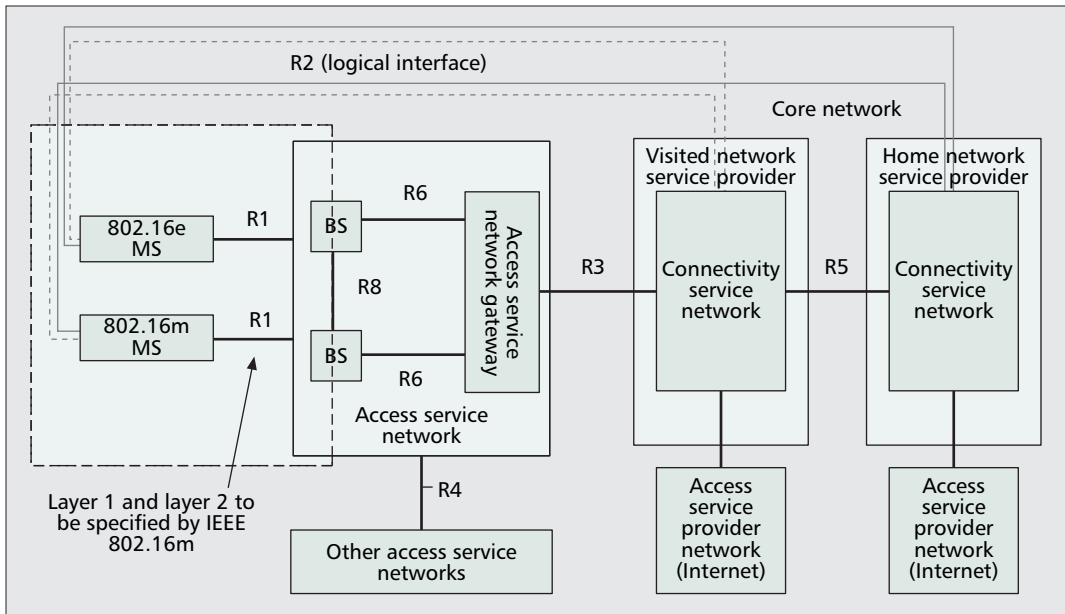
embarked on the development of a new amendment of the IEEE 802.16 standard (i.e., IEEE 802.16m) as an advanced air interface to meet the requirements of the International Telecommunication Union — Radiocommunication/International Mobile Telecommunications (ITU-R/IMT)-advanced for fourth-generation (4G) systems, as well as the next-generation mobile network operators.

Depending on the available bandwidth and multi-antenna mode, the next-generation mobile WiMAX will be capable of over-the-air data-transfer rates in excess of 1 Gb/s and support a wide range of high-quality and high-capacity IP-based services and applications while maintaining full backward compatibility with the existing mobile WiMAX systems to preserve investments and continuing to support first-generation products. There are distinctive features and advantages such as flexibility and the extensibility of its physical and medium-access-layer protocols that make mobile WiMAX and its evolution more attractive and more suitable for the realization of ubiquitous mobile Internet access.

The next-generation mobile WiMAX will build on the success of the existing WiMAX technology and its time-to-market advantage over other mobile broadband wireless access technologies. In fact, all OFDM-based, mobile broadband access technologies that have been developed lately exploit, enhance, and expand fundamental concepts that were originally utilized in mobile WiMAX.

The IEEE 802.16m will be suitable for both green-field and mixed deployments with legacy mobile stations (MSs) and base stations (BSs). The backward compatibility feature will allow smooth upgrades and an evolution path for the existing deployments. It will enable roaming and seamless connectivity across IMT-advanced and IMT-2000 systems through the use of appropriate interworking functions. In addition, the IEEE 802.16m system utilizes multi-hop relay architectures for improved coverage and performance. This article briefly describes the salient technical features of IEEE 802.16m and the potential for successful deployment of the next generation of mobile WiMAX in 2011+.

The evaluation scenarios and associated parameters specified in IEEE 802.16m evaluation methodology document are based mainly on those that were used for the evaluation of the reference system to benchmark the relative improvements.



■ Figure 1. Mobile WiMAX network reference model [9].

IEEE 802.16M SYSTEM REQUIREMENTS AND EVALUATION METHODOLOGY

Full backward compatibility and interoperability with the reference system is required for IEEE 802.16m systems, although the network operator can disable legacy support in green-field deployments. The reference system is defined as a system that is compliant with a subset of the IEEE 802.16e-2005 [1] features as specified by the WiMAX Forum mobile system profile, Release 1.0 [2]. The backward compatibility feature ensures a smooth migration from legacy to new systems without any significant impact on the performance of the legacy systems, as long as they exist. Furthermore, the requirements for IEEE 802.16m were selected to ensure competitiveness with the emerging 4G radio-access technologies while improving and extending existing functionalities of the reference system.

The IMT-advanced requirements defined and approved by ITU-R/Working Party 5D and published as Report ITU-R M.2134 [3] are referred to as target requirements in the IEEE 802.16m system requirement document and will be evaluated based on the methodology and guidelines specified by Report ITU-R M.2135 [4]. The baseline performance requirements will be evaluated according to the IEEE 802.16m evaluation methodology document [5]. A careful examination of the IMT-advanced requirements reveals that they are a subset of, and less stringent than, the IEEE 802.16m system requirements; therefore, the IEEE 802.16m standard can qualify as an IMT-advanced technology. Table 1 summarizes the IEEE 802.16m baseline system requirements and the corresponding requirements specified by Report ITU-R M.2134 [3]. In the next sections, we briefly discuss how these requirements can be met or exceeded through the extension and enhancements of the legacy-system functional features.

The IEEE 802.16m systems are expected to meet all service requirements for IMT-advanced systems as specified by Report ITU-R M.1822 [7], where compliance is verified through inspection [4].

The IEEE 802.16m evaluation methodology document [5] provides simulation parameters and guidelines for evaluation of the candidate proposals against the IEEE 802.16m system requirements [6]. The evaluation scenarios and associated parameters specified in IEEE 802.16m evaluation methodology document are based mainly on those that were used for the evaluation of the reference system [8] to benchmark the relative improvements. There are similarities and differences between the evaluation guidelines, test scenarios, and configuration parameters specified by Report ITU-R M.2135 [4] and the IEEE 802.16m evaluation methodology document [5]. Compliance with Report ITU-R M.2134 requirements in at least three test environments is required.

NETWORK REFERENCE MODEL AND IEEE 802.16M PROTOCOL STRUCTURE

The WiMAX Network Architecture Release 1.0 [9] specifies a non-hierarchical, end-to-end network reference model (Fig. 1) for mobile WiMAX that can be expanded to further include optional relay entities for coverage and performance enhancement in future releases.

The IEEE 802.16 standard [1] describes medium-access-control (MAC) and physical-layer (PHY) protocols for fixed and mobile broadband wireless-access systems. The MAC and PHY functions can be classified into three categories, namely, data plane, control plane, and management plane. The data plane comprises functions in the data processing path such as header compression, as well as MAC

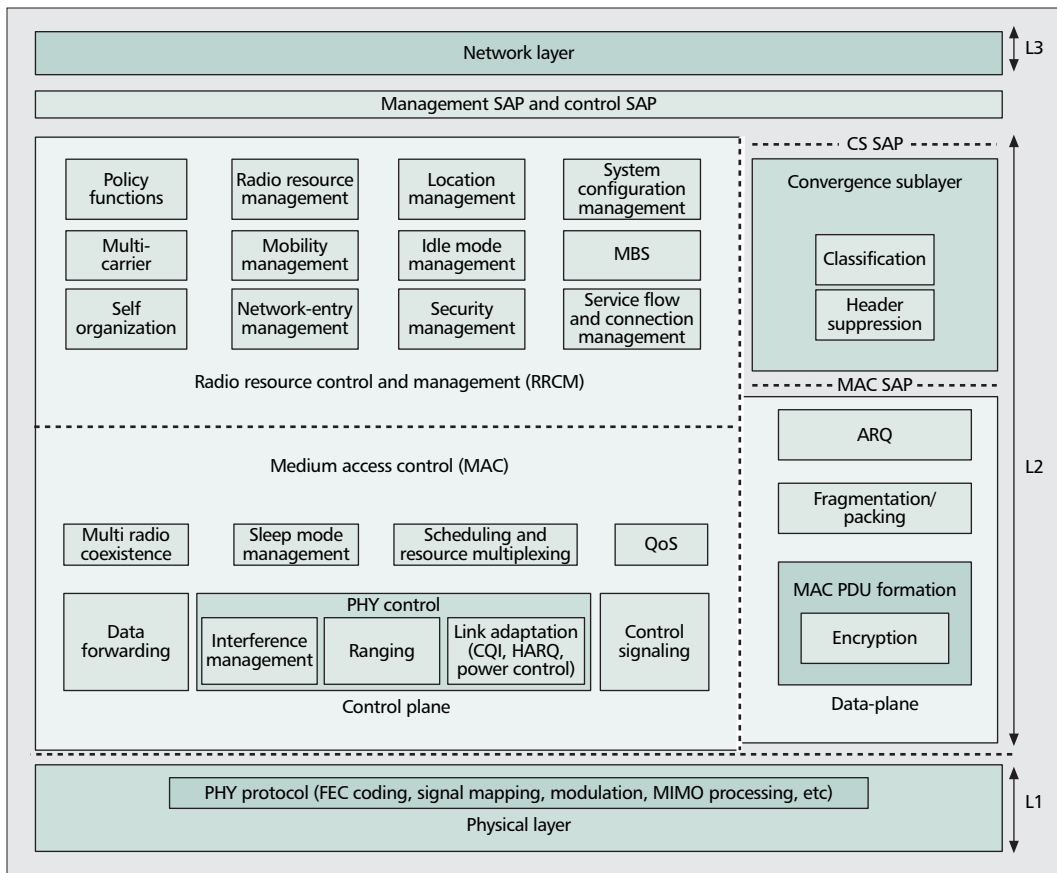
Requirements	IMT-Advanced [3]	IEEE 802.16m [6]
Peak data rate (b/s/Hz)	DL: 15 (4 × 4) UL: 6.75 (2 × 4)	DL: 8.0/15.0 (2 × 2/4 × 4) UL: 2.8/6.75 (1 × 2/2 × 4)
Cell spectral efficiency (b/s/Hz/sector)	DL (4 × 2) = 2.2 UL (2 × 4) = 1.4 (base coverage urban)	DL (2 × 2) = 2.6 UL (1 × 2) = 1.3 (mixed mobility)
Cell-edge user spectral efficiency (b/s/Hz)	DL (4 × 2) = 0.06 UL (2 × 4) = 0.03 (base coverage urban)	DL (2 × 2) = 0.09 UL (1 × 2) = 0.05 (mixed mobility)
Latency	C-plane: 100 ms (idle to active) U-plane: 10 ms	C-plane: 100 ms (idle to active) U-plane: 10 ms
Mobility (b/s/Hz at km/h)	0.55 at 120 km/h 0.25 at 350 km/h (link-level)	Optimal performance up to 10 km/h Graceful degradation up to 120 km/h Connectivity up to 350 km/h Up to 500 km/h depending on operating frequency
Handover interruption time (ms)	Intrafrequency: 27.5 Interfrequency: 40 (in a frequency band) 60 (between frequency bands)	Intrafrequency: 27.5 Interfrequency: 40 (in a frequency band) 60 (between frequency bands)
VoIP capacity (Active users/sector/MHz)	40 (4 × 2 and 2 × 4) (Base coverage urban)	60 (DL 2 × 2 and UL 1 × 2)
Antenna configuration	Not specified	DL: 2 × 2 (baseline), 2 × 4, 4 × 2, 4 × 4, 8 × 8 UL: 1 × 2 (baseline), 1 × 4, 2 × 4, 4 × 4
Cell range and coverage	Not specified	Up to 100 km Optimal performance up to 5 km
Multicast and broadcast service (MBS)	Not specified	4 b/s/Hz for ISD 0.5 km 2 b/s/Hz for ISD 1.5 km
MBS channel reselection interruption time	Not specified	1.0 s (intrafrequency) 1.5 s (interfrequency)
Location-based services (LBS)	Not specified	Location determination latency < 30 s MS-based position determination accuracy < 50 m Network-based position determination accuracy < 100 m
Operating bandwidth	Up to 40 MHz (with band aggregation)	5 to 20 MHz (up to 100 MHz through band aggregation)
Duplex scheme	Not specified	TDD, FDD (support for H-FDD terminals)
Operating frequencies (MHz)	IMT bands 450–470 698–960 1710–2025 2110–2200 2300–2400 2500–2690 3400–3600	IMT bands 450–470 698–960 1710–2025 2110–2200 2300–2400 2500–2690 3400–3600

■ **Table 1.** IEEE 802.16m and IMT-Advanced system requirements.

and PHY data packet-processing functions. A set of layer-2 (L2) control functions is required to support various radio resource configuration, coordination, signaling, and management. This set of functions is collectively referred to as the control-plane functions. A management plane also is defined for external management and

system configuration. Therefore, all management entities fall into the management-plane category.

The IEEE 802.16 MAC layer is composed of two sublayers: the convergence sublayer (CS) and the MAC common-part sublayer (MAC CPS) [1]. For convenience, we logically classify



■ **Figure 2.** IEEE 802.16m protocol stack [10].

The IEEE 802.16 MAC layer is composed of two sublayers: the convergence sublayer (CS) and the MAC common-part sublayer (MAC CPS). For convenience, we logically classify MAC CPS functions into two groups based on their characteristics.

MAC CPS functions into two groups based on their characteristics as shown in Fig. 2.

The upper and lower classes are called the resource control and management functional group and the MAC functional group, respectively. The control-plane functions and data-plane functions also are classified separately. As shown in Fig. 2, the radio-resource control and management functional group comprises several functional blocks including:

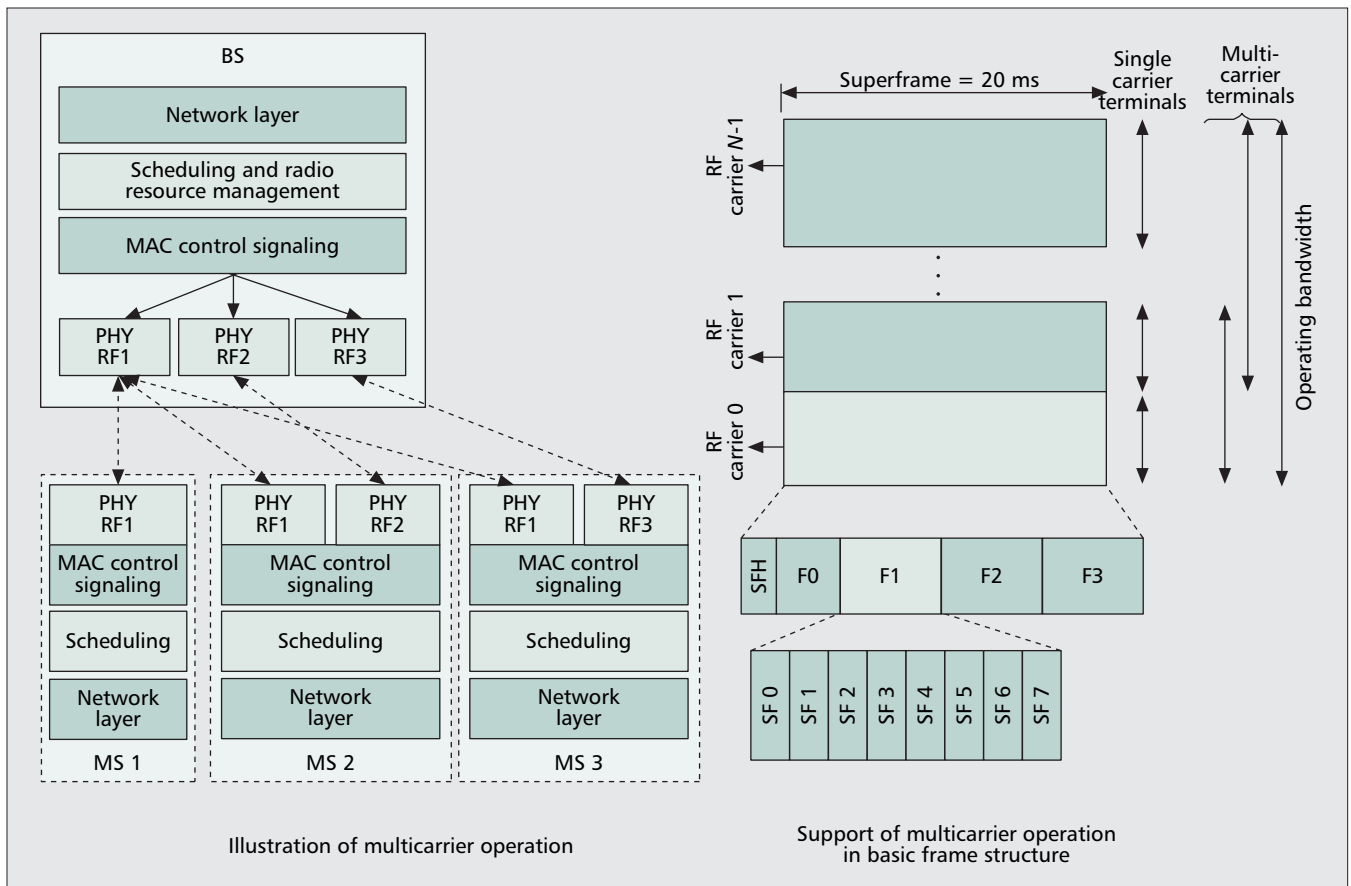
- **Radio-resource management:** This block adjusts radio network parameters related to the traffic load and also includes the functions of load control (load balancing), admission control, and interference control.
- **Mobility management:** This block scans neighbor BSs and decides whether an MS should perform a handover operation.
- **Network-entry management:** This block controls initialization and access procedures and generates management messages during initialization and access procedures.
- **Location management:** This block supports location-based service (LBS), generates messages including the LBS information, and manages the location-update operation during idle mode.
- **Idle-mode management:** This block controls idle-mode operation and generates the paging-advertisement message, based on a paging message from the paging controller in the core network.
- **Security management:** This block performs key management for secure communica-

tion. Using a managed key, traffic encryption/decryption and authentication are performed.

- **System configuration management:** This block manages system-configuration parameters and generates broadcast-control messages, such as a DL/UL channel descriptor.
- **Multicast and broadcast service (MBS):** This block controls and generates management messages and data associated with the MBS.
- **Connection management:** This block allocates connection identifiers (CIDs) during initialization/handover service-flow creation procedures; interacts with the convergence sublayer to classify MAC service data units (MSDUs) from upper layers; and maps MSDUs into a particular transport connection.

The MAC functional group includes functional blocks that are related to physical layer and link controls such as:

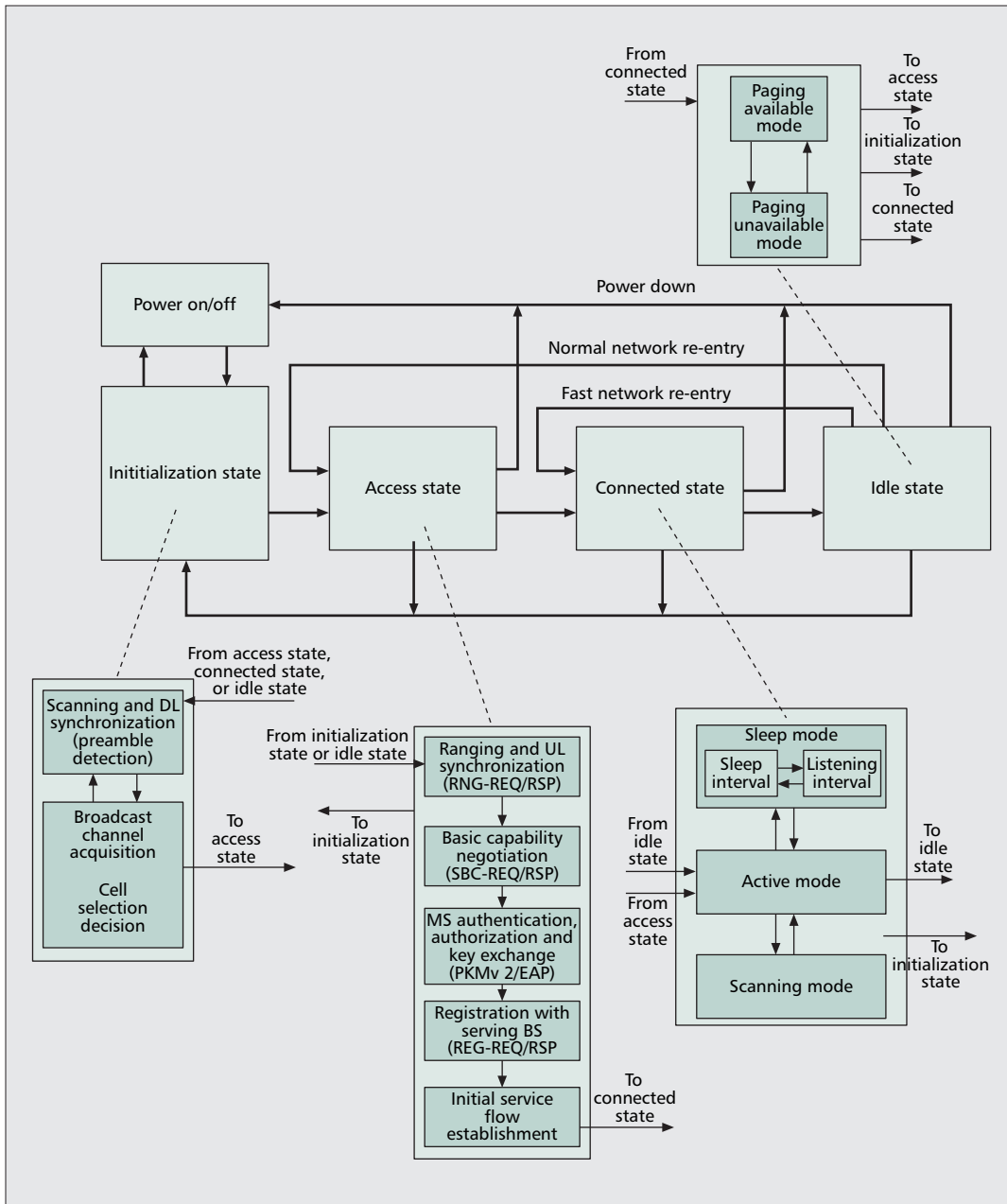
- **PHY control:** This block performs PHY signaling such as ranging, channel quality measurement/feedback (CQI), and hybrid automatic repeat request (HARQ) acknowledgment (ACK) or negative acknowledgment (NACK) signaling.
- **Control signaling:** This block generates resource-allocation messages such as DL/UL medium-access protocol (MAP), as well as specific control signaling messages, and other signaling messages not in the form of general MAC messages (e.g., a DL frame control header).



■ **Figure 3.** IEEE 802.16m multicarrier protocol stack and frame structure [10].

- **Sleep mode management:** This block handles sleep mode operation and generates management messages related to sleep operation and can communicate with the scheduler block to operate properly according to the sleep period.
 - **Quality-of-service (QoS):** This block performs rate control based on QoS input parameters from the connection management function for each connection.
 - **Scheduling and resource multiplexing:** This block schedules and multiplexes packets based on the properties of the connections.
 - **Automatic repeat request (ARQ):** This block performs the MAC ARQ function. For ARQ-enabled connections, the ARQ block splits MSDUs logically and sequences logical ARQ blocks.
 - **Fragmentation/packing:** This block performs the fragmentation or packing of MSDUs based on input from the scheduler block.
 - **MAC PDU formation:** This block constructs MAC protocol data units (PDUs) so that a BS/MS can transmit user traffic or management messages via PHY channels.
- The IEEE 802.16m protocol structure is similar to that of IEEE 802.16 with additional functional blocks for new features including the following:
- **Relay functions:** Relay functionality and packet routing in relay networks
 - **Self-organization and self-optimization functions:** a plug-and-play form of operation for an indoor BS (i.e., a femtocell)
- **Multi-carrier functions:** Control and operation of a number of adjacent or non-adjacent radio-frequency (RF) carriers where the RF carriers can be assigned to unicast and/or multicast and broadcast services. A single MAC instantiation is used to control several physical layers. If the MS supports multi-carrier operation, it can receive control and signaling, broadcast, and synchronization channels through a primary carrier, and traffic assignments can be made on the secondary carriers. A generalization of the protocol structure for multi-carrier support using a single MAC instantiation is shown in Fig. 3. The load-balancing functions and the RF-carrier mapping and control are performed by the radio-resource control and management functional class. From the perspective of an MS, the carriers utilized in a multi-carrier system can be divided into two categories:
 - A primary RF carrier is the carrier that is used by the BS and the MS to exchange traffic and full PHY/MAC control information.
 - A secondary RF carrier is an additional carrier that the BS may use for traffic allocations for mobile stations capable of multi-carrier support.
 Based on the primary and/or secondary usage, the carriers of a multi-carrier system can be configured differently as follows:

The MS state diagram for IEEE 802.16m is similar to that of the reference system with the exception of the initialization state that was simplified to reduce the scan latency and to enable fast cell selection or reselection.



■ Figure 4. Mobile station state diagram [10].

–Fully configured carrier: A carrier for which all control channels including synchronization, broadcast, multicast, and unicast control signaling are configured. The information and parameters related to multi-carrier operation and the other carriers also can be included in the control channels.

–Partially configured carrier: A carrier with only essential control-channel configuration to support traffic exchanges during multi-carrier operation.

If the user-terminal RF front end and/or its baseband is not capable of processing more than one RF carrier simultaneously, the user terminal may be allowed, in certain intervals, to monitor secondary RF carriers and to resume monitoring of the primary carrier prior to transmission of the synchro-

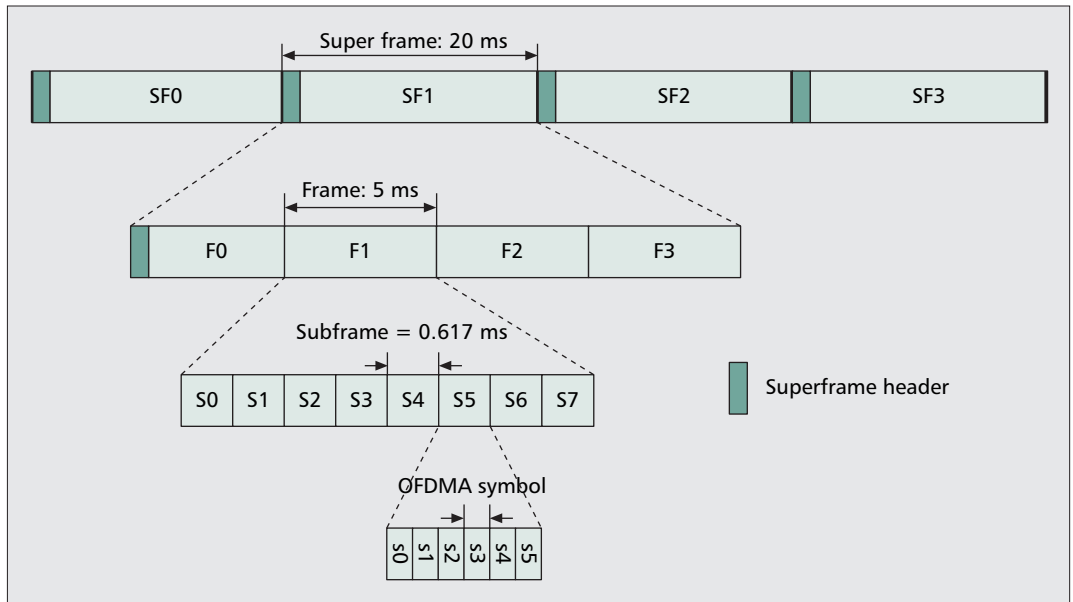
nization, broadcast, and non-user-specific control channels.

- Multi-radio coexistence functions: Protocols for multi-radio coexistence, where the MS generates management messages to report the information about its co-located radio activities obtained from the inter-radio interface, and the BS responds with the corresponding messages to support multi-radio coexistence operation.

IEEE 802.16M MOBILE STATION STATE DIAGRAM

A mobile state diagram (i.e., a set of states and procedures between which the MS transits when operating in the system to receive and transmit data) for the reference system, based on a com-

The superframe is a new concept introduced in IEEE 802.16m, where a superframe is a collection of consecutive, equally-sized radio frames, where the beginning is marked with a superframe header. The superframe header carries short-term and long-term system-configuration information.



■ Figure 5. IEEE basic frame structure [10].

mon understanding of its behavior, can be established as follows (Fig. 4):

- Initialization state: A state where an MS without any connection performs cell selection by scanning and synchronizing to a BS preamble and acquires the system-configuration information through the DL and UL channel descriptors.

- Access state: A state where the MS performs network entry to the selected BS. The MS performs the initial ranging process to obtain UL synchronization. Next, the MS performs basic capability negotiation with the BS. Then, the MS performs the authentication and authorization procedure. Next, the MS performs the registration process. The mobile station receives user identification as part of the access-state procedures. The IP address assignment can follow using appropriate procedures.

- Connected state: A state consisting of the following modes: sleep mode, active mode, and scanning mode. During the connected state, the MS maintains at least one connection as established during the access state, while the MS and BS can establish additional transport connections. In addition, to reduce power consumption of the MS, the MS or the BS can request a transition to sleep mode. Also, the MS can scan neighbor BSs to reselect a cell that provides more robust and reliable services.

- Idle state: A state comprising two separate modes, paging-available mode and paging-unavailable mode. During the idle state, the MS can attempt power saving by switching between paging-available mode and paging-unavailable mode. In the paging-available mode, the MS can be paged by the BS. If the MS is paged, it transitions to the access state for its network re-entry. The MS performs the location update procedure during the idle state.

The MS state diagram for IEEE 802.16m is similar to that of the reference system with the exception of the initialization state that was simplified to reduce the scan latency and to enable

fast cell selection or reselection. The location of the system-configuration information was fixed so that upon successful DL synchronization, the system-configuration information can be acquired; this enables the MS to make a decision for attachment to the BS without acquiring and decoding a DL MAP and waiting for the acquisition of the system parameters, resulting in power saving in the MS due to the reduction and simplification of the initialization procedure.

OVERVIEW OF IEEE 802.16M PHY MULTIPLE-ACCESS SCHEMES

IEEE 802.16m uses OFDMA as the multiple-access scheme in the DL and UL. It further supports both time-division duplex (TDD) and frequency-division duplex (FDD) schemes including the half-duplex FDD (HFDD) operation of the mobile stations in the FDD networks. The frame structure attributes and baseband processing are common for both duplex schemes.

FRAME STRUCTURE

The superframe is a new concept introduced in IEEE 802.16m, where a superframe is a collection of consecutive, equally-sized radio frames, where the beginning is marked with a superframe header. The superframe header carries short-term and long-term system-configuration information.

To decrease the air-link access latency, the radio frames are further divided into a number of subframes where each subframe comprises an integer number of OFDMA symbols. The transmission time interval is defined as the transmission latency over the air-link and is equal to a multiple of subframe length (default one subframe). There are three types of subframes depending on the size of the cyclic prefix:

- Type-1 subframe, which consists of six OFDMA symbols
- Type-2 subframe, which consists of seven OFDMA symbols

MCS index	Modulation	Code rate
0000	QPSK	31/256
0001	QPSK	48/256
0010	QPSK	71/256
0011	QPSK	101/256
0100	QPSK	135/256
0101	QPSK	171/256
0110	16QAM	102/256
0111	16QAM	128/256
1000	16QAM	155/256
1001	16QAM	184/256
1010	64QAM	135/256
1011	64QAM	157/256
1100	64QAM	181/256
1101	64QAM	205/256
1110	64QAM	225/256
1111	64QAM	237/256

■ **Table 2.** MCS table for downlink and uplink data channels [10].

- Type-3 subframe, which consists of five OFDMA symbols

In all of the subframe types, some of the symbols can be idle symbols.

In the basic frame structure, the superframe length is 20 ms (comprising four radio frames), radio frame size is 5 ms (comprising eight subframes), and subframe length is 0.617 ms. The use of the subframe concept with the latter parameter set would reduce the one-way air-link access latency from 18.5 ms (corresponding to the reference system) to less than 5 ms [10].

The concept of time zones that is applied to both TDD and FDD systems was introduced in IEEE 802.16m. The new and legacy time zones are time-division multiplexed across the time domain for the DL. For UL transmissions, both time- and frequency-division multiplex approaches are supported for the multiplexing of legacy and new terminals. The non-backward compatible improvements and features are restricted to the new zones. All backward compatible features and functions are used in the legacy zones. In the absence of a legacy system, the legacy zones disappear, and the entire frame is allocated to the new zones.

The legacy and new radio frames are offset by a fixed number of subframes to accommodate new features such as the IEEE 802.16m preambles, superframe header (system-configuration information), and control channels [10].

Multiple-RF carriers can be accommodated with the same frame structure that is used for single-carrier operation. All RF carriers are time aligned at the frame, subframe, and symbol level (Fig. 3). Alternative frame structures for CP = 1/16 and CP = 1/4 are used, which incorporate a different number of OFDMA symbols per subframe or a different number of subframes per frame [10].

PHYSICAL AND LOGICAL RESOURCE BLOCKS

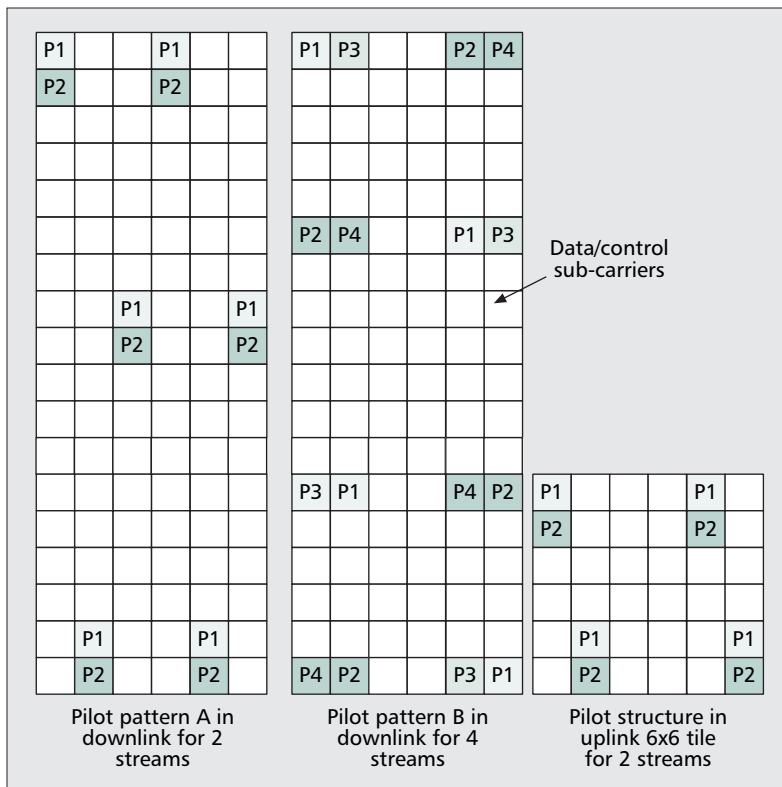
A physical resource unit is the basic physical unit for resource allocation that comprises 18 contiguous subcarriers by six contiguous OFDMA symbols. A logical resource unit is the basic logical unit for distributed and localized resource allocations. A logical resource unit comprises 18 × 6 subcarriers occupying a 196.88-kHz bandwidth in the frequency domain.

Distributed resource units are used to achieve frequency diversity gain. A distributed resource unit contains a group of subcarriers that are spread across a frequency partition. The size of the distributed resource units is equal to that of a physical resource unit. Localized resource units are used to achieve a frequency-selective scheduling gain. A localized resource unit comprises a group of subcarriers that are contiguous across frequency. The size of the localized resource units is equal to that of the physical resource units. To form distributed and localized resource units, the subcarriers over an OFDMA symbol are partitioned into guard and used subcarriers. The direct current (DC) subcarrier is not used. The used subcarriers are divided into physical resource units. Each physical resource unit contains pilot and data subcarriers. The number of used pilot and data subcarriers depends on multiple-input multiple-output (MIMO) mode, rank and number of multiplexed MSs, as well as the number of symbols within a subframe.

An outer permutation is applied to the physical resource units in the groups of physical resource units depending on system bandwidth. The direct mapping of outer permutation is exclusively applied to localized allocations. The permuted physical resource units are distributed in frequency partitions. Each frequency partition is divided into a localized and/or distributed block of resources. Sector-specific permutation can be used, and direct mapping of the resources is utilized for localized resources. The size of the distributed or localized resources is flexibly configured per sector. Adjacent sectors are not required to have the same configuration of localized and distributed resources.

An inner permutation is performed where the localized and distributed groups are mapped into logical resources by direct mapping of localized blocks and by subcarrier-level permutation of distributed resource units. The inner permutation defined for the DL distributed resource allocations within a frequency partition spreads the subcarriers of the distributed units over the entire distributed resource allocation. The granularity of the inner permutation is equal to the minimum unit for forming a distributed block (i.e., a subcarrier).

A logical resource unit is the basic logical unit for distributed and localized resource allocations. A logical resource unit comprises 18 x 6 subcarriers occupying a 196.88-kHz bandwidth in the frequency domain.



■ **Figure 6.** Downlink/uplink pilot structures for one, two, and four streams.

In the UL, the subframes are divided into a number of frequency partitions, where each partition consists of a set of physical resource units over the available number of OFDMA symbols in the subframe. Each frequency partition can include localized and/or distributed physical resource units. This is different from the legacy system where each zone can accommodate only localized or distributed subchannels. The UL resource petitioning and mapping is similar to that of the DL.

The UL distributed units comprise a group of subcarriers that are spread over a frequency partition. The size of the distributed unit is equal to the logical resource blocks. The minimum unit for constructing a distributed resource unit is a tile. The UL tile sizes are six subcarriers by six OFDMA symbols. The size of the localized resource unit equals the size of the logical resource units for localized allocations (i.e., 18 subcarriers by six OFDMA symbols). The tile permutation defined for the UL distributed resources spreads the tiles over the entire allocated frequency band.

MODULATION AND CODING

IEEE 802.16m supports quadrature-phase shift-keying (QPSK), 16-QAM, and 64-QAM modulation schemes in the DL and UL. The performance of adaptive modulation generally suffers from the power inefficiencies of multi-level-modulation formats. This is due to the variations in bit reliabilities caused by the bit-mapping onto the signal constellation. To overcome this issue, a constellation rearrangement scheme is utilized where a signal constellation of quadrature amplitude modulation

(QAM) signals between retransmissions is rearranged; that is, the mapping of the bits onto the complex-valued symbols between successive HARQ retransmissions is changed, resulting in averaging the bit reliabilities over several retransmissions and lower packet-error rates. The mapping of bits to the constellation point depends on the constellation rearrangement type used for HARQ retransmissions and also can depend on the MIMO scheme. The complex-valued modulated symbols are mapped to the input of the MIMO encoder. Incremental-redundancy HARQ is used in determining the starting position of the bit selection for HARQ retransmissions.

Both convolutional code and convolutional turbo code with variable code rate and repetition coding are supported.

The modulation and coding schemes used in a data transmission are selected from a set of 16 modulation coding schemes (MCSs). Note that rate matching is used to create more MCS granularity [10].

PILOT STRUCTURE

Transmission of pilot subcarriers in the DL is required to enable channel estimation, channel quality measurement (e.g., CQI), frequency offset estimation, and so on. IEEE 802.16m supports both common and dedicated pilot structures. The common pilots can be used by all mobile stations. Dedicated pilots can be used with both localized and distributed allocations. The dedicated pilots are associated with a specific fractional-frequency-reuse (FFR) group and can be used only by the mobile stations assigned to that group; therefore, they can be precoded or beamformed similarly to the data subcarriers. The pilot structure is defined for up to eight transmission streams, and there is a unified design for common and dedicated pilots. Pattern A of dedicated or common pilots is used for one or two spatial streams, and pattern B is utilized in the case of four spatial streams (Fig. 6).

For the type 2 and type 3 subframes, one of OFDMA symbols is deleted or repeated. To overcome the effects of pilot interference among the neighboring sectors or BSs, an interlaced pilot structure is utilized by cyclically shifting the base pilot pattern such that the pilots of neighboring cells do not overlap.

The UL pilots are dedicated to localized and distributed resource units and are precoded using the same precoding as the data subcarriers. The UL pilot structure is defined for up to four spatial streams with orthogonal patterns. When pilots are power-boostered, each data subcarrier has the same transmission power across all OFDMA symbols in a resource block. The 18×6 UL resource blocks use the same pilot patterns as the DL counterpart for up to four spatial streams. The pilot pattern for 6×6 tile structure is different, and it is shown in Fig. 6.

CONTROL CHANNELS

DL Control Channels — DL control channels carry essential information for system operation. Depending on the type of control signaling, information is transmitted in different time intervals. The system configuration parameters

are transmitted at the superframe intervals, whereas control signaling related to user data allocations is transmitted at the frame/subframe intervals.

The superframe header carries system-configuration information. The content of the superframe header is divided into two categories, the primary and secondary superframe header. The primary superframe header is transmitted every superframe, whereas the secondary superframe header is transmitted over one or more superframes. The primary and secondary superframe headers are located in the first subframe within a superframe and are time-division-multiplexed with the preamble. The superframe header occupies 5 MHz of bandwidth. The primary and secondary superframe headers are frequency-division-multiplexed with data within the same subframe and are transmitted using predetermined modulation and coding schemes. Multiple-antenna schemes are used for transmission of the primary and secondary superframe headers to improve coverage and reliability (i.e., two spatial streams and space-frequency block coding [10]).

The advanced MAP consists of both user-specific and non-user-specific control information. Non-user-specific control information includes information that is not dedicated to a specific user or a specific group of users. It contains information required to decode the user-specific control signaling. Non-user-specific control information that is not carried in the superframe header can be included in this category.

User-specific control information consists of information intended for one or more users. It includes scheduling assignment, power control, HARQ feedback, or ACK/NACK information. Resources can be allocated persistently to the mobile stations. The periodicity of the allocation is configurable. Group control information is used to allocate resources and/or configure resources to one or multiple mobile stations within a user group. Each group is associated with a set of resources. Voice over IP (VoIP) is an example of the class of services that can take advantage of group allocation. Within a subframe, control and data channels are frequency-division-multiplexed. Both control and data channels are transmitted on logical resource units that span over all OFDMA symbols within a subframe [10].

The control regions are located per subframe or every other subframe. This control structure is conceptually similar to legacy sub-MAPs containing one user [1]; however, the modulation and coding of the control blocks are not known to the MS; hence, blind detection must be utilized by the MS.

An advanced MAP allocation information element (MAP-IE) is defined as the basic control element and can be addressed to one user using a unicast identifier or to multiple users using a broadcast identifier. It contains information corresponding to resource allocation, HARQ, MIMO transmission mode, and so on. Each unicast control information element is coded separately. Note that this method is different from the legacy system control mechanism where the information elements of all users are

jointly coded. There is a trade-off between reliability and efficiency when one uses separate versus joint coding. Non-user-specific control information is encoded independently from the user-specific control information. The transmission format of non-user-specific control information is predetermined.

UL Control Channels — The UL control channels carry various types of control information to support air-interface procedures. The information carried in the UL control channels is classified as shown in Table 3.

ADVANCED PREAMBLES

IEEE 802.16m utilizes a new hierarchical structure for the DL synchronization where two sets of preambles at superframe and frame intervals are transmitted (Fig. 7). The first set of preamble sequences mark the beginning of the superframe and are common to a group of sectors or cells. They carry partial cell ID information (e.g., BS type, sector information, or grouping of cell ID) and can support limited signaling (e.g., system bandwidth, carrier information, etc.). The primary advanced preamble has a fixed bandwidth of 5 MHz and can be used to facilitate location-based services. A frequency reuse of one is applied to the primary advanced preamble in frequency domain.

The second set of advanced preamble sequences (secondary advanced preamble) is repeated every frame and spans the entire system bandwidth and carries the cell ID. A frequency reuse of three is used for this set of sequences to mitigate inter-cell interference. A minimum number of 512 distinct cell IDs are carried through the secondary advanced preambles.

MULTI-ANTENNA TECHNIQUES IN IEEE 802.16M

DL MIMO Structure — IEEE 802.16m supports several advanced multi-antenna techniques including single- and multi-user MIMO (spatial multiplexing and beamforming), as well as a number of transmit diversity schemes. In a single-user MIMO (SU-MIMO) scheme, only one user can be scheduled over one resource unit, whereas in multi-user MIMO (MU-MIMO), multiple users can be scheduled in one resource unit [11]. Vertical encoding (or single codeword) utilizes one encoder block (or layer), whereas horizontal encoding (or multi-codeword) uses multiple encoders (or multiple layers). Each of the various SU-MIMO or MU-MIMO open-loop or closed-loop schemes is defined as a MIMO mode.

The DL MIMO transmitter structure is shown in Fig. 8. The encoder block contains the channel encoder, interleaving, rate-matching, and modulating blocks per layer. A layer is defined as an encoding and modulation input path to the MIMO encoder. The resource mapping block maps the complex-valued modulation symbols to the corresponding time-frequency resources. The MIMO encoder block maps the layers onto the streams, which are further processed through the beamforming or the precoder block. The beam-

The control regions are located per subframe or every other subframe. This control structure is conceptually similar to legacy sub-MAPs containing one user; however, the modulation and coding of the control blocks are not known to the MS; hence, blind detection must be utilized by the MS.

Uplink control channel	Description
MIMO feedback	MIMO feedback provides wideband and/or narrowband spatial characteristics of the channel that are required for MIMO operation. The MIMO mode, precoder matrix index, rank adaptation information, channel covariance matrix elements, power loading factor, eigenvectors, and channel sounding are examples of MIMO feedback information.
HARQ feedback	HARQ feedback (ACK/NACK) is used to acknowledge downlink data transmissions. The uplink HARQ feedback channel starts at a predetermined offset with respect to the corresponding downlink transmission. The HARQ feedback channel is frequency-division-multiplexed with other control and data channels.
Bandwidth request	Bandwidth requests are used to indicate the amount of bandwidth required by a user and are transmitted through indicators or messages. Contention or non-contention based random access is used to transmit bandwidth request information. A five-step regular procedure or an optional three-step quick access procedure is utilized.
Channel quality indicators	Channel quality feedback provides information about channel conditions as seen by the user. This information is used by the base station for link adaptation, resource allocation, power control, etc. There are two types of uplink fast feedback control channels: primary and secondary fast feedback channels. The primary fast feedback channel provides wideband feedback information, including channel quality and MIMO feedback. The secondary fast feedback control channel carries narrowband CQI and MIMO feedback information.
Uplink sounding channel	The sounding channel is used by a user terminal to transmit sounding reference signals to enable the base station to measure uplink channel conditions. The sounding channel occupies either specific uplink sub-bands or the entire bandwidth over an OFDMA symbol.
Ranging channel	The ranging channel is used for uplink synchronization. The ranging channel can be further classified into ranging for nonsynchronized and synchronized mobile stations. A random access procedure, which can be contention- or non-contention-based, is used for ranging. Contention-based random access is used for initial ranging and handover. Non-contention based random access is used for periodic ranging and handover. The ranging channel for non-synchronized mobile stations is frequency-division multiplexed with other uplink control and data channels.
Power control	The base station controls the transmit power per subframe and per user in the downlink and uplink. The downlink advanced MAPs are power-controlled based on the terminal uplink channel quality feedback. The per-pilot-subcarrier and per-data-subcarrier power can jointly be adjusted for adaptive downlink power control. The uplink power control is supported to compensate the path loss, shadowing, fast fading, and implementation loss, as well as to mitigate inter-cell and intra-cell interference levels. The uplink power control includes open-loop and closed-loop power control mechanisms.

■ **Table 3.** *Uplink control channels.*

forming/precoding block maps the streams to antennas by generating the antenna-specific data symbols according to the selected MIMO mode. The OFDMA symbol construction block maps antenna-specific data to the OFDMA symbols. The feedback block contains feedback information such as CQI or channel state information (CSI) from the mobile station.

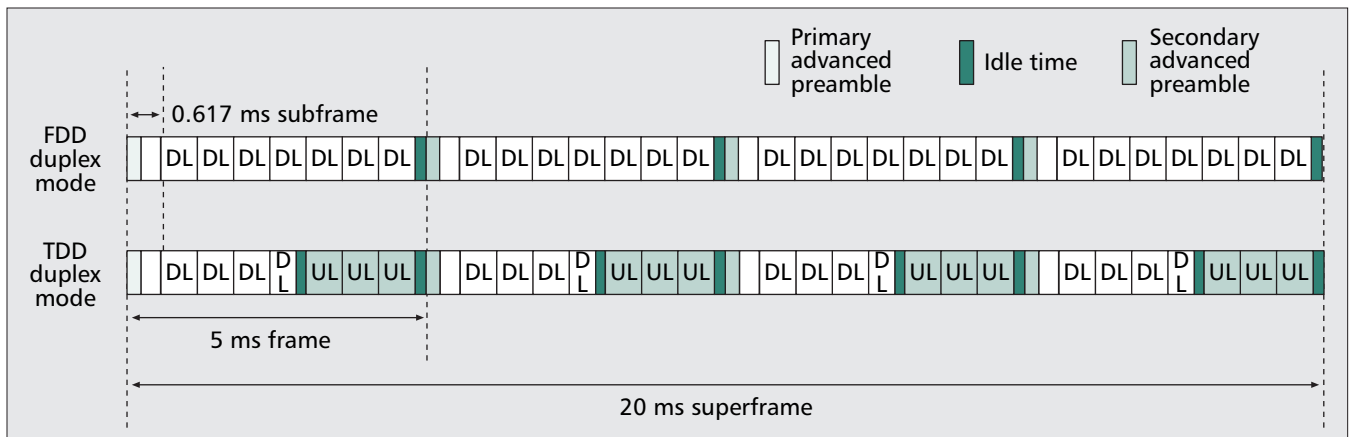
The minimum antenna configuration in the DL and UL is 2×2 and 1×2 , respectively. For open-loop spatial multiplexing and closed-loop SU-MIMO, the number of streams is constrained to the minimum of number of transmit or receive antennas. For open-loop transmit diversity modes, the number of streams depends on the space-time coding (STC) schemes that are used by the MIMO encoder. The MU-MIMO can support up to two streams with two transmit antennas and up to four streams for four and eight transmit antennas.

For SU-MIMO, vertical encoding is utilized, whereas for MU-MIMO horizontal encoding is employed at the BS, and only one stream is transmitted to each MS. The stream-to-antenna mapping depends on the MIMO scheme that is used. Note that in this case, CQI and rank feed-

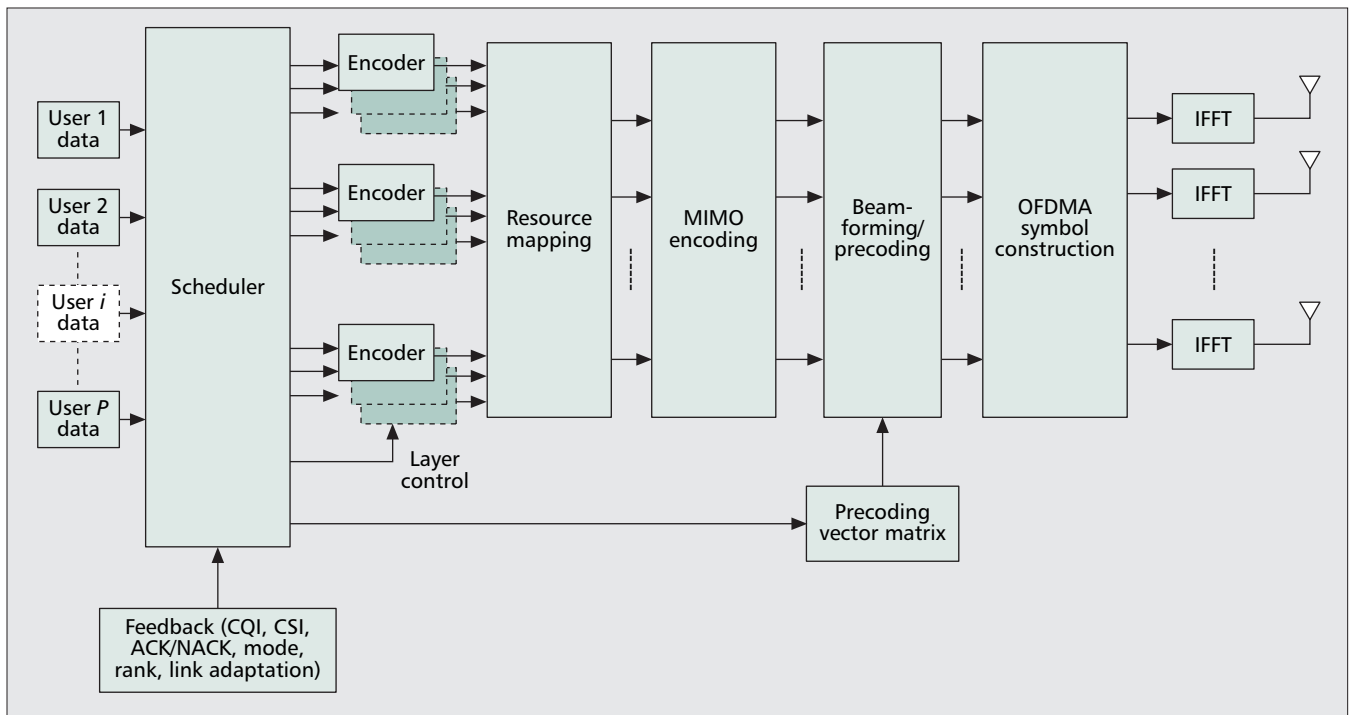
back are transmitted to assist the BS in rank adaptation, mode switching, and rate adaptation. For spatial multiplexing, the rank is defined as the number of streams to be used for each user. In FDD and TDD systems, unitary codebook-based precoding is used for closed-loop SU-MIMO. An MS can feed back some information to the BS in closed-loop SU-MIMO such as rank, sub-band selection, CQI, precoding matrix index (PMI), and long-term CSI.

The MU-MIMO transmission with one stream per user is supported. The MU-MIMO schemes include two transmit antennas for up to two users, and four and eight transmit antennas for up to four users. Both unitary and non-unitary MU-MIMO schemes are supported in IEEE 802.16m. If the columns of the precoding matrix are orthogonal to each other, it is defined as unitary MU-MIMO. Otherwise, it is defined as non-unitary MU-MIMO [11]. Note that beamforming is enabled with this precoding mechanism.

IEEE 802.16m has the capability to adapt between SU-MIMO and MU-MIMO in a predefined and flexible manner. Multi-BS MIMO techniques also are supported for improving sector and cell-edge throughput using multi-BS col-



■ Figure 7. Structure of advanced preambles.



■ Figure 8. Illustration of downlink MIMO structure [10].

laborative precoding, network coordinated beam-forming, or inter-cell interference cancellation. Both open-loop and closed-loop multi-BS MIMO techniques are under consideration.

UL MIMO — The block diagram of the UL MIMO transmitter is illustrated in Fig. 9. Note the similarities of MIMO baseband processing in the DL and UL.

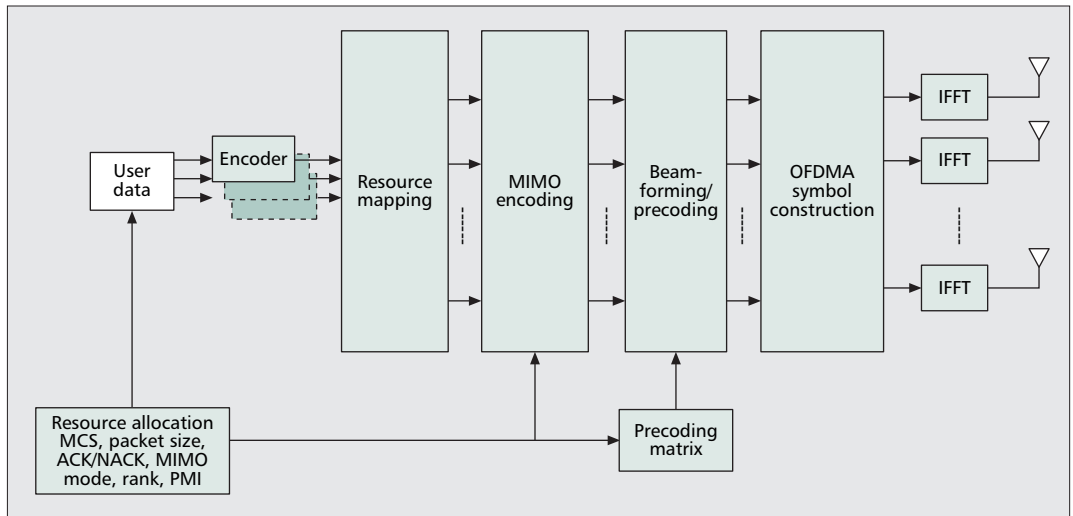
The BS schedules users to resource blocks and determines the MCS level and MIMO parameters (mode, rank, etc.). The supported antenna configurations include one, two, or four transmit antennas and more than two receive antennas. In the UL, the MS measurements of the channel are based on DL reference signals (e.g., common pilots or a mid-ambler).

A number of antenna configurations and transmission rates are supported in UL open-loop SU-MIMO including two and four transmit

antennas with rate one (i.e., transmit diversity mode) and two and four transmit antennas with rates two, three, and four (i.e., spatial multiplexing). The supported UL transmit diversity modes include two and four rate one transmit antenna schemes such as space-time-block coding (STBC) or space-frequency-block coding (SFBC) and rank one precoder.

The multiplexing modes supported for open-loop single-user MIMO include two and four transmit-antenna rate two schemes with and without precoding, four transmit-antenna rate three schemes with precoding, and four transmit-antenna rate four schemes. In FDD and TDD systems, unitary codebook-based precoding is supported. In this mode, an MS transmits a sounding reference signal in the UL to assist the UL scheduling and precoder selection in the BS. The BS signals the resource allocation, MCS, rank, preferred precoder index, and packet size

UL MU-MIMO enables multiple MSs to be spatially multiplexed on the same radio resources. Both open-loop and closed-loop MU-MIMO are supported. The MSs with single transmit antenna can operate in open-loop MU-MIMO mode. Unitary codebook-based precoding is supported for both TDD and FDD.



■ Figure 9. Illustration of the UL MIMO structure [10].

to the MS. The codebook on the UL is the same or a subset of the SU-MIMO codebook in the DL.

UL MU-MIMO enables multiple MSs to be spatially multiplexed on the same radio resources. Both open-loop and closed-loop MU-MIMO are supported. The MSs with single transmit antenna can operate in open-loop MU-MIMO mode. Unitary codebook-based precoding is supported for both TDD and FDD.

OVERVIEW OF THE IEEE 802.16M MAC LAYER

There are various MAC functionalities and features that are specified by the IEEE 802.16m standard, some of which are extensions of the existing features in mobile WiMAX [2]. The following sections briefly describe selected MAC features.

MAC ADDRESSING

The IEEE 802.16m standard defines permanent and temporary addresses for a mobile station that identify the user and its connections during operation. The MS is identified by a unique 48-bit identifier. The MS is further assigned the following temporary identifiers:

- A station identifier during network entry (or network re-entry) that uniquely identifies the MS within the cell
- A flow identifier that uniquely identifies the management and transport connections with the MS.

NETWORK ENTRY

Network entry is the procedure through which an MS detects a cellular network and establishes a connection with that network. The network entry includes the following steps (Fig. 4):

- Synchronization with the BS by acquiring the preambles
- Acquiring the required information such as BS and network service provider identifiers for initial network entry and cell selection
- Ranging

- Authentication and registration
- Service-flow set up

Neighbor search is based on the same DL signals as initial network search except that some information is provided through neighbor advertisement messages by the serving BS.

CONNECTION MANAGEMENT

Connections are identified by the combination of the station identifier and the flow identifier. Two types of connections (i.e., management and transport connections) are specified. Management connections are used to carry MAC management messages. Transport connections are used to carry user data including upper-layer signaling messages and data-plane signaling such as ARQ feedback. Fragmentation and augmentation of the MSDUs are supported on transport connections.

Management connection is bidirectional, and the predefined values of the flow identifier are reserved for unicast management connection(s). Management connections are established automatically after the station identifier is assigned to an MS during initial network entry. Transport connection, on the other hand, is unidirectional and is established with a unique flow identifier assigned during the service-flow establishment procedure. Each active service flow is uniquely mapped to a transport connection.

QoS

The IEEE 802.16m MAC assigns a unidirectional flow of packets with specific QoS requirements with a service flow. A service flow is mapped to a transport connection with a flow identifier. The QoS parameter set is negotiated between the BS and the MS during the service flow set up/change procedure. The QoS parameters can be used to schedule traffic and allocate radio resources. The UL traffic can be regulated based on the QoS parameters. IEEE 802.16m supports adaptation of service-flow QoS parameters. The MS and BS negotiate the possible QoS parameter sets during service-flow set up.

MAC MANAGEMENT MESSAGES

To satisfy the latency requirements for network entry, handover, and state transition, IEEE 802.16m supports the fast and reliable transmission of MAC management messages. The transmission of MAC management messages using HARQ is under consideration, where retransmissions can be triggered by an unsuccessful outcome from the HARQ entity in the transmitter. If the MAC management message is fragmented into multiple MSDUs, only unsuccessful fragments are retransmitted.

MAC HEADER

IEEE 802.16m specifies an efficient MAC header for small payload applications comprising fewer fields with a shorter size, reducing the header size to two bytes. The new MAC header consists of the *extended header indicator*, *flow identifier*, and *payload length* fields.

ARQ AND HARQ FUNCTIONS

An ARQ block is generated from one or multiple MSDUs or MSDU fragment(s). ARQ blocks can be variable in size and are sequentially numbered. If the HARQ entity in the transmitter determines that the HARQ process was terminated with an unsuccessful outcome, the HARQ entity in the transmitter informs the ARQ entity in the transmitter about the failure of the HARQ burst. Then, the ARQ entity in the transmitter can initiate retransmission and resegmentation of the appropriate ARQ blocks.

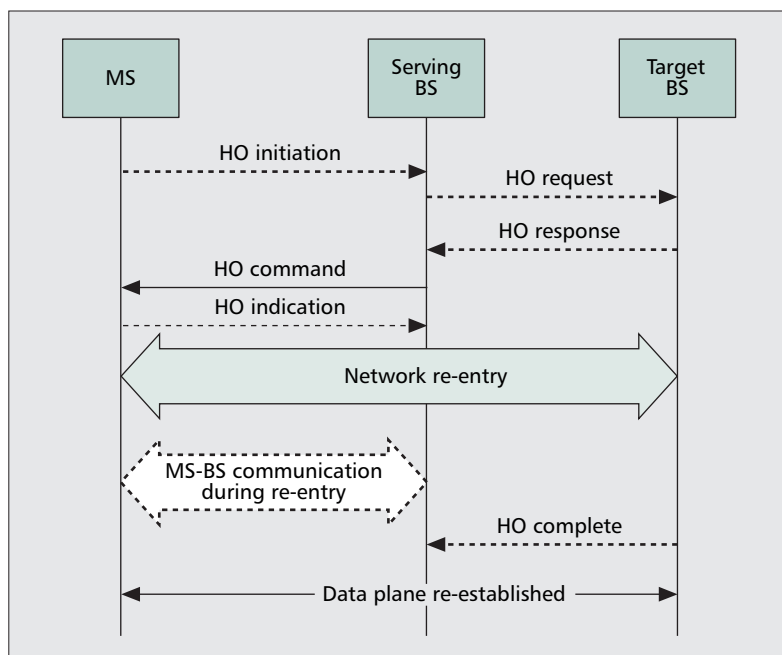
IEEE 802.16m uses adaptive asynchronous and non-adaptive synchronous HARQ schemes in the DL and UL, respectively. The HARQ operation is relying on an N -process (multi-channel) stop-and-wait protocol. In adaptive asynchronous HARQ, the resource allocation and transmission format for the HARQ retransmissions may be different from the initial transmission. A non-adaptive synchronous HARQ scheme is used in the UL, where the parameters and the resource allocation for the retransmission are known a priori.

MOBILITY MANAGEMENT AND HANDOVER

IEEE 802.16m supports both network-controlled and MS-assisted handover (HO). The MS executes the HO as directed by the BS or cancels the procedure through the HO cancellation message. The MS also can maintain communication with the serving BS while performing network re-entry at the target BS as directed by the serving BS. Figure 10 illustrates the general HO procedure.

The HO procedure is divided into three stages: HO initialization, HO preparation, and HO execution. Upon completion of the HO execution, the MS is ready to perform network re-entry with the target BS. In addition, the HO cancellation procedure is defined to allow an MS to cancel the HO procedure [10].

The HO preparation is completed when the serving BS informs the MS of its HO decision through the HO control command. The control signaling includes an action time for the MS to start network re-entry with the target BS and an indication whether the MS should maintain com-



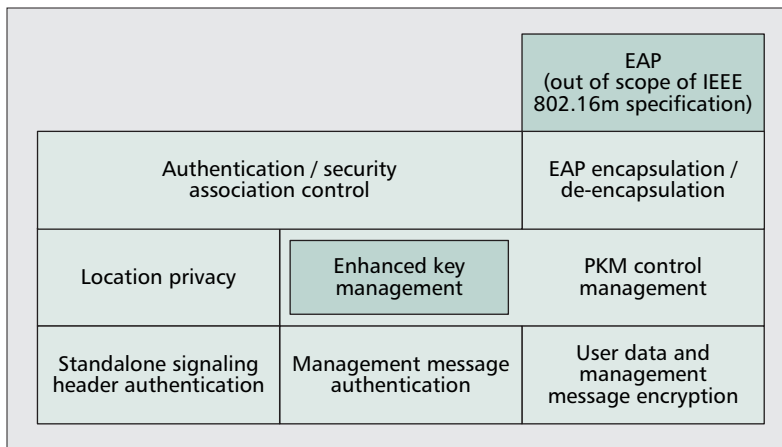
■ Figure 10. General handover procedure [10].

munication with the serving BS during network re-entry. If the communication cannot be maintained between the MS and the serving BS during network re-entry, the serving BS stops allocating resources to the MS for transmission in action time. If directed by the serving BS through the HO control command, the MS performs network re-entry with the target BS during action time while continuously communicating with the serving BS. The MS cannot exchange data with the target BS prior to completion of network re-entry.

POWER MANAGEMENT

Sleep mode is a state in which an MS performs pre-negotiated periods of absence from the serving BS. Using the sleep mode, the MS is provided with a series of alternative listening and sleep windows. The listening window is the time interval in which the MS is available for transmit/receive of control signaling and data. The IEEE 802.16m has the capability of dynamically adjusting the duration of sleep and listening windows within a sleep cycle based on changing traffic patterns and HARQ operations. When the MS is in active mode, sleep parameters are negotiated between the MS and the BS. The BS instructs the MS to enter the sleep mode. MAC management messages can be used for the sleep mode request/response [1]. The period of the sleep cycle is measured in units of frames or superframes and is the sum of the sleep and listening windows. During the MS listening window, the BS can transmit the traffic indication message intended for one or multiple MSs [10].

Idle mode enables the MS to become periodically available for DL broadcast-traffic messaging such as paging a message without registering with the network. The network assigns MSs in the idle mode to a paging group during idle mode entry or location update. The MS monitors the paging message during the listening



■ **Figure 11.** Functional blocks of the IEEE 802.16m security architecture.

interval. The start of the paging-listening interval is calculated based on the paging cycle and the paging offset [10]. The serving BS transmits the list of paging group identifiers at the predetermined location at the beginning of the paging-available interval. The IEEE 802.16m paging mechanism uses a two-step procedure that includes a paging indication followed by a full paging message. The paging indications, if present, are transmitted at the predetermined location. The paging message contains the identification of the MSs to be notified of pending traffic or a location update.

SECURITY

Security functions provide subscribers with privacy, authentication, and confidentiality across the IEEE 802.16m network. The MAC packet data units are encrypted over the connections between the MS and the BS. Figure 11 shows the functional blocks of IEEE 802.16m security architecture.

The security architecture is divided into security management and encryption and integrity logical entities. The security management functions include overall security management and control, Extensible Authentication Protocol (EAP) encapsulation/de-encapsulation, privacy key-management (PKM) control, security-association management, and identity/location privacy. The encryption and integrity protection entity functions include user data encryption and authentication, management message authentication, and message confidentiality protection [10].

CONCLUSIONS

The IEEE 802.16m is designed to provide state-of-the-art mobile broadband wireless access in the next decade and to satisfy the growing demand for advanced wireless multimedia applications and services. The standardization of IEEE 802.16m and Release 2.0 of the mobile

WiMAX profile are expected to be completed by 2011.

Multihop relay architecture, multi-carrier operation, self-configuration, advanced single-user/multi-user multi-antenna schemes and interference mitigation techniques, enhanced multicast-broadcast service, increased VoIP capacity, improved cell-edge user throughput, and support of vehicular speeds up to 500 km/h, and so on are among the most prominent features that would make IEEE 802.16m one of the most successful and advanced broadband wireless-access systems in the next decade.

ACKNOWLEDGMENT

The author would like to acknowledge and thank contributions from several of his Intel Corporation colleagues to the content of this article.

REFERENCES

- [1] Part 16: Air Interface for Broadband Wireless Access Systems, P802.16Rev2/D9, (Revision of IEEE Std 802.16-2004 and consolidates material from IEEE Std 802.16e-2005, IEEE Std 802.16-2004/Cor1-2005, IEEE Std 802.16f-2005, and IEEE Std 802.16g-2007), Jan. 2009.
- [2] WiMAX Forum, "Mobile System Profile, Release 1.0 Approved Specification," Revision 1.7.1, Nov. 7, 2008; <http://www.wimaxforum.org/technology/documents>
- [3] ITU-R M.2134, "Requirements Related to Technical System Performance for IMT-Advanced Radio Interface(s) [IMT.TECH]," draft new report, Nov. 2008.
- [4] ITU-R M.2135, "Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced [IMT.EVAL]," draft new report, Nov. 2008.
- [5] IEEE 802.16m-08/004r5, "IEEE 802.16m Evaluation Methodology Document," Jan. 2009; <http://ieee802.org/16/tgm/index.html>
- [6] IEEE 802.16m-07/002r8, "IEEE 802.16m System Requirements," Jan. 2009; <http://ieee802.org/16/tgm/index.html>
- [7] ITU-R M.1822 Rec., "Framework for Services Supported by IMT," Oct. 2007.
- [8] WiMAX Forum, "WiMAX System Evaluation Methodology," July 2008; <http://www.wimaxforum.org/technology/documents>
- [9] WiMAX Forum, "Network Architecture Stage 2-3," Rel. 1, v. 1.2, Jan. 2008; <http://www.wimaxforum.org/technology/documents>
- [10] IEEE 802.16m-08/003r7, "IEEE 802.16m System Description Document," Jan. 2009; <http://ieee802.org/16/tgm/index.html>
- [11] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

BIOGRAPHIES

SASSAN AHMADI [M'85, SM'07] (sasan.ahmadi@intel.com) received his B.S., M.S., and Ph.D. degrees, all in electrical engineering, in 1988, 1990, and 1997, respectively. He has been with Intel Corporation since 2005 where, as chief system architect — 4G wireless systems and technical lead of the IEEE 802.16m project, he has significantly contributed to the next generation of mobile WiMAX radio air interface design, and has led the IEEE 802.16m standard development in IEEE and ITU-R/IMT-advanced. His research interests include mobile WiMAX and 4G cellular systems, the design and development of physical and medium access control layers of MIMO-OFDMA-based cellular systems, and packet data/voice/multimedia transport protocols. From 1997 to 2005 he was a principal engineer with Nokia Inc., where he successfully contributed to and led Nokia wireless multimedia initiatives. He also was affiliated as a lecturer with the University of California, San Diego from 2000 to 2003.