Chapter 3

WIMAX/802.16 BROADBAND WIRELESS NETWORKS

MICHEL BARBEAU, PAUL BOONE and EVANGELOS KRANAKIS*
School of Computer Science, Carleton University,
1125 Colonel By Drive, Ottawa, ON, Canada, K1S 5B6
**kranakis@scs.carleton.ca*

WiMAX/802.16 is a kind of network providing IP-based broadband wireless access to infrastructure networks such as the Internet. The two main envisioned applications are Web access and voice over IP. Highlights of WiMAX/802.16 are secure communications and broadband access in remote areas. This chapter covers the background, physical layer, medium access control layer, mobility support, mesh mode and multihop relay operation of WiMAX/802.16. The chapter ends with a presentation of thoughts for practitioners and a discussion of directions for future research.

1. Introduction

Worldwide Interoperability for Microwave Access/IEEE 802.16 (WiMAX/802.16) is a technology for fixed or mobile and secure broadband wireless access (BWA). It is a wireless alternative to digital subscriber line (DSL). The main envisioned applications are wireless voice over IP, wireless Internet access and broadband wireless access in rural areas. Topics covered in this chapter include an overview of WiMAX/802.16, the physical (PHY) and medium access control (MAC) layers, mobility, mesh mode, multihop relay and security.

2. Background

WiMAX/802.16 was first published as the IEEE 802.16 standard that defines fixed Point-to-Multipoint (PMP) BWA in the 10–66 gigahertz (GHz) range. Amendment 802.16a adds control enhancements, mesh mode support, more frequencies in the 2–11 GHz range, non-line-of-sight communications and licensed or unlicensed operation [6]. Amendment 802.16c specifies implementation profiles [7]. This work has been combined into the 802.16-2004 standard [5]. Mobility support and improved security are introduced in 802.16e-2005 [8]. A management information base (MIB) is documented in 802.16f. Management plane procedures and services are covered in 802.16g. At this time, the latest full standard has been incorporated

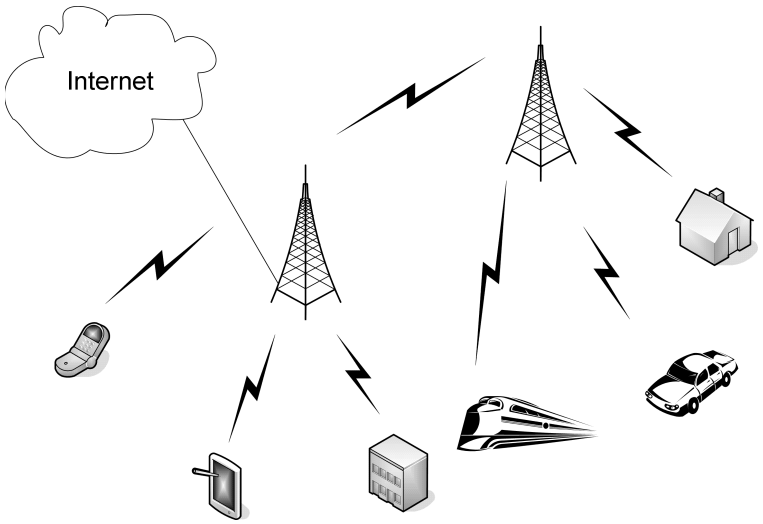


Fig. 1. A WiMAX/802.16 PMP network.

as 802.16-2008 [IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed broadband wireless access systems, IEEE Standard 802.16-2008 (2008)]. Other IEEE standards are being drafted. The WiMAX Forum is an organization that addresses implementation certification of WiMAX/802.16 products and integration with second and third generation cellular technologies.

The main mode of operation of WiMAX/802.16 is PMP, pictured in Fig. 1. Subscriber stations (SSs), e.g., laptops, handhelds, cars or residences, get access to the network through an association with a base station (BS), pictured as a tower. The association physically takes the form of a wireless link. Logically, the WiMAX/802.16 service, at the link layer, is connection oriented. There are two categories of connections: management connections and transport connections. The former is used for control purposes while the latter is used to carry data traffic.

In contrast to WiFi/802.11, WiMAX/802.16 is designed for long-range wireless access covering several kilometers. WiFi/802.11 is designed for ranges in the order of about a hundred meters. WiMAX/802.16 is predominantly deployed by cellular or Internet service providers. Typically, WiFi/802.11 networks are used to cover smaller areas ranging from small hotspots in shops to larger areas such as an organization's campus. WiMAX/802.16 radios primarily operate in the licensed spectrum, although there is support for unlicensed operation. WiFi/802.11 operates exclusively in unlicensed spectrum ranges. The data rates of WiMAX/802.16 and WiFi/802.11 are both in the Mbps range. However, the WiFi/802.11 bandwidth is shared by all devices attached to the access point while the data rate is exclusive for WiMAX/802.16 devices. This results from WiFi/802.11 network access being contention-based. Devices in the network use carrier sense multiple access with collision avoidance (CSMA/CA) to compete for the channel and transmit. WiMAX/802.16 devices must establish a connection with a BS, competing a

Table 1. WiMAX/802.16 versus WiFi/802.11.

	WiMAX/802.16	WiFi/802.11
Regulation	Licensed and unlicensed spectrum	Unlicensed spectrum
Range	Kilometers	Meters
Rate	Exclusive Mbps	Shared Mbps
Access	FDD, TDD, TDMA	CSMA/CA
Service	Connection, QoS	Contention-based, best effort

single time during network entry. Once connected with a BS, all transmissions are scheduled by the BS. This leads to a much better quality of service (QoS) support in WiMAX/802.16 networks. A summary of the comparison can be reviewed in Table 1.

WiMAX/802.16 and WiFi/802.11 should be thought of as complementary technologies that can co-exist. WiMAX/802.16 is better suited for metropolitan area networks (MANs). WiFi/802.11 is better suited for local area networks (LANs). One possible scenario is to have WiMAX/802.16 providing the backbone Internet for WiFi/802.11 networks.

The WiMAX/802.16 network architecture is structured into two main layers: the physical (PHY) layer and the Medium Access Control (MAC) layer. The PHY is responsible for the transmission of streams of bits as electromagnetic signals. It is also responsible for the organization of the bit streams into periodic units of fixed lengths called frames. Frames are further divided into smaller units called bursts. The physical layer is a two-way mapping between MAC protocol data units (PDUs) and physical layer frames received and transmitted through coding and modulation of an RF signal. The clients of the MAC layer push data units such as IP packets or ATM cells at identified ports of the WiMAX/802.16 service interface called service access points (SAPs).

The MAC layer is subdivided into three sublayers as shown in Fig. 2. The convergence sublayer adapts data units, accepted through the SAPs, to the requirements of the MAC service data units (SDUs), and *vice versa*. That might involve, for instance, tasks such as fragmentation and reassembly. There could be several instances of the convergence sublayer according to the kind of adaptation required.

The convergence sublayer also de-multiplexes the MAC SDUs to the connections to which they belong. MAC SDUs are pushed on the common part sublayer. This layer contains the core of the WiMAX/802.16 logic. For instance, it encompasses all aspects of connection management. It creates the protocol data units (PDUs) that are mapped to frames by the PHY layer. The security sublayer addresses authentication, access control and confidentiality issues of the MAC layer traffic.

3. Physical Layer

Four choices are available at the PHY layer: WirelessMAN SC, SCa, OFDM or OFDMA. WirelessMAN-SC uses a single radio carrier. It is intended for frequencies

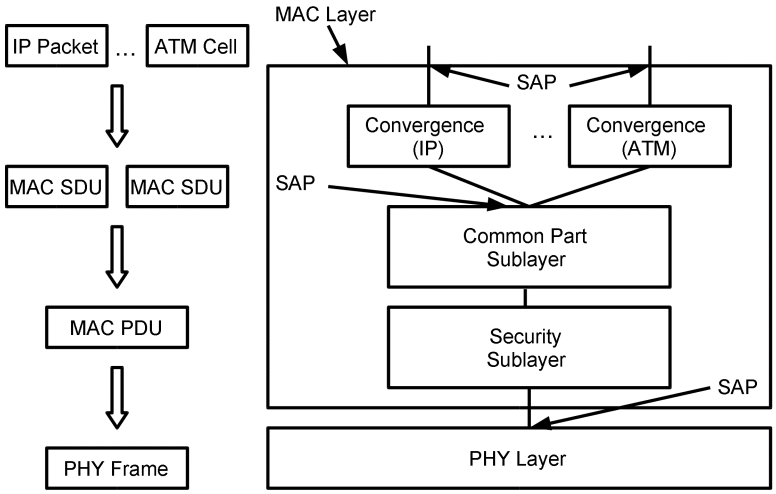


Fig. 2. Layered architecture of WiMAX/802.16.

in the range of 11–66 GHz. It requires line-of-sight (LOS). WirelessMAN-SCa uses a single-carrier at frequencies below 11 GHz and non-line-of-sight (NLOS). WirelessMAN orthogonal frequency division multiplexing (OFDM) and Orthogonal Frequency Division Multiple Access (OFDMA) are NLOS at frequencies below 11 GHz.

PMP communications consist of a BS and several SSs. The downlink, i.e., the BS to SSs channel, is determined by a frequency and a sector. The uplink, i.e., the SSs to BS channel, is determined by a frequency. If the downlink frequency and uplink frequencies are the same, then time division duplexing (TDD) is used to control the access to the channel. If the frequencies are different, then frequency division duplexing (FDD) is used. In addition, a time division multiple access (TDMA) technique is used on the uplink.

The flow of bits is structured as a sequence of frames of equal length. There is a downlink subframe and an uplink subframe. In FDD the downlink subframe and uplink subframe are simultaneous, but do not interfere because they are sent on different frequencies. In TDD, the downlink subframe and uplink subframe are consecutive and alternate. Various frame durations are supported between 2 and 20 ms in length. In TDD, the portion allocated to the downlink and portion allocated to the uplink may vary. A downlink subframe consists of a preamble, maps and data bursts (see Fig. 3). The preamble is a sequence of bits used for synchronization purposes. There are two maps. A downlink map announces the start position and transmission characteristics of the following data bursts. An uplink map announces the allocation of the bandwidth to the SSs for their transmission.

In the sequence of data bursts, each burst is transmitted according to a kind of modulation and a kind of forward error correction (FEC). They are sent in increasing

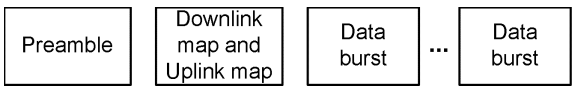


Fig. 3. A TDD downlink subframe.

Table 2. WiMAX data rates example.

	Data rate (Mbps)		
Channel bandwidth (MHz)	QPSK	16-QAM	64-QAM
20	32	64	96
25	40	80	120
30	44.8	89.6	134.4

degree of decoding difficulty, that is, in increasing level of data rate. Hence, an SS may only decode the bursts up to its reception capabilities and may ignore the bursts it cannot demodulate.

Table 2 gives the transmission characteristics of three modulation schemes. Quadrature Phase-Shift Keying (QPSK) uses four different symbols and can hence encodes two bits per symbol. 16-quadratute amplitude modulation (16-QAM) uses 16 different symbols and encodes four bits per symbol. It doubles the data rate of QPSK. 64-quadrature amplitude modulation (64-QAM) uses 64 different symbols and can encode six bits per symbol. It triples the data rate of QPSK. The degree of decoding difficulty increases with the data rate because the received signal needs to be stronger. In a downlink subframe, data bursts encoded with QPSK would be transmitted first, followed by the 16-QAM bursts and finally the 64-QAM data bursts.

An FDD downlink subframe contains a TDMA portion assigned to half duplex SSs. The TDMA portion consists of data bursts. Each data burst is assigned to a half duplex SS. A half duplex SS can transmit in a BS assigned TDMA data burst. Also, it needs to receive only when the BS transmits its data burst. A TDMA data burst on the downlink is prefixed by a preamble, transmitted by the BS, to allow the target SS to synchronize. The start position of every TDMA data burst is defined in the downlink map. QPSK is used to send the preambles.

The exact structure of an uplink subframe is determined by the BS. An uplink subframe is made of an arrangement of three kinds of bursts, which may be repeated and appear in any order. The three kinds of bursts are ranging, bandwidth requests and transmission. They are all for SS transmissions destined to the BS. A ranging burst is multiple contention-based access. When an SS initializes its attachment to the network, the SS sends range request messages to the BS with increasing power levels until reception is acknowledged by the BS, which is done by returning a range response message. The requests are sent within a ranging burst. A bandwidth request burst is multiple access contention based, but access must be in response to

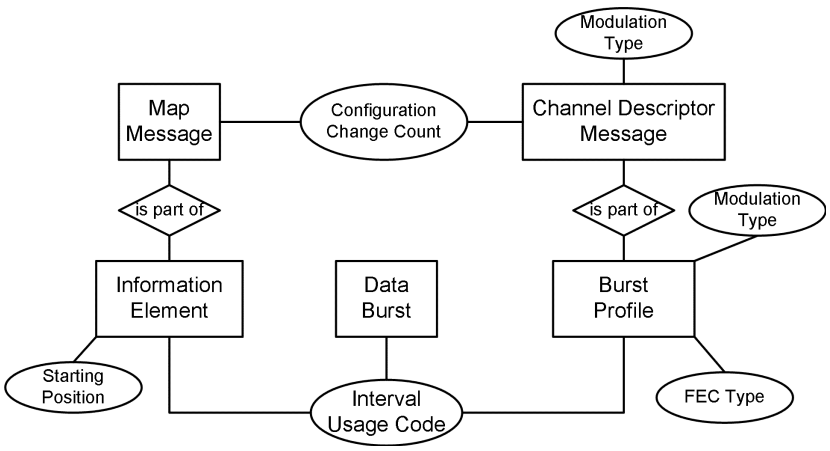


Fig. 4. Model of PHY and MAC concepts.

polling initiated by the BS. An SS may send a request for more bandwidth within such a burst. A transmission burst is a time interval granted to an SS by the BS.

The model of information of Fig. 4 represents the relationships existing among three key entities, that is, data burst, map message and channel descriptor message. Entities are represented by rectangles. Ovals represent their attributes. Lines depict relationships between entities. Data bursts are in downlink subframes and uplink subframes. A data burst is an interval of time during which data is transmitted. A data burst is identified by an interval usage code. A map message is either of type downlink or uplink and prefixes every downlink subframe. It is versioned with a number called a configuration change count. The version number matches the one of the corresponding channel descriptor message. A map message has a BS identifier. It contains a number of information elements. Each information element contains the starting position of a data burst, which is associated with the interval usage code.

A channel descriptor message is either of type downlink or uplink. It is sent periodically (maximum of 10 s between messages) in a MAC PDU. It is versioned with a number called a configuration change count. The version number matches the one of the corresponding map message. A channel descriptor message contains burst profiles. Each burst profile specifies the transmission characteristics of a data burst, which is associated to an information element with an interval usage code. The specification includes a modulation type and a FEC type.

OFDM is a transmission technique used to share the radio spectrum by several users. Another goal of OFDM is to construct communication channels using, for each of them, several segments distributed over a radio spectrum. This spreading mitigates, on the channel, the undesirable effects of fading or interference that may be stronger on some spectrum segments and weaker on others; it averages them in a sense.

The radio spectrum is divided into N adjacent subchannel elements. Each subchannel element is defined by a radio frequency that is used by a carrier. A channel consists of a number of subchannel elements. Each channel has its specific modulation parameters (rate, coding). The set of subchannel elements is divided into N_G groups. Each group contains N_E subchannel elements, i.e., N is equal to N_G times N_E. Each channel is constructed as the random selection of subchannel elements, each from different groups.

Example 1

Nine subchannel elements are used to constitute three channels. N_G is three and N_E is three. There are three groups: G_1, G_2 and G_3, each of them contains three subchannel elements. TDD is used for framing. Each frame is divided into five time slots. Each time slot on a channel can be assigned to an individual MAC layer connection. For instance, an individual connection may be assigned a sequence of time slots on a channel.

Example 2

A WiMAX/802.16 OFDMA scheme consists of 2,048 subchannel elements, i.e., N is equal to 2,048 [11]. Each subchannel occupies 187 KHz. TDD is used for framing, with TDMA on the uplink. Some subchannel elements are reserved for control purposes. For the downlink N_G is 48 and N_E is 32. For the uplink, N_G is 53 and N_E is 32.

4. Medium Access Control Layer

The PHY layer is responsible for sending bits of data from transmitters to receivers. It has no knowledge of the types of data that are transferred between parties. This is the responsibility of the MAC layer. The MAC layer controls access to the PHY layer from the higher layers in the protocol stack. The MAC layer has several other responsibilities including:

- building PDUs to be sent via the PHY layer,
- scheduling transmissions over the PHY layer,
- selecting PHY burst profiles and transmit power levels,
- providing QoS,
- retransmitting PDUs in case of error, and
- providing support for mobility.

4.1. *MAC Layer Concepts*

Figure 2 shows the three components of the MAC layer. The central element is the common part sublayer. It exchanges SDUs with the convergence sublayer. It also constructs PDUs, establishes connections and manages bandwidth. The common part is tightly integrated with the security sublayer. The security sublayer

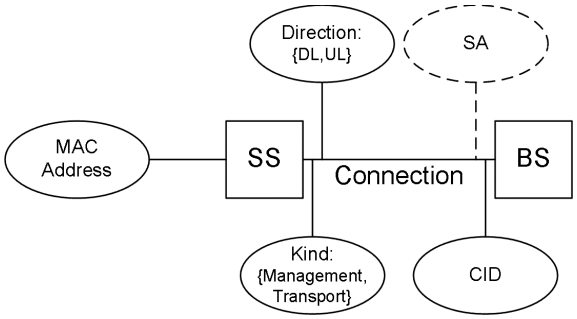


Fig. 5. Model of MAC concepts.

addresses authentication, authorization, key establishment and encryption. The security sublayer exchanges PDUs with the PHY layer.

A model of MAC concepts is shown in Fig. 5. The MAC layer is connection oriented. In order to offer services to SSs with various levels of QoS, all data communications are related to connections. An SS has a universal MAC address. A connection with a BS has a connection identifier (CID), direction [uplink (UL) or downlink (DL)], kind (management or transport) and an optional security association (SA) with the BS.

4.2. *Management Connections*

Table 3 summarizes the management connections. Following the ranging process, two pairs of management connections — one pair member for the uplink and one pair member for the downlink — are established between an SS and a BS. These are called the basic and primary uplink and downlink connections. An optional third pair of management connections, called the secondary management connections, may be created at the end of the network entry. The three pairs of connections handle the different QoS requirements of the management traffic between an SS and a BS. The basic connections are used by an SS and a BS to exchange short, time-urgent management messages, such as requests and responses for changes in burst profiles. The primary connections are used by an SS and a BS to exchange longer, delay-tolerant management messages.

Table 3. Management connections.

Type	Usage
DL basic UL basic	Short and urgent management messages
DL primary UL primary	Delay tolerant management messages
DL secondary UL secondary	IP encapsulated management messages

Finally, the secondary management connections are used by an SS and a BS to transfer delay tolerant messages encapsulated in IP packets, such as DHCP, SNMP or TFP messages.

4.3. *MAC Protocol Data Units*

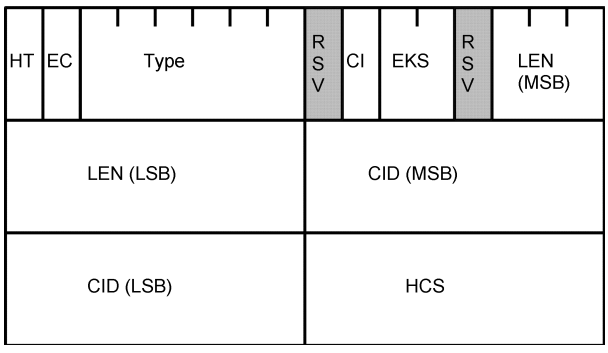
The common part sublayer is responsible for building PDUs with the SDUs received from the convergence sublayer. Depending on size, a SDU may be fragmented among multiple PDUs or several SDUs may be encapsulated inside a single PDU. There are two types of PDUs: the generic PDU and bandwidth request PDU.

The generic PDU carries user data and MAC layer control messages. The generic PDU contains a header, shown in Fig. 6, followed by a payload and a cyclic redundancy check (CRC) for error detection purposes. The bandwidth request PDU is used by SSs to request additional uplink bandwidth from a BS. It consists of a bandwidth request header, without payload and CRC.

The bandwidth request header is similar to the generic header shown in Fig. 6. The header is modified so that the 19 bits, starting at bit 5 and ending at the bit before the CID, holds the bandwidth request (BR) field. The BR field holds the number of bytes an SS is requesting from a BS. The header type (HT) field is set to one. The encryption control (EC) field is set to zero.

4.4. *Bandwidth Request and Allocation*

There are several messages sent by a BS in the downlink subframes. These include the downlink and uplink maps (DL-MAP and UL-MAP) and the downlink and uplink channel descriptors (DCD and UCD). The DL-MAP message defines the start time of every burst in the TDD downlink subframe and its TDMA portion. The UL-MAP message defines the structure of the uplink subframe by giving the offset of every burst. The DCD and UCD describe the PHY layer characteristics for the downlink and uplink.



Field	Description
HT	Header type (set to 0)
EC	Encryption control (0=not encrypted, 1=encrypted)
Type	Type
CI	CRC Indicator (0=not present, 1=present)
EKS	Encryption key sequence
LEN	Length (including header)
CID	Connection ID for sending PDU
HCS	Header check sequence

Fig. 6. The generic MAC header.

Requests are sent by the SSs to ask for new uplink allocation from the BS. A request may be in the form of a standalone bandwidth request header or it may be a piggyback request, an optional feature. All bandwidth requests are made in terms of the number of bytes required to contain a MAC header and payload. The request may be transmitted during any allocated uplink slot, but not during the initial ranging interval. Bandwidth requests may be aggregate or incremental. If the BS receives an aggregate bandwidth request, then the number of bytes requested will be the new bandwidth allocation. If the BS receives an incremental bandwidth request, then the number of bytes requested will be added to the existing allocation.

A requested bandwidth is for an SS's individual connection, but the bandwidth granted by the BS is addressed to the SS's basic CID, not to individual CIDs. Since an SS does not necessarily know which request is being granted, it must either assign the bandwidth to its outstanding requests or decide to do a backoff and send additional bandwidth requests if short of bandwidth.

Polling is a mechanism where SSs are allocated bandwidth by the BS in order for them to make bandwidth requests. The allocations are not sent in an explicit message, but are contained as information elements (IEs) inside the UL-MAP. Polling may be done using unicast, multicast or broadcast. The UL-MAP and DL-MAP indicate the uplink and downlink allocations for all the SSs currently communicating with a BS.

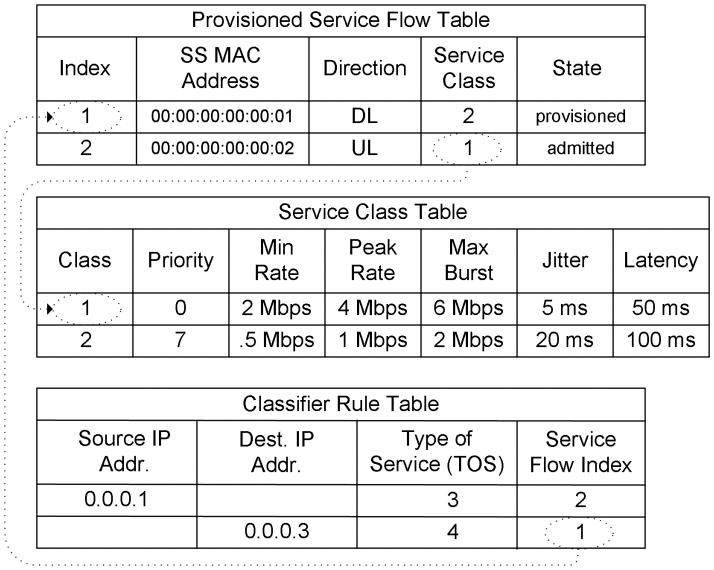
4.5. *Service Flows*

Service flows define the parameters of transport connections. They are an important part of the bandwidth allocation process. An SS requests uplink bandwidth on a per connection basis, identifying the service flow, and the BS grants bandwidth as an aggregate of all the per connection requests. Service flows are characterized by a set of QoS parameters. These include latency, jitter and guaranteed throughput parameters. A service flow has a service flow ID (SFID).

The MAC layer maintains information about provisioned service flows that are used to create connections. This information is stored in the service provisioning tables, depicted in Fig. 7. From the provisioned service flow table, using the service class index, the QoS parameters for a flow are looked up in the service class table. These parameters include priority, minimum and peak rates, maximum burst, jitter and latency. The classifier rule table provides the rules for mapping IP packets to service flows. Their IP addresses and type of service value determine the service flow, using the service flow index.

4.6. *Quality of Service*

One of the distinguishing features of the MAC layer is the built-in QoS support for five kinds of service flows. An Unsolicited Grant Service (UGS) flow transports realtime traffic consisting of periodic fixed size packets. This is intended for T1/E1 or VoIP traffic. Bandwidth grants are of fixed size. An SS is not required to explicitly



Provisioned Service Flow Table

Index	SS MAC Address	Direction	Service Class	State
1	00:00:00:00:00:01	DL	2	provisioned
2	00:00:00:00:00:02	UL	1	admitted

Service Class Table

Class	Priority	Min Rate	Peak Rate	Max Burst	Jitter	Latency
1	0	2 Mbps	4 Mbps	6 Mbps	5 ms	50 ms
2	7	.5 Mbps	1 Mbps	2 Mbps	20 ms	100 ms

Classifier Rule Table

Source IP Addr.	Dest. IP Addr.	Type of Service (TOS)	Service Flow Index
0.0.0.1		3	2
	0.0.0.3	4	1

Fig. 7. Service provisioning tables.

request bandwidth. A realtime polling service (rtPS) flow transports realtime traffic consisting of periodic variable size packets. This is suited for MPEG video traffic. To guarantee that the realtime needs are met, a BS provides realtime, periodic unicast request opportunities allowing an SS to specify the size of required bandwidth grants. An extended realtime polling service (ertPS) flow is an enhancement of both UGS and rtPS with unsolicited unicast bandwidth grants (as for UGS), while allowing variable size packets. This is intended for voice over IP services with silence suppression. A BS also allows an SS to request additional bandwidth as required. A non-realtime polling service (nrtPS) flow has unicast polls, on a regular basis, to ensure that the SS has bandwidth request opportunities even when the network is congested. It is intended for applications such as file transfers. A best-effort (BE) flow is designed for uplink traffic that has no QoS requirements.

4.7. *Network Entry*

An SS wanting to establish a connection with a BS follows the network entry procedure depicted in Fig. 8. The SS first scans to find a frequency in use by a BS. It listens to each possible frequency until it hears a frame preamble. The SS synchronizes with the BS by waiting for the DL-MAP and obtains the uplink channel characteristics from the UCD messages that are periodically broadcast by the BS. The SS then gets an Initial Ranging Interval from a DL-MAP message sent by the BS in order to perform initial ranging with the BS. Initial ranging determines the transmit power requirements and timing offsets of the SS in order to synchronize communication with the BS.

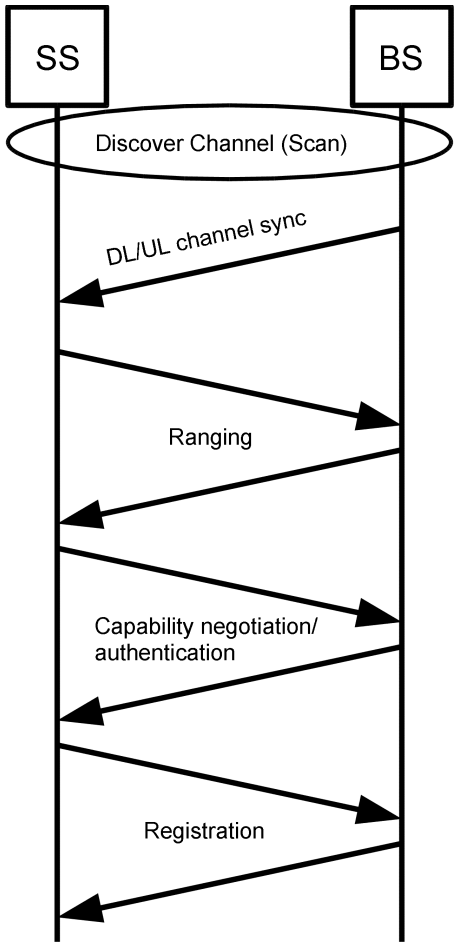


Fig. 8. Network entry procedure.

4.7.1. *Scanning*

The goal of scanning is to acquire a downlink signal from a BS. The exact number of frequencies depends on the regulatory provisioned bandwidth. It varies from one country to another. The number of frequencies is also determined by the choice of physical layer specification. Scanning continues periodically to aid SSs in the selection of suitable target BSs for a possible handover in order to maintain network connectivity in mobile WiMAX.

4.7.2. *Initial Ranging*

The initial ranging process, executed by an SS, determines the transmit power level and timing offset. The timing offset is used to synchronize each SS's transmission to a symbol marking the beginning of a minislot boundary for the SC and SCa PHY

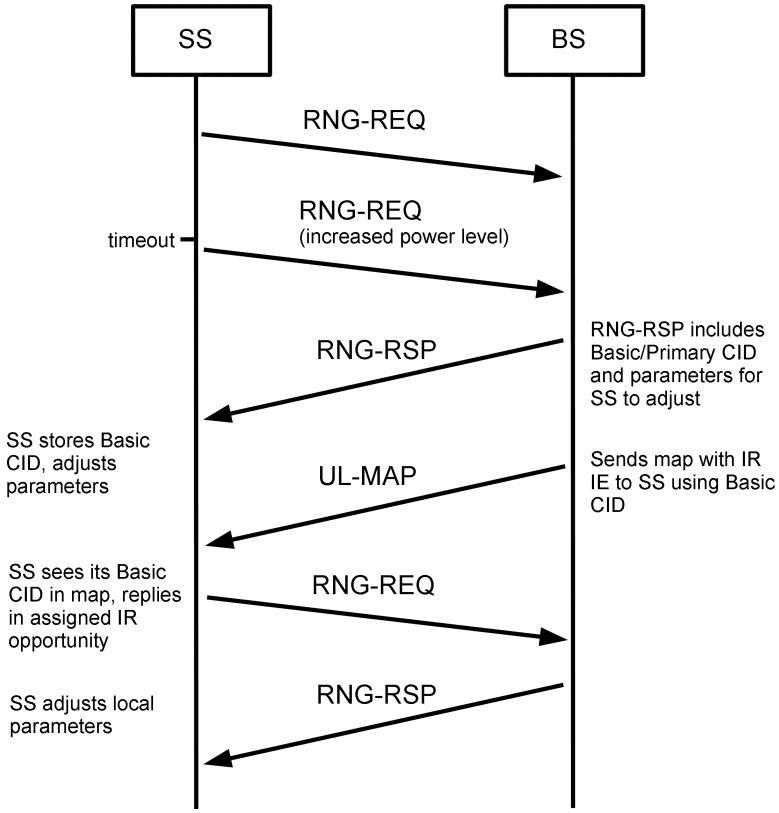


Fig. 9. Initial ranging process.

layers, or align its transmission with the BS received frame for the OFDM(A) PHY layers. Figure 9 shows the initial ranging procedure.

Once the SS has synchronized with the downlink channel and discovered the uplink channel parameters in the UCD message, it uses the UL-MAP message to determine the next initial ranging interval. Next, the SS sends a range request (RNG-REQ) message during the next interval with a CID set to zero. For OFDMA, the request is sent via code division multiple access (CDMA) using codes for initial ranging. The SS adjusts its transmission timings to account for delays, so that it appears to be collocated with the BS. The SS calculates the maximum signal strength P used for initial ranging as follows:

$$P = Max + (EIRP - RSS).$$

Max is the maximum equivalent isotropic received power for the BS receiver. *EIRP* is the equivalent isotropic radiated power of the BS. Both are obtained from the DCD. *RSS* is the received signal strength measured at the SS. If the values of *Max* and *EIRP* are unknown, then the SS starts sending the RNG-REQ message

at a minimum power level, retransmitting and increasing power, during future ranging opportunities after a backoff period, until a BS receives the message. Once a BS has received an RNG-REQ message, it must return a range response (RNG-RSP) message with the same CID. The RNG-RSP message includes the basic and primary management CIDs assigned to the SS, as well as information on power level adjustment, offset frequency adjustment and timing offset corrections. If the status of the RNG-RSP is success, then the ranging procedure completes. If not, then the SS uses scheduled ranging intervals to complete the ranging process.

4.7.3. *Negotiating Basic Capabilities*

Negotiating basic capabilities is a necessary step because WiMAX/802.16 is broad, with core functionality that must be supported by all vendors along with optional functionality that different vendors may or may not support. In order for equipment from different vendors to operate together, they must negotiate the set of capabilities that the BS and SS both support. An example is what PHY burst profiles (e.g., 64 QAM) are supported. The negotiation is a handshake beginning with the SS sending a basic capabilities request (SBC-REQ) message to the BS indicating its capabilities. The BS replies with a basic capabilities response (SBC-RSP) message that includes the capabilities common to both the SS and BS.

4.7.4. *Registration*

Registration is the process of allowing an SS to enter the network and to receive a secondary management CID. In order to register with the network, the SS sends a registration request (REG-REQ) message to the BS. This message includes information such as uplink CID support, IP version, SS capabilities, convergence sublayer support, ARQ settings and vender specific information. The BS replies with a registration response (REG-RSP) message that includes the secondary management CID. At this point, the SS can establish IP connectivity via DHCP over the secondary management connection.

4.7.5. *Establishing a Service Flow*

Service flows can be established either from the SS or BS side depending on whether the traffic arriving is on the uplink or downlink. Figure 10 shows the creation of a service flow initiated by a BS. Here the BS checks to determine if the SS is authorized for service and if the QoS requirements can be met with the available resources. It then creates a new SFID with the required class of service and sends this information in a dynamic service addition request (DSA-REQ) message to the SS. If the SS supports the service, then it will respond using a dynamic service addition response (DSA-RSP) message. The SS also reserves the required resources and enables downlink reception. When it receives the response, the BS enables uplink reception. Finally, the BS completes the procedure by sending an

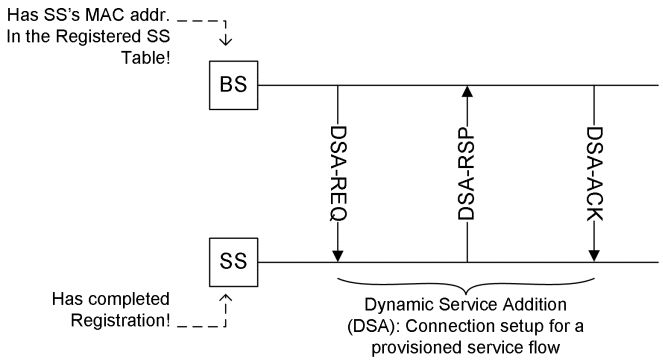


Fig. 10. Establishing a service flow: BS initiated.

acknowledgment (DSA-ACK) message. Both the BS and SS are now ready for reception and transmission can start in both directions.

When an SS wishes to establish a service flow it sends the BS a DSA-REQ message. Upon reception, the BS checks the integrity of the message and responds with a dynamic service received (DSX-RVD) message. The BS then determines if the SS is authorized for the service and if there are resources available to meet the QoS parameters. The BS then creates a new SFID with the required class of service and sends this in a DSA-RSP response message to the SS. The final step is for the SS to acknowledge the response using a DSA-ACK acknowledgment message. Both the BS and SS are now ready for reception and transmission can start in both directions.

5. Mobile WiMAX and Handovers

WiMAX/802.16 supports mobility. Mobility specific issues for the PHY layer, network support for mobility and handovers are discussed hereafter.

5.1. *Physical Layer*

For mobility support purposes, the OFDMA PHY layer has been adapted to use SOFDMA (scalable OFDMA). Here the number of subchannel elements can be chosen to optimize performance according to the channel and the radio conditions. The number of subchannel elements can be 128, 512, 1,024 or 2,048.

5.2. *Determining Network Topology*

To aid mobile stations (MSs) in discovering neighboring BSs, the BSs periodically broadcasts information about the network topology using neighbor advertisement (MOB_NBR-ADV) messages. BSs can obtain this information over the backbone network. These advertisements provide the MS with information that they would otherwise obtain by scanning.

BSs assign time for MSs to scan for neighboring BSs. The time that an MS has allocated for scanning is called the scanning interval. A BS may also schedule time to perform interleaving scanning for neighboring BSs. Interleaving scanning is the alternation of sending data and scanning. The MS requests a scanning interval by sending a MOB_SCN-REQ message to the BS including a requested amount of time. The BS responds with a MOB_SCN-RSP message either granting or denying the request. When the MS has been granted a scanning opportunity, it starts scanning for one or more neighboring BSs during the time allotted, beginning at the start frame indicated in the response message.

5.3. *Association Procedure*

Association is an optional initial ranging operation that can be done during the scanning interval. This is done between an MS and one of its neighboring BSs. There are three levels of association possible during the scanning interval.

Level 0 (scan/association without coordination): The MS performs ranging without assistance from the network. Ranging is contention based. It is successful if the MS receives the success RNG-RSP message from the neighboring BS.

Level 1 (association with coordination): The serving BS provides the MS with association parameters (ranging codes and transmission opportunity) for target BSs. This allows the MS to perform collision free ranging and succeeds when it receives the success RNG-RSP message from a neighboring BS.

Level 2 (network assisted association reporting): Similar to Level 1, but after the MS has sent its RNG-REQ message it does not need to wait for the RNG-RSP message from a BS. Instead, the RNG-RSP containing information about the PHY layer timing offsets of neighboring BSs is sent to the serving BS over the backbone network. The serving BS may send an aggregation of all this information to the MS in an association report (MOB_ASC-REP) message.

5.4. *Handovers*

As an MS moves throughout the coverage area, maintaining connectivity is done via performing handovers between neighboring BSs. An example of a handover, where an MS must choose one of six neighboring BSs, is shown in Fig. 11. In this case neighbor six is chosen. Selection of the best handover target can be complex since the MS must scan for neighboring BSs to find a suitable target based on a number of criteria such as signal strength or error rates. Since a handover is an important function, an MS should perform the scanning and determine a target BS before beginning the handover. Uplink and downlink communication can be temporarily suspended, between an MS and a BS, in order to allow the MS to perform scanning for neighboring BSs. While communication is suspended, the data streams must be buffered on either side.

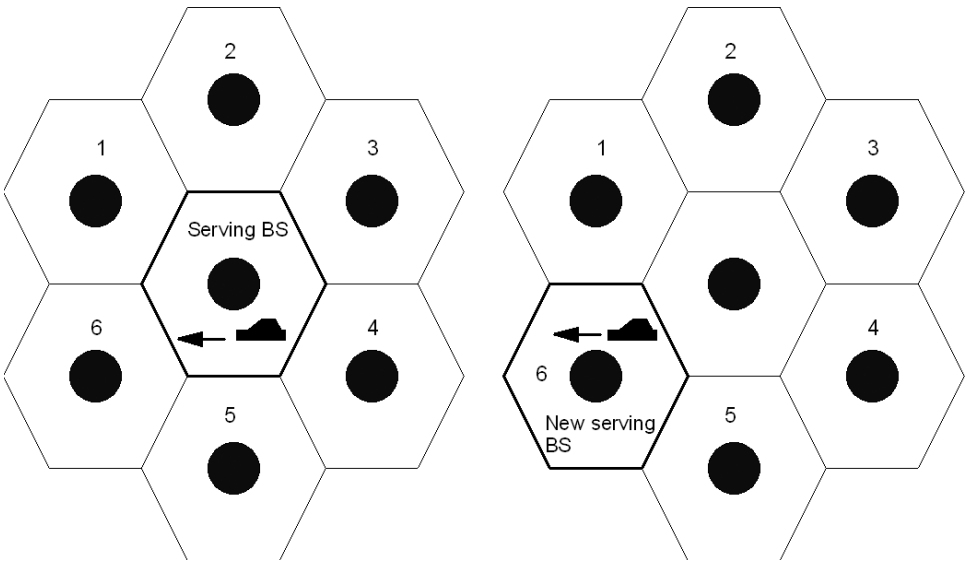


Fig. 11. A mobile station handover.

5.5. *Handover Process*

A handover is a process during which an MS changes its connections from one BS to another neighboring BS. From the perspective of an MS, the handover process is comprised of a number of steps including cell reselection, handover decision and initiation, synchronization with a target BS, ranging, negotiating capabilities, and authentication.

Network assisted handovers are supported in mobile WiMAX. Every BS obtains information about neighboring BSs over the backbone network. The BS periodically sends that information as a mobility neighbor advertisement (MOB_NBR-ADV) message to the MSs. For cell reselection, an MS may use information acquired from the MOB_NBR-ADV messages or may make a request to schedule scanning intervals to scan and perform ranging with a potential target neighboring BS. The handover decision and initiation may originate either at the MS or the serving BS. The notification of intent to handover is sent in either a mobility MS handover request (MOB_MSHO-REQ) message from the MS or a mobility BS handover request (MOB_BSHO-REQ) message from the BS. Scanning and obtaining the downlink and uplink transmission parameters of the neighboring BSs allows an MS to synchronize with the target neighbor BS. If the MS has received a MOB_NBR-ADV message including a potential target BSs identifier, frequency and channel descriptors, then this procedure can be shortened.

It is possible that a target neighboring BS has received a handover notification from the serving BS over the backbone allowing the target to assign the MS a scheduled ranging opportunity. The MS and target BS must conduct the ranging

process in order to determine transmission power level and timing offsets. If the MS includes the serving BS identifier in the RNG-REQ message, then the target BS may use the backbone network to request information about the MS from the serving BS. The target BS can then decide which of the remaining network entry steps may be skipped.

6. Mesh Mode and Multihop Relay

WiMAX/802.16 has an optional fixed mesh mode, and multihop relay (MR) is being developed as part of IEEE 802.16j [9]. Mesh mode is a mechanism that increases network coverage without growing the number of BSs. MR addresses the same issue, but is it more compatible with the PMP mode and supports MSs.

6.1. *Mesh Mode*

In the PMP mode, service is established only between SSs and BSs. In mesh mode, service can be also established between SSs. Basic functions such as network synchronization and transmission scheduling are based on the neighbor information that all nodes (BSs and SSs) are required to maintain. All nodes must keep a physical neighborhood list storing information such as MAC addresses, hop counts and node IDs. This information is used to determine how to synchronize and to schedule data transmissions.

6.1.1. *Network Entry Procedure*

The mesh network configuration (MSH-NCFG) and mesh network entry (MSH-NENT) messages are used to advertise the mesh network and aid new SSs attempting to synchronize and join. MSH-NENT is sent on the basic channel by an SS wishing to join the network. MSH-NCFG is an advertisement that is broadcast to all one-hop neighbors. Nodes currently participating in a mesh must periodically send a MSH-NCFG beacon including a network descriptor that contains information on the base channel in use as well as the BS identifier. A new SS entering a mesh (called a candidate node) must scan for MSH-NCFG messages in order to synchronize and begin the network entry procedure.

An SS remains synchronized with a network as long as it is receiving MSH-NCFG messages from neighbors. The SS collects the MSH-NCFG messages until it has received one from the same neighbor twice, as well as containing the ID of an operator matching one of its own.

From all of the MSH-NCFG messages heard, a candidate node chooses a potential sponsor node. The sponsoring node is used to negotiate basic capabilities and get authorized on the network. This process, shown in Fig. 12, begins with the candidate node sending an MSH-NENT:NetEntryRequest message to the sponsor node to open a sponsor channel. If the sponsor node accepts the request, them it opens a sponsor channel by responding with an MSH-NCFG:NetEntryOpen

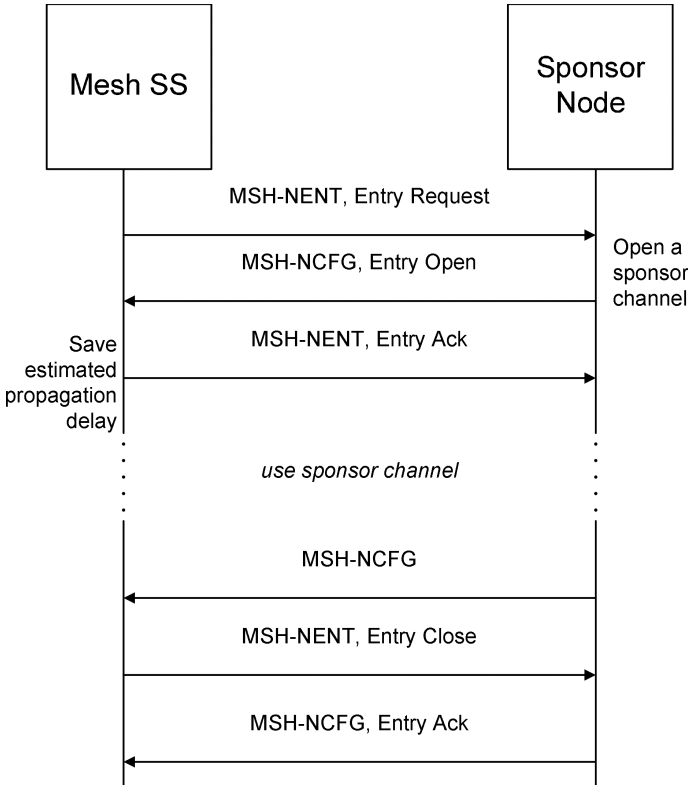


Fig. 12. Mesh mode network entry.

message including the candidate node's MAC address and scheduling information. The candidate node then acknowledges with an MSH-NENT:NetEntryAck message. The sponsor channel can now be used according to the scheduling information received from the network entry open message to negotiate basic capabilities, get authorized and register with the network, obtain an IP address, the time of day and download operational parameters. Once completed, the sponsor channel can be closed by the candidate node sending an MSH-NENT:EntryClose message to the sponsoring node and receiving an MSH-NCFG:EntryAck in response from the sponsor node. The SS is now a full member of the mesh.

Once an SS has entered the network, it is allowed to form new links with its other neighbors by using the MSH-NCFG message with the neighbor link establishment option. Figure 13 shows this in operation. When Node A wishes to establish a new link with one of its neighbors it listens for a MSH-NCFG message and obtains the frame number. Node A then sends a challenge in the form of *HMAC{Operator shared secret, frame number, NodeIDA, NodeIDB}* along with a frame number to Node B. HMAC is the hashed message authentication calculated with the secure hash algorithm SHA-1.

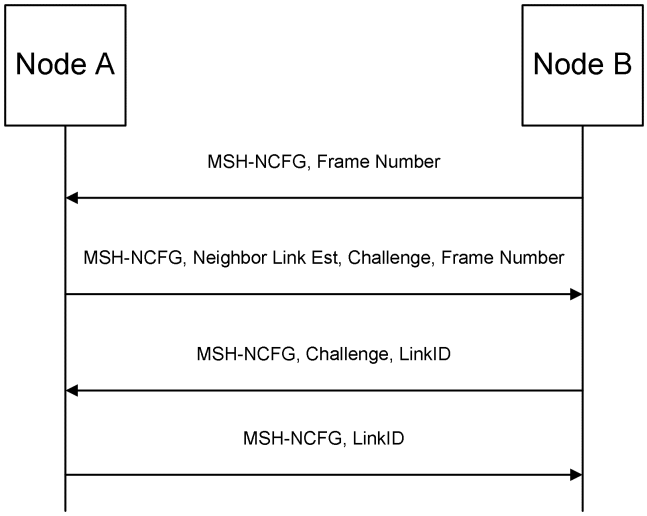


Fig. 13. Establishing new links in mesh mode.

Node B performs the same computation and compares the results to the value sent by Node A. If the values match, then Node B accepts the connection implicitly by sending a new challenge similar to the one sent by Node A. This time the frame number used in the calculation is the one sent by Node A. Node B also sends a random unused link ID with the challenge response. Node A checks validity of the challenge response and sends an MSH-NCFG:Accept along with a random unused link ID to complete the creation of a new link between A and B.

6.1.2. *Scheduling Transmissions*

Transmissions in mesh mode are TDMA based and, as such, they are coordinated and scheduled. Transmission scheduling can be done in a centralized fashion (coordinated by the mesh BS), in distributed fashion, or a combination of both. Transmissions are assigned to one of three types of logical channels: basic, broadcast or data. The basic channels are used for SS ranging and network entry messages. The broadcast channel is used for control information. The data channels are used for mostly data traffic.

6.1.2.1. Centralized scheduling

In centralized scheduling, the transmission schedule for all mesh SSs is determined by the mesh BS. The mesh BS determines the assignments of flows based on the resource requests of the mesh SSs and distributes this information to nodes in the mesh via a routing tree topology. The SSs can determine the schedule based on the assignment of flows by using a predetermined algorithm. The mesh BS acts similarly to normal PMP BSs with the exception that they may not be

in direct communication with the SSs. The resource requests and assignments are sent in the control portion of the frames. Two control messages are used in establishing centralized scheduling: mesh centralized scheduling (MSH-CSCH) and mesh centralized schedule configuration (MSH-CSCF). MSH-CSCF is used to distribute a global tree topology to mesh nodes. As nodes receive MSH-CSCF messages, the multicast routing tree can be constructed (with the BS as the root), and transmission opportunities calculated. The MSH-CSCH message is created by the mesh BS and is broadcast to all its neighbors. All these neighbors with a hop count lower than a hop count threshold forward the message to neighbors further away from the mesh BS. Nodes in the mesh use MSH-CSCH messages to request bandwidth from the mesh BS by setting a grant/request flag bit and sending this message to their parents in the routing tree. All nodes report the bandwidth request of neighbors of their children in their routing subtree up to the mesh BS.

When a node receives a new schedule, it computes the time it must transmit the schedule further down the routing tree, the frame at which the last node in the tree receives the message and the mesh BS transmission time of the message. This is completed by performing the following actions:

(i) the mesh BS transmits first;
(ii) eligible children of mesh BS transmit — in the order they appear in the routing tree;
(iii) eligible children of the nodes from Step (ii) transmit — in the order they appear in the routing tree;
(iv) the process repeats until all eligible nodes in the routing tree have transmitted.

All nodes are required to compute timings of uplink requests. Uplink requests may begin in the last frame that a node has received the previous schedule. All nodes except the mesh BS are eligible to transmit and must do so starting from the bottom of the routing tree.

6.1.2.2. Distributed scheduling

Distributed scheduling may be coordinated or uncoordinated. In coordinated distributed scheduling, both the BS and SSs must coordinate their transmissions within their two-hop neighborhood to avoid interference. Every node in coordinated mode uses a portion or all of the control part of each frame to regularly send its schedule, and any schedule changes it proposes, to all its neighbors. On a given channel all neighbor stations receive the same schedule transmissions. All stations in the mesh must use this same channel. This ensures that transmissions can be scheduled in a manner independent of any BS.

Both coordinated and uncoordinated modes must use a three-way handshake:

(i) an MSH-DSCH:Request is made with an MSH-DSCH:Availabilities message indicating potential time slots for a reply as well as the current schedule;

(ii) an MSH-DSCH:Grant message is sent in reply and includes a subset of the availabilities, if any, that match the request. Any neighbor not involved in the schedule assumes that it has been approved;
(iii) an MSH-DSCH:Grant message is sent by the original node and includes a copy of the grant message from Step (ii) as a confirmation.

The main differences between coordinated and uncoordinated distributed scheduling are that in the coordinated case the MSH-DSCH messages are scheduled in the control frame while in the uncoordinated case the MSH-DSCH messages must contend for the channel. In the uncoordinated mode, any node responding to a request must wait a certain amount of time before sending a grant so that any node listed earlier in the request is able to respond. The grant confirmation is sent in the first available slot after receiving the grant message.

6.2. *Multihop Relay*

Mesh mode has a few disadvantages that need to be overcome, including a lack of compatibility with the PMP mode due to a different PHY frame structure, a different MAC network entry procedure, and no mobility support. Multihop relay (MR) addresses these disadvantages. The goal of MR is to enhance coverage, throughput and system capacity by the introduction of relay stations and enhanced protocols to support multihop relay. The specifications for MR capabilities as well as the interoperability between relay stations and BSs are currently being defined.

6.2.1. *MR Base Stations and Relay Stations*

An MR base station (MR-BS) is a BS that supports multihop relay. Unlike mesh mode, where SSs relay messages for each other in the mesh, MR introduces relay stations (RSs) to provide a path to the MR-BS. All RSs are managed by an MR-BS but have some control over relaying traffic in their neighborhood. An RS provides the following functionality:

- relays user data and/or control information between other stations, and
- maintains processes that indirectly support MR.

There are three levels of RS mobility: fixed, nomadic and mobile. Fixed RSs are deployed in a fixed position and are never moved (tower, rooftop, side of building). Nomadic RSs are portable — they can be moved about — but only operate from a stationary position. Mobile RSs operate when they are in motion. If the route is fixed, such as an RS located on a train, additional support RSs can be planned along the route. However, RSs may also move about without a predetermined path.

6.2.2. *Usage Models*

There are a number of usage models being defined to handle various multihop relay modes of operation. Topologies in MR can support routes of two or more hops and

redundant routes may be maintained in order to improve network throughput and balance the network traffic.

6.2.2.1. Fixed infrastructure usage model

In this scenario, depicted in Fig. 14, a service provider deploys RSs and MR-BSs in order to enhance coverage, capacity, or throughput in areas that are not adequately serviced by a BS or to extend coverage outside the range of a BS. Fixed RSs owned by the service provider are used in this model. The provider is able to plan the deployment of RSs to provide line-of-sight conditions to other RSs or the MR-BS wherever possible. Examples of this usage model include cell edge coverage, covering holes (from shadowing) or coverage outside a cell's area.

6.2.2.2. In-building coverage usage model

In this scenario, one or more RSs are deployed to enhance coverage within a building, tunnel or underground facility. The RSs can be either provider or customer owned and typically only need to have basic functionality. The links are expected to be non-line-of-sight and the RS may have more than one antenna on the outside of a building to connect to another RS or the MR-BS. RSs may be deployed inside or outside buildings, and RSs owned by the customer may enter and leave the network at any time.

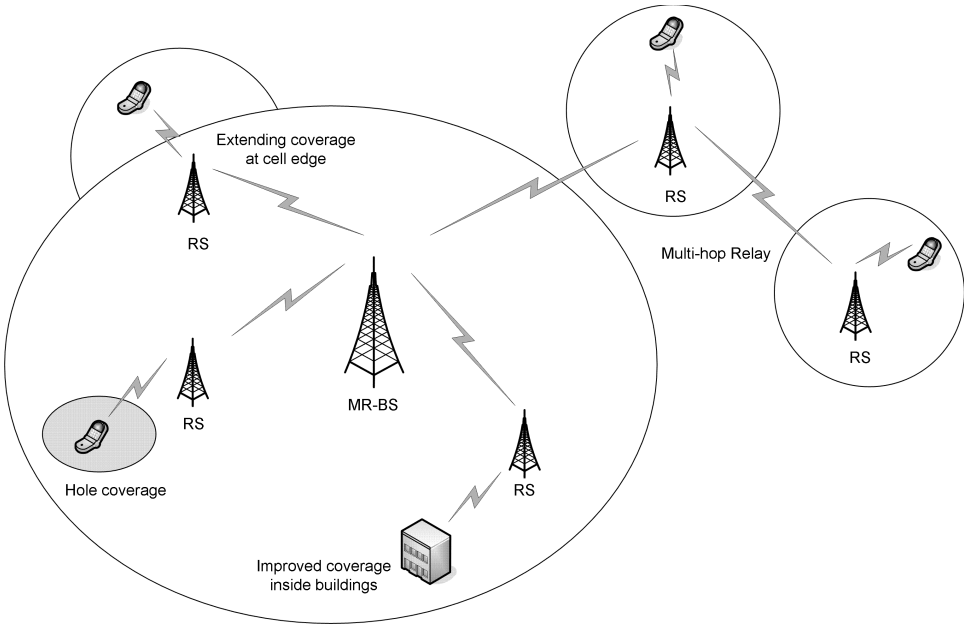


Fig. 14. MR fixed infrastructure usage model.

6.2.2.3. Temporary coverage usage model

In this scenario, nomadic RSs are deployed to provide temporary coverage to desired areas that may not have sufficient coverage by existing MR-BSs or other RSs. RSs enter the network upon deployment, and leave the network when the service is no longer required. Both line-of-sight and non-line-of-sight conditions can be expected for the RS communications with the MR-BS. Examples of this scenario include emergency, disaster recovery or special event coverage.

6.2.2.4. Mobile vehicle coverage usage model

This scenario covers situations where multiple MSs are traveling together on a vehicle such as a bus or a train. A mobile RS is located on the vehicle to connect with a MR-BS or RS. The mobile RS provides a fixed access link to MSs on the vehicle over a mobile link to the MR-BS or another RS. The mobile RS may enter and leave the network as it travels, moving in or out of the area of coverage. The mobile RS may also enter or leave the network on a fixed schedule such as the first/last run of the vehicle for the day.

7. WiMAX Security Model

Wireless LANs have attained significant commercial success because of their ease of use and deployment as well as the wider possibilities offered to users. WiFi/802.11 and WiMAX/802.16 have had a great impact on the market. They have changed the user's perception of ease and convenience in networking. Despite the obvious advantages offered by mobility to end users, drawbacks in user security and access have become evident. In this section, we focus on the security characteristics of WiMAX/802.16. We outline various security features including encryption, authentication, security associations, and key establishment and exchange. We begin with a brief introduction to WiFi/802.11.

7.1. *IEEE 802.11*

IEEE 802.11 is currently the most widespread wireless LAN standard. Recent versions allow for data rates of up to 54 Mbps. One of its major security disadvantages has been the Wired Equivalent Privacy (WEP) protocol [supporting only one encryption mechanism, the stream cipher RC4, using a pre-shared secret key and a per-packet initialization vector (IV) as input]. It was broken by the end of 2001. In addition, a cyclic redundancy check (CRC-32) is appended to a packet before it is encrypted in order to provide integrity protection. This too was shown to be ineffective by 2001. In WEP, the authentication server (AS), encryption and integrity protection endpoint (EIPE) and network access point coincide. Since the pre-shared secret key is used directly for authentication and encryption, neither key transfer from AS to EIPE nor security associations and

negotiations are necessary. In fact no key agreement or key establishment is being used. Meyer gives a comprehensive introduction [15].

Many different (and sometimes incompatible) proprietary solutions were since developed until finally the new IEEE 802.11i standard was adopted in 2004. It supports additionally the Advanced Encryption Standard (AES).

7.2. *IEEE 802.16*

IEEE 802.16 was developed to address wireless access for the last mile and its working group sought from the beginning to avoid the design mistakes of IEEE 802.11 by incorporating into it a pre-existing standard. Nevertheless, Johnston and Walker in their recent article criticize the adoption of Data Over Cable Service Interface Specifications (DOCSIS) [10]. Although designed to solve the last mile problem for cable, it has an otherwise different threat model from IEEE 802.16 since the former is wired while the latter is wireless technology [13].

IEEE 802.16 defines a separate security sublayer within the MAC layer that is in charge of authentication, secure key exchange and encryption. It encrypts connections between an SS and a BS. It also provides strong protection to operators against theft of service. It uses a key management protocol whereby the BS distributes keys to the SS.

7.2.1. *Traffic Encryption Key*

Traffic encryption keys (TEKs) are used to encrypt data. Encryption of a TEK is made possible using any of the following three algorithms: (a) two-key 3-DES in EDE mode, (b) RSA and (c) 128-bit AES in ECB mode. The value fields of a TEK, in a reply message sent to a client SS by a BS, are encrypted using one of the previous three algorithms. After authorization, BSs provide separate TEK for each Security Association ID (SAID) in the authorization reply message. The TEK state machine sends key request messages to the BS periodically, requesting refreshed keying material for their respective SAIDs. The TEK is encrypted using the key exchange key derived from an authorization key. The key reply contains the TEK, a CBC initialization vector (IV), as well as the remaining lifetime of each of the two sets of keying. The BS is also generating authorization keys (AKs), TEKs and IVs using a pseudo-random number generator. AKs in authorization reply messages shall be RSA public-key encrypted using the SS's public key.

7.2.2. *Encryption*

The encryption protocol encrypts only the packet MAC PDU payload while leaving the MAC header unencrypted. Public key cryptography is used to derive a shared AK, but this key is used thereafter only as a means to derive the subsequent keying material. IEEE 802.16-2004 supports CBC-Mode 56-bit DES and CCM mode AES. The former was first supported in the IEEE 802.16-2001 to encrypt the MAC PDU

payload, while the latter was first supported in the IEEE 802.16-2004, again for the encryption of the MAC PDU payload.

For DES CBC-Mode, different IVs are used for the downlink and uplink. In the downlink, the CBC initial value is calculated using the XOR of the IV included in the TEK key and content of the PHY synchronization field of the latest DL-MAP. In the uplink, the CBC initial value is calculated with XOR of the initial value in the TEK key and content of the PHY synchronization field of the DL-MAP that is in effect when the UL-MAP for the uplink transmission is created/received.

For AES in CCM mode, the MAC PDU payload is preprocessed by prepending an unencrypted four-byte packet number (PN). An eight-byte integrity check value (ICV) is appended to the payload. The PDU plaintext and ICV are encrypted and authenticated using the active TEK key.

7.2.3. *Authentication*

The basic model used is based on the IEEE 802.1X security model for authentication and access control which has been implemented in IEEE 802.11i and IEEE 802.16e networks. As depicted in Fig. 15, three entities are defined in the network: (i) the client (requesting access to the network), (ii) an access point and (iii) an authentication server (AS), i.e., a database containing the credentials needed to accept or deny access to clients. 802.1X enables port-based authentication: all stations have associated ports and traffic is blocked in the station port until an AS authenticates the client. The authentication process is summarized in Fig. 16.

The Privacy Key Management (PKM) protocol provides for secure distribution of keys between a BS and an SS whereby the former acts as a server and the latter

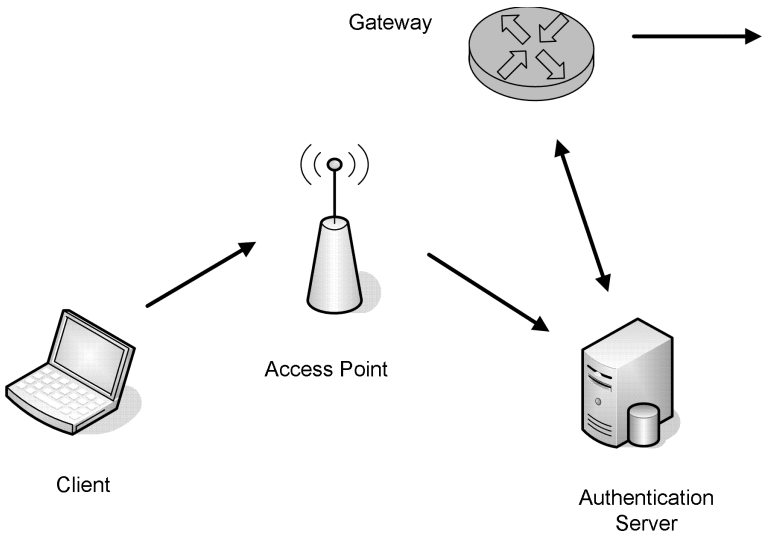


Fig. 15. Basic IEEE 802.1X scheme.

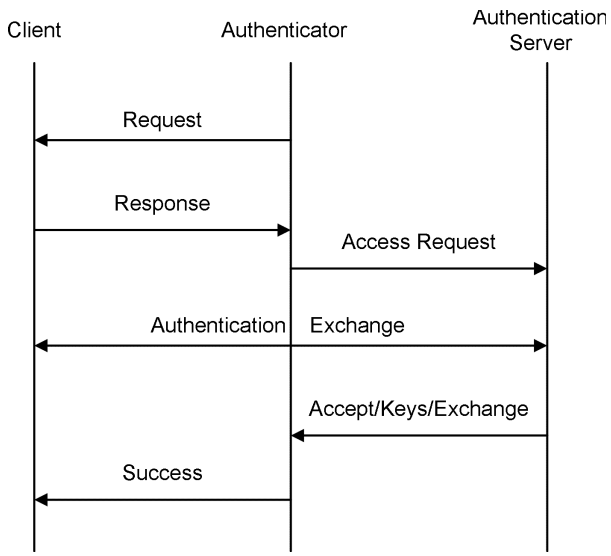


Fig. 16. Protocol authentication in the IEEE 802.1X scheme.

as a client. Information exchanged also includes how to access the several network services.

Periodic reauthorization and key refresh is also supported when the SS requests material from the BS. WiMAX/802.16 uses Public-Key Infrastructure (PKI) authentication, in which trusted authorities identify parties to a transmission via digital certificates. In addition, the PKM uses X.509 digital certificates, RSA public-key encryption algorithms and strong encryption algorithms to perform key exchanges. The digital certificate contains both the SS's public key and MAC address. Upon receipt of an authorization request from an SS, the BS verifies the digital certificate and, if valid, generates an AK which it encrypts with the SS's public key before sending it back to the SS.

7.2.4. *Security Associations*

A security association (SA) concerns the establishment of shared security information between an SS and a BS in order to support secure communications. SAs maintain the security state and keying material used to protect a connection [16]. When the two entities communicate, the subscriber may be interested in more than one service, each of which may have different service primitives like a data encryption algorithm, public key or initialization vector. Therefore an SA includes cryptographic keys, key lifetime, initialization vectors and/or digital certificates. A new authentication request is required if the SS's keying material has expired. It is the subscriber's responsibility to request this material from the BS. Service disruptions (during reauthorization) are prevented by maintaining successive

generations of the authentication keys with overlapping lifetimes. SSs and BSs must be able to support up to two active authorization keys simultaneously. In response to en(dis)abling service, a BS can dynamically establish and eliminate SAs. This can be done by issuing an SA add message and further establishing a TEK.

7.3. *Other Security Issues*

As suggested by Xu *et al.,* mobility increases the vulnerability of authentication and key management protocols [18, 19]. Multicast is another issue, where authentication and key management protocols should be revised to facilitate the multicast functions. Yuksel has completed a detailed analysis of the PKM protocol [20]. An important issue also concerns secure roaming and handover procedures, which is studied by Meyer [15].

8. Thoughts for Practitioners

Wireless networks are of growing omnipresence in every facet of our lives and the threats to their security must be uncovered and addressed. Their performance must be maximized. Broadband wireless access mesh networks enable the deployment of multimedia services in rural areas and emergency situations. Applications of broadband wireless networks include emergency communications, last-mile networks and numerous Web-based services. Broadband wireless networks can provide a backup network infrastructure when the wired infrastructure is damaged and unusable. Hurricane Katrina destroyed the telecommunications infrastructure in Louisiana. WiMAX/802.16 was used after the hurricane to link to the Internet isolated rural communities of Louisiana. An important application is wireless last-mile networks that deliver data at high speed to subscribers that cannot be serviced by digital subscriber loops, because it is not economically viable. The high speed of WiMAX/802.16 addresses the needs of a wide range of subscribers, including individuals, public services, and enterprises. Potential services comprise fast Internet access, high quality audio and video communications, education and entertainment, tele-medicine, tele-metering and tele-surveillance.

Standardization organizations have established a framework for broadband wireless networks. However, several issues are intentionally left open and their solutions lead to differentiation factors between suppliers of equipment. Research is required to support this growing area.

9. Directions for Future Research

Operation in mesh modes raises security issues. Regarding authentication, authorization and accounting, the central server approach is not appropriate in the mesh mode. Mobility support uses the PKMv2 for key management, but in the mesh mode key management is limited to PKMv1, which makes subscribers theoretically vulnerable to the following attacks: sponsor node impersonation and

person-in-the-middle (because of the absence of mutual authentication), message replay (because of the absence of nonce in messages) and denial of service and message modification attacks (because of the absence of authentication codes in messages, sinkhole and wormhole (because network topology information is broadcasted), security level rollback (because of the security capability negotiation process), neighbor bandwidth reduction (by faking resource allocation requests) and lack of privacy during authentication in mesh mode. Research is required to address all these security vulnerabilities.

Either a centralized or a distributed scheduling scheme is possible in the mesh mode. For distributed scheduling, the standard defines an election algorithm for the allocation of control subframes, which are used to coordinate the scheduling of data frames. In general, the scheduling problem is fairly open and can be addressed according to various and conflicting performance criteria.

10. Conclusions

In this chapter, we have given an overview of WiMAX/802.16 networking for broadband wireless access. We have discussed how it compares with other wireless technologies such as WiFi/802.11. WiMAX/802.16 operates in both PMP and mesh mode. The specifications define the PHY and MAC layers. The PHY layer is responsible for sending the individual bits across the air interface. There are multiple PHY layers depending on intended use. The MAC layer defines the access to the PHY layer from the higher layer protocols. The MAC layer has many responsibilities including network entry, scheduling transmissions and selection of burst profiles for the PHY layer, providing QoS for data flows, as well as supporting mobility.

We have described mobility support, discussing the MS association procedures, as well as handovers. We have described mesh mode operation, the mesh network entry procedure and scheduling, and we have discussed the new initiative of IEEE 802.16j, multihop relay, which aims to improve coverage and connectivity through the deployment of relay stations. Relays can be deployed based upon several usage models. Finally, we have discussed the security model, including security associations, packet encryption protocol, privacy key management protocol and authentication procedures.

Terminology

Base station (BS): A piece of equipment that provides connectivity, management and control of SSs.

Broadband Wireless Access (BWA): Wireless access providing broadband connections, typically with bandwidth greater than one MHz supporting data rates of at least 1.5 Mbps.

Burst profile: A set of parameters that describe the physical characteristics of the uplink and downlink channel. This includes such things as modulation type and forward error correction type.

Frequency Division Duplex (FDD): Transmission scheme where uplink and downlink transmissions can occur simultaneously, but on different frequencies.

Handover: The process where an MS changes its point of attachment to the network from one BS to a neighboring BS, usually due to the mobility of the MS.

Point-to-Multipoint (PMP): The typical cellular telephony model where a BS is at the center and communicates with one or more SSs that are within communications range.

Mobile station (MS): A subscriber station that is intended to be used while in motion.

Security association (SA): The set of security information that is shared between the BS and an SS in order to provide secure communications.

Service access point (SAP): The place in the protocol stack where the services of a lower level layer are accessible from the one above.

Subscriber station (SS): The piece of equipment that provides connectivity with a BS.

Time Division Duplex (TDD): A transmission scheme where uplink and downlink transmissions share a common frequency but occur at different times.

Exercises

1. QPSK, 16-QAM and 64-QAM use respectively four, 16 and 64 different pulse values. How many bits per pulse can be encoded with each modulation scheme?
2. In WiMAX/802.16, a downlink sub-frame consists of three data bursts: a QPSK burst, a 16-QAM burst and a 64-QAM burst. From first to last, in what order must these bursts be sent? Why is this ordering required?
3. In Example 2, what is the share of the bandwidth for each channel? What is the value of the bandwidth in Hertz for each channel? How much bandwidth is lost for control purposes, on the downlink? On the uplink?
4. Describe some advantages and disadvantages of centralized versus distributed scheduling in the WiMAX/802.16 mesh mode.
5. WiMAX/802.16 supports three forms of authentication: device authentication, user authentication and message authentication. Why is each form of authentication required? Explain the authentication procedure for each case.
6. Compare WiMAX/802.16 with Long Term Evolution. Consider technical criteria such as speed and radio spectrum.

7. Let Fi be the geometric region covered by frequency fi, $i = 1, 2, \ldots, k$. Given that the mobile user attempts to connect to the base station by selecting frequencies in the order $f1, f2, \ldots, fk$, the expected number of attempts until a successful connection is established is given by the formula

$$E(f1, f2, \ldots, fk) = \sum_{i=1}^{k} i \cdot \Pr[Fi \backslash (F1 \cup \cdots \cup Fi - 1)]$$
$$= \sum_{i=1}^{k} i \cdot \Pr[Fi] - \sum_{i=1}^{k} i \cdot \Pr[Fi \cap (F1 \cup \cdots \cup Fi - 1)].$$

   (a) Write the formula for the special case of two and three frequencies (i.e., $k = 2, 3$) and discuss its meaning.
   (b) Assume the three regions are the circles F_1, F_2, F_3 depicted in Fig. 17 and that frequency $fi$is available only within circle Fi. Explain what frequencies are available at the corresponding intersections. (*Hint*: There are four intersections.)
   (c) A given mobile is moving within the union of the three regions and occupies a position with the uniform distribution. Compute $\Pr[Fi]$ for $i = 1, 2, 3$. (*Hint*: Find the area delimited by the union of the three circles as well as the corresponding intersections. Observe that the desired probability is equal to the proportion of area covered by the frequency.)
   (d) What is the expected number of attempts before a frequency is found by the mobile station. [*Hint*: Compute $E(f_1, f_2, f_3)$ for the regions depicted in Fig. 17.]

8. For any permutation σ of $1, 2, \ldots, k$ let us define

$$E\sigma(f_1, f_2, \ldots, f_k) = E(f\sigma_{(1)}, f\sigma_{(2)}, \ldots, f\sigma_{(k)}).$$

   Why should we be interested in the problem of finding a permutation that minimizes the expected number of attempts $E\sigma(f_1, f_2, \ldots, f_k)$? (*Hint*: Give an example of three circles of unequal size.)

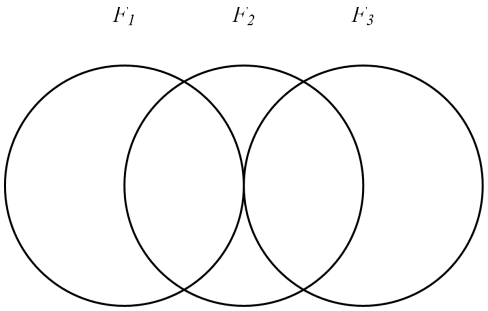


Fig. 17. Three circular regions.

9. Assume that we have two frequencies, i.e., $k = 2$. There are two possibilities for the user: either (i) to select frequencies in the order f_1, f_2 or (ii) in the order f_2, f_1. What is a good frequency selection criterion? (*Hint*: The expected number of steps until successful connection is established is

$$E(f_1, f_2) = \Pr[F_1] + 2(\Pr[F_2] - \Pr[F_2 \cap F_1]),$$
$$E(f_2, f_1) = \Pr[F_2] + 2(\Pr[F_1] - \Pr[F_1 \cap F_2])$$

in the two cases, respectively. This leads to us selecting the most likely frequency.)

10. Show that the competitive ratio of the previously described strategy is at most 2. (*Hint*: Given that $\Pr[F_1] \geq \Pr[F_2]$, we can easily find an upper bound on the competitive ratio between the worst and best choices

$$\frac{E(f_2, f_1)}{E(f_1, f_2)} = \frac{\Pr[F_2] + 2\Pr[F_1] - 2\Pr[F_1 \cap F_2]}{\Pr[F_1] + 2\Pr[F_2] - 2\Pr[F_2 \cap F_1]} \leq 2.$$

In other words, in the worst case the mobile user will take at most twice as long as the best case.)

11. Assume the regions are pairwise disjoint, i.e., $Fi \cap Fj = \emptyset$, for $i \neq j$.
    (a) Give the formula for $E(F_1, F_2, \ldots, F_k)$.
    (b) Show that for the case of k frequencies and pairwise mutually disjoint regions, a mobile user will minimize the expected number of attempts until a successful connection is established by selecting a permutation σ such that

$$\Pr[F\sigma(1)] \geq \Pr[F\sigma(2)] \geq \cdots \geq \Pr[F\sigma(k)].$$

    The permutation σ having this property defines the monotone strategy.
    (c) The mobile user cannot know the permutation that will optimize the expected number of steps (i.e., the monotone strategy). Conclude from this that the competitive ratio of the strategy in the previous exercise is at most k.

Acknowledgments

The authors gratefully acknowledge the financial support received from Mathematics of Information Technology and Complex Systems (MITACS) and Natural Sciences and Engineering Research Council of Canada (NSERC). The authors also thank Christine Laurendeau for her comments about the text of this chapter.

References

1. J. G. Andrews, A. Ghosh and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking* (Prentice Hall PTR, 2007).

2. P. Djukic and S. Valaee, 802.16 mesh networking, in *WiMAX*: *Standards and Security*, eds. S. Ahson and M. Ilyas, (CRC Press, 2007), pp. 147–174.
3. C. Eklund, R. B. Marks, K. L. Stanwood and S. Wang, IEEE standard 802.16: A technical overview of the wirelessman air interface for broadband wireless access, *IEEE Commun. Mag.*, June, 98–107 (2002).
4. R. Fantacci, L. Maccari, T. Pecorella and F. Frosali, A Secure and Performant Token-Based Authentication for Infrastructure and Mesh 802.1X Networks IEEE Infocom Poster (2006).
5. IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed broadband wireless access systems, October, IEEE Standard 802.16-2004, Revision of IEEE Std. 802.16-2001 (2004).
6. IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed broadband wireless access systems, Amendment 2: Medium Access Control Modifications and Additional Physical Layer Specifications for 2–11 GHz, IEEE Standard 802.16a-2003 (2003).
7. IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed broadband wireless access systems, Amendment 1: Detailed System Profiles for 10–66 GHz, IEEE Standard 802.16c-2002 (2002).
8. IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed broadband wireless access systems, Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, February, IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004) (2006).
9. IEEE 802.16j, IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air interface for fixed and mobile broadband wireless access systems, Multihop Relay Specification, Unapproved draft document, June (2007).
10. D. Johnston and J. Walker, Overview of IEEE 802.16 security, *IEEE Secur. Priv.* May/June, 40–48 (2004).
11. I. Koffman and V. Roman, Broadband wireless access solutions based on OFDM access in IEEE 802.16, *IEEE Commun. Mag.* **40**(4), 96–103 (2002).
12. H. Labiod, H. Afifi and C. De Santis, *Wi-Fi, Bluetooth, Zigbee and WiMAX* (Springer, 2007).
13. C. Laurendeau and M. Barbeau, Threats to security in DSRC/WAVE, *5th Int. Conf. Ad-hoc Networks*, Lecture Notes in Computer Science (Springer, Berlin, 2006), pp. 266–279.
14. L. Maccari, M. Paoli and R. Fantacci, Security analysis of IEEE 802.16. *IEEE Int. Conf. Communications* (*ICC' 07*), 24–28 June (2007), pp. 1160–1165.
15. U. Meyer, Secure roaming and handover procedures in wireless access networks, dissertation, Fachgebiet Informatik der Technischen Universitaet Darmstadt (2005).
16. F. A. Perez, Security in Current Commercial Wireless Networks: A Survey (2006).
17. J. Sydir (ed.), *Harmonized Contribution on 802.16j (Mobile Multihop Relay) Usage Models*, June IEEE 802.16 Relay Task Group, http://www.ieee802.org/16/relay/ (2007).
18. S. Xu and C.-T. Huang, Attacks on PKM protocols of IEEE 802.16 and its later versions, *Wireless Communication Systems* (*ISWCS '06*), 6–8 September (2006), pp. 185–189.
19. S. Xu, M. Matthews and C.-T. Huang, Security Issues in Privacy and Key Management Protocols of IEEE 802.16, *ACM SE06*, 10–12 March, Melbourne, Florida, USA (2006).
20. E. Yuksel, Analysis of the PKMv2 protocol in IEEE 802.16e-2005 using static analysis, thesis, Technical University of Denmark, Lyngby, February (2007).