

Analysis and Performance Evaluation of IEEE 802.16 Enhanced with Decentrally Controlled Relays

Pavel Mach, Robert Bestak

Department of Telecommunications Engineering
Czech Technical University in Prague, Faculty of Electrical Engineering
Prague, Czech Republic
machp2@fel.cvut.cz, bestar1@fel.cvut.cz

Abstract— In order to enhance throughput and to extend coverage of wireless networks based on IEEE 802.16 standards, relay stations may be implemented. This paper primarily focuses on decentrally controlled relays when a relay itself schedules all transmissions for its subordinate stations. Needed system resources are allocated to individual relays by means of already defined scheduling services such as UGS, rtPS or ertPS. The paper analyzes suitability of scheduling services. Studied performance metrics include system throughput, signaling overhead and packet delay depending on offered traffic load. The obtained results indicate that ertPS service considerably outperforms UGS or rtPS.

Keywords— IEEE 802.16; Decentrally Controlled Relay; Scheduling; UGS; rtPS; ertPS

I. INTRODUCTION

WiMAX technology is widely known as wireless networking standard that addresses interoperability across IEEE 802.16 standard-based products. So far, IEEE Std. 802.16-2004 [1] intended for fixed subscriber stations (SSs) was approved. To provide mobility support for users, an amendment to former standard labeled as IEEE 802.16e [2] was ratified at the end of 2005. Its purpose is to enrich WiMAX by further features such as handover and power management modes to enable user's mobility.

In order to facilitate QoS requirements of individual users, five QoS scheduling services are specified; i) UGS (Unsolicited Grant Service), ii) rtPS (real time Polling Service), iii) ertPS (extended real time Polling Service), iv) nrtPS (non real time Polling Service) and last one v) BE (Best Effort). A large number of recent studies deal with QoS and scheduling issues in WiMAX networks (e.g. [3-5]).

To enhance the overall system performance both in throughput and in coverage, relay stations (RSs) can be introduced into WiMAX system. Implementation of RSs is within the scope of IEEE 802.16j [6]. The RSs are in most cases build in, owned and controlled by service provider. An RS is not directly connected to wire infrastructure and has the minimum functionalities to support multihop communication. To integrate RSs into WiMAX, modifications of the original IEEE 802.16 frame structure has to be modified (e.g., [7-10]). The main objective is to remain compatible with the legacy WiMAX devices, i.e. no modifications to SS are allowed.

The RSs may be classified by two aspects. The first aspect distinguishes if RS's position is fixed, nomadic or mobile. While the fixed and nomadic RSs position is assumed to be fixed when operating, the mobile RSs are implemented at moving object, for example buses or trains. The second aspect differentiates RSs by their utilization. A RS can be either transparent or non-transparent. The transparent RS (T-RS) does not transmit control information at the beginning of frame. Thus, a SS has to be in the Base Station (BS) range. Consequently, the only purpose of T-RS is to enhance throughput within BS cell. Therefore, a T-RS is implemented in cooperation scenarios when data are sent through several independent radio channels. In comparison with the T-RS, a non-transparent RS (NT-RSs) sends control information at the beginning of every frame similarly as the BS does. Hence, the NT-RSs are suitable for scenarios where SS or MS (Mobile Station) are not able to receive BS's signal with sufficient quality, for example due to shadowing or distance. Nonetheless, the NT-RS can be used for enhancement of throughput as well.

The NT-RSs are further divided on a centrally controlled RS (CC-RS) and de-centrally controlled (distributed) RS (DC-RS). The CC-RS is fully controlled by the BS. This means that the BS handles and schedules all data and controls transmissions between the RS and its own users. On the other hand, when we consider DC-RS, the RS itself (without BS help) schedules data transmission of its users.

According to [10], the DC-RS has to ask for radio resources by means of existing scheduling services such as UGS, rtPS or ertPS. As these services are primarily intended for simple Point to Multipoint (PMP) architecture their suitability for RSs should be analyzed in more details. This paper evaluates performance of individual scheduling services. The considered metrics include overall system throughput, signaling overhead and packet delay.

The rest of the paper is structured accordingly. The next section describes frame structure of DC-RS. The third section analyzes individual scheduling services and contemplates their suitability for allocation of RS's resources. Simulation scenario and obtained results are discussed in section fourth. The last section concludes the paper and briefly describes our future work.

II. FRAME STRUCTURE OF DECENTRALLY CONTROLLED RELAY

The implementation of DC-RS into the system is depicted in Fig. 1. The example assumes the BS enhanced by three RSs. While the RS1 and RS2 are attached directly to the BS, the RS3 is connected to the BS via RS2. Though, OFDM is considered for better understanding, the whole concept may also be extended to OFDMA.

To guarantee a compatibility with the existing IEEE 802.16 standards, the frame is divided into downlink (DL) and uplink (UL) subframes. At the beginning of every MAC frame, the BS broadcasts its control information to all subordinate stations within communication range. The broadcast part is composed of long preamble, FCH (Frame Control Header) field and broadcasted MAC management messages such as DL/UL maps and DCD (Downlink Channel Descriptor)/UCD (Uplink Channel Descriptor). Subsequently, the BS transmits data to stations on the first hop, i.e. to MSs and RSs. At the end of DL subframe a short gap is scheduled. The gap enables antennas to switch between transmitting to receiving mode and vice versa.

The UL subframe is initiated by contention interval (CI). Within a CI, users can ask for additional bandwidth or initiate network entry procedure. The CI is followed by individual UL data bursts. Besides regular UL bursts, certain part of frame called Multihop Subframe (MuS) is dedicated to RSs. For every RS in the system one MuS has to be scheduled. To satisfy backward compatibility, the structure of MuS is exactly the same as described in the IEEE 802.16 standards.

The MuS is assigned to the RS by means of particular scheduling services that make possible to allocate radio resources in periodic manner; UGS, rtPS or ertPS. The periodicity of MuS is necessary as RS has to periodically broadcast control information. Since every scheduling service has its strong and weak points, further analysis of their suitability is addressed in the next section.

III. ANALYSIS OF MULTIHOP SUBFRAME APPLICATION

A. UGS

The basic idea behind UGS is to assign in the UL a constant amount of radio resources on periodic bases. The meaning of "unsolicited" is that allocated intervals are given to a MS automatically, i.e. MS does not have to send any bandwidth requests. This leads to saving of system resources as no signaling overhead is introduced and to minimizing of packet delay in the UL direction. Consequently, UGS is designed to support service flows which generate fixed data packets at strictly periodic intervals, e.g., E1/T1 connections or VoIP (Voice over IP) without silence suppression. The mandatory

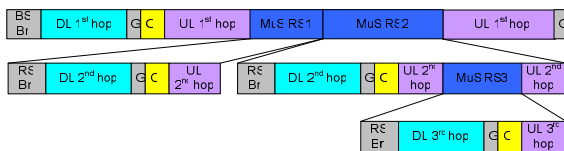


Figure 2. Frame structure for DC-RS

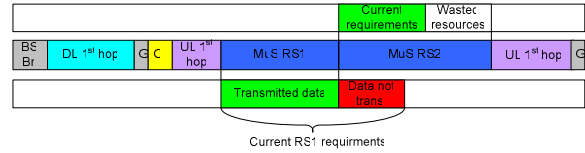


Figure 1. MuS scheduled by means of UGS

QoS requirements for UGS are maximum sustained traffic rate, maximum latency, tolerated jitter and request/ transmission policy (for more details see [1]).

Fig. 2 illustrates what happen if MuSs are assigned to RSs through UGS service. If the traffic load of individual RS is low, the system performs well. The important facts are that no data are lost and packet delay is very short (in comparison with rtPS and ertPS). However, as system load increases the shortage of UGS become evident. The problem is indicated in Fig. 2. While RS2 is able to cope with current requirements of its subordinate station, RS1 is at the same time overloaded. Consequently, the data are either lost (discarded by the RS/MS in the DL/UL direction) or stored in stations memory (prepared for transmission in time when resources are available). Anyway, if requirements on the RS are high for a long time, the data may be lost.

The above mentioned drawback is due to UGS's original purpose which is to serve traffic flows with constant bit-rate. As RS's requirements significantly vary in time, the system performs poorly when the conventional UGS is used. This setback cannot be permanently solved but may be mitigated by appropriate allocation of MuS.

B. rtPS

In comparison with UGS, rtPS scheduling service is much more effective in exploitation of given radio resources. While UGS always allocates fixed intervals, rtPS assigns to every traffic flow intervals with a variable size. To be able to support rtPS service, the BS provides a periodic unicast request opportunities (RO). During a RO, the MS can ask for required bandwidth. The mandatory QoS requirements for rtPS are maximum sustained traffic rate, minimum reserved traffic rate, maximum latency and request/transmission policy [1]. Compare to UGS, use of rtPS brings two drawbacks; i) signaling overhead as the MS has to send bandwidth requests with current flow requirement and ii) longer packet delays in the UL direction. Hence, rtPS is designated to support real time service flows which generate variable data packets (e.g., MPEG video).

When rtPS scheduling service is used for MuS, the BS is able to cope with dynamic RS's requirements on radio resources (see Fig. 3). As a result, no radio resources are needlessly wasted when compare to UGS. Nevertheless, as

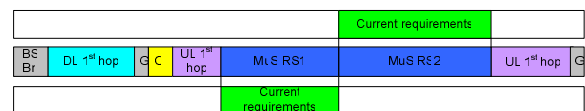


Figure 3. MuS scheduled by means of rtPS

already mentioned before, the rtPS constrain represents inherent long packet delay. This issue is emphasized when multihop communication is employed and data have to travel through more than two hops. This is a drawback especially for data streams that require short packet duration (e.g., VoIP packets).

C. ertPS

The last scheduling service that make possible to send data packets at periodic intervals is called ertPS. The ertPS is a sort of hybrid scheduling service combining the best features of UGS and rtPS. The principle of ertPS is similar to the UGS one; the BS assigns to the flows unsolicited allocation of fixed size intervals. However, the allocation does not have to be necessarily of fixed size, the allocation can vary in time. The change of allocation size is managed either by extended piggyback request, BR field in the MAC signaling header, or over CQICH (Channel Quality Information Channel) [2]. This way, packet delay and signaling overhead are kept at reasonable level. Thus, ertPS is reserved for real time data flows that generate packets of variable size (e.g., VoIP with silence suppression). The mandatory QoS requirements for rtPS are the same as in case of rtPS.

If ertPS is applied for MuS, the radio resources are distributed among the RSs in the most efficient manner similarly to rtPS case (see Fig. 3). The main advantage of ertPS is that there is no need to ask every frame for allocation. Thus, signaling overhead and packet delay are minimized. However, this is true just as long as RS requirements do not change rapidly. If this condition is not fulfilled, the adjustment of MuS has to be done very often (in worst case scenario every MAC frame). In this case, the performance of ertPS may be comparable to rtPS one.

IV. SIMULATIONS

A. Simulation model

To determine how individual scheduling service (UGS, rtPS, ertPS) performs when used for RS's MuS, MATLAB system level simulator has been developed. The used parameters during simulation are summarized in Tab 1. As indicated in Fig. 4, the simulation model is composed of one BS and eight fixed RSs. Additionally, individual mutual connections (i.e. possible routes to/from the BS) and distances are depicted. The maximum distance between the RS and BS is restricted to two hops. The RSs positions are chosen in such way that all MSs are always in a transmission range of at the least one station (BS or RS).

There is implemented a mobility model for every MS movement. At the beginning of simulation, an initial position of each MS is randomly determined in such a manner, that the MS is located within a defined range, i.e. between 0 to 800 m from the BS. Additionally, a velocity and random movement direction are determined for all individual MSs in the system. The MSs are moving along straight line until the distance from the BS is equal or larger than defined BS cell area. In such circumstance a new MS direction is established. This

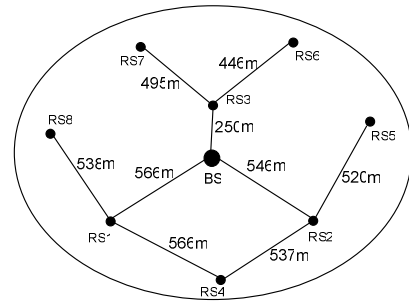


Figure 4. Deployment of RSs within BS cell

mechanism guarantees, that no MS moves out of the BS range during the simulation process.

The path between the BS and MS is determined according to the minimum Radio Resource Cost (RRC) metric (more detail may be found in [11]). The RRC is measured in a number of OFDM symbols needed for transmission of certain amount of data burst (e.g., 1000 bits). To decide which point of attachment is the best for the system performance, the RRC compares all available routes from (to) the BS and determines how much system resources have to be allocated.

There are implemented four traffic models. The first traffic model (in simulation labeled as Traffic model I) is based on VoIP with suppression of silence interval as defined in [12]. The size of packets generated during active/inactive state is in simulation denoted as AS/IS (see Tab. 1). The packet size covers both, voice payload and protocol headers (RTP, UDP, IPv4, 802.16 generic MAC header and CRC). As performance of ertPS strongly dependent on the rate of traffic size changes (signaling overhead and packet delay), other three traffic models shortening active/inactive interval are implemented. Average duration of active/inactive state is depicted in Tab. 2.

As already stated, UGS performance consists in the appropriate MuS size allocations. For this purpose, four types of UGS (labeled from UGS-A to UGS-D) characterized by various RS's allocation size are considered. UGS-A/B/C assume that all RSs in the system have the same size of MuS

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Frequency band [GHz]	3.5
Channel bandwidth [MHz]	20
Number of MS [-]	1-100
MS's velocity [m/s]	10-50
Frame duration [ms]	20
BS transmit power P_t [dBm]/height [m]	30/30
RS transmit power P_t [dBm]/height [m]	30/30
MS transmit power P_t [dBm]/height [m]	30/2
BS cell area [m]	800
Max number of hops between the BS and MS [-]	3
Channel model between BS-RS, RS-RS	LOS [12]
Channel model between BS-MS, RS-MS	NLOS [12]
Length of simulation [min.]	30
Noise [dBm]	-100.97
Packet size during active (AS)/inactive state (IS) [b]	696/456

TABLE II. PARAMETERS OF TRAFFIC MODELS

Traffic model type	Average active state duration [ms]	Average inactive state duration [ms]
I	1000	1500
II	200	300
III	100	150
IV	67	100

which corresponds to 5%/7.5%/10% of overall BS resources. UGS-D takes into account RS's hop distance from the BS, thus RSs on the first hop have 10% of overall resources and RSs on the second hop dispose with 5% of resources. Note if the size of MuS is smaller than 5%, an insufficient amount of radio resources are allocated to the RS. On the other hand, more than 10% for MuS allocation is too much since no resources are reserved for the BS (8 RSs is assumed in the simulations).

The performance metrics used in the simulation are system throughput, packets delay in DL and UL directions and signaling overhead. A major factor which influences the behavior of individual scheduling services represents the offered traffic load, i.e. number of active MSs in the system. The maximal offered traffic (MOT) considered in the simulation corresponds to situation when all MSs in the system are active with two VoIP connections. The system throughput is derived from the ratio between overall amounts of transferred data TT^L and offered traffic load OT^L during the whole simulation time. The parameter OT^L can be derived from individual MS's requirement MSR_i which can be expressed as:

$$MSR_i^L = \int_0^t MSD_i(t) dt. \quad (1)$$

where t is the length of simulation and $MSD_i(t)$ is amount of data generated by i -th MS during t . Since VoIP packets are generated at periodic interval, Eq. 1 can be rewritten as:

$$MSR_i^L = (n_i^L * AS + m_i^L * IS). \quad (2)$$

where n_i^L/m_i^L is number of packets generated by i -th MS during active/inactive state. By summation of individual MSR_i , total offered traffic load OT^L can be derived from:

$$OT^L = \sum_{i=1}^k (n_i^L * AS + m_i^L * IS). \quad (3)$$

where k is the number of currently active MS. The quantity of data not transferred within simulation time LT^L can be expressed as:

$$LT^L = \sum_{i=1}^k (n_{LOST_i}^L * AS + m_{LOST_i}^L * IS). \quad (4)$$

where $n_{LOST_i}^L/m_{LOST_i}^L$ represents number of lost packets. Finally, overall transferred throughput TT^L within the whole simulation time is given as:

$$TT^L = \frac{OT^L - LT^L}{OT^L}. \quad (5)$$

The packet delay considered in simulation corresponds to a time interval between packet arrival and reception at the station's MAC layer. Thus, delays introduced by higher protocol layers and the rest of network are not considered. Furthermore, the packet delay is evaluated under assumption that MSs use ertPS scheduling services in the uplink directions. Generally, the packet delay pd is composed of queuing time t_q and transfer time t_t and can be formulated as:

$$pd = t_q + t_t. \quad (6)$$

Note that packets eventually discarded by the system due to congestion are not considered in pd . Thus, only packets successfully arriving at the destination are assumed. The signaling overhead so caused by RS's MuS size modification is calculated in a number of requests per second. The parameter so is derived as follows:

$$so = NoR^T / T. \quad (7)$$

where NoR^T corresponds to a number of requests per simulation time and T represents duration of simulation in seconds.

B. Simulation results

Fig. 5 depicts the transferred throughput depending on the offered traffic load for each scheduling service. The best results are achieved by ertPS which outperforms both rtPS and UGS. Since the results for ertPS with different traffic model are comparable, only the first traffic model is considered here. While difference between ertPS and rtPS is insignificant, all types of analyzed UGSs provide poor performance. A marginal difference between rtPS and ertPS is due to higher signaling overhead of rtPS. However, as long as offered traffic load is kept below 70% of MOT, the system utilizing rtPS or ertPS is able to serve all MS's demands, i.e. no data are lost. If the traffic load increases above 70%, even system with rtPS and ertPS is incapable to carry all offered data packets. The system saturation is reached approximately at 80% of MOT.

Fig. 5 also shows that no matter which MuS's allocation is considered, the system is saturated much sooner when compare

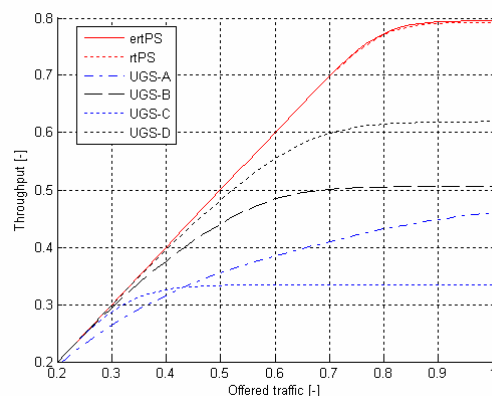


Figure 5. Transferred system throughput per individual scheduling service

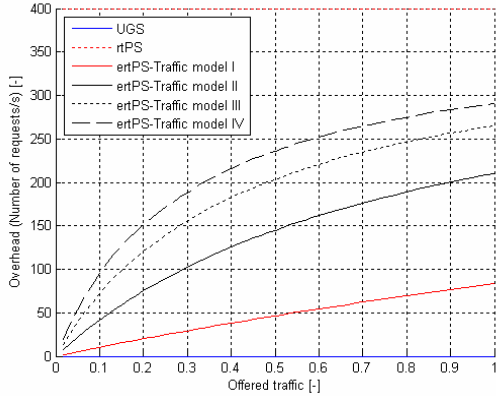


Figure 7. Signaling overhead per individual scheduling services

to (e)rtPS (between 33% to 62% of MOT). Generally, the smaller RS's MuS is used, the smoother the saturation is. If the MuS duration is small (UGS-A), significant amount of packets are discarded as RSs are unable to cope with all users requirements. On the other hand, when a lot of resources are dedicated directly to RSs (UGS-C), the BS manages a few resources. But, as soon as the BS is congested, the whole system is as well as congested since data can not be transmitted between BS and any other station. In case of UGS, the best performance is achieved with UGS-D type which takes into account that RSs on the first hop need more radio resources than RSs on the second hop. This is due to fact that RSs closer to the BS serve as aggregation points to its subordinate stations.

Another aspect considered during the simulation process is signaling overhead depicted in Fig. 6. As RSs employing rtPS has to send the request every frame, rtPS produces the highest overhead that is constant regardless the current offered traffic load. On the other hand, by implementing UGS, the system resources are saved as no signaling overhead is generated. In case of ertPS, the overhead strongly depends on individual traffic model. If the traffic model I is used, the results are only slightly worse in comparison with UGS. However, as RSs requirements vary more frequently, the advantage of ertPS over rtPS is smaller.

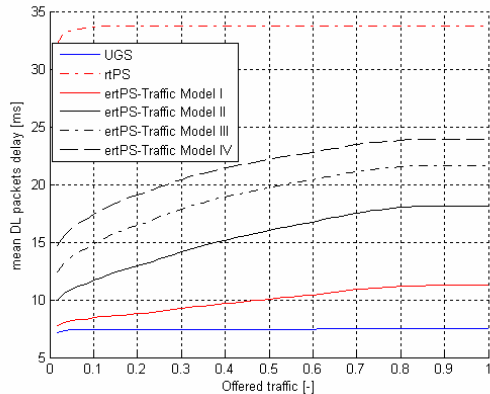


Figure 8. Mean DL packet delay per individual scheduling services

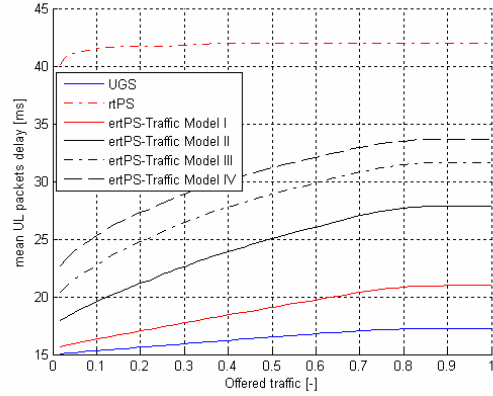


Figure 6. Mean UL packet delay per individual scheduling services

Fig. 7 and Fig. 8 illustrate the mean packet delay for the DL and UL direction depending on the offered traffic load. Similarly as with the signaling overhead, the best performance provides UGS. Packets are delivered at most within 8ms in the DL and up to 17ms in the UL direction. The highest mean packet delays are obtained for rtPS service, i.e. up to 34ms for DL and 42ms for UL. In case of ertPS, the mean packet delay varies according to the offered traffic load; from 8ms to 25ms for the DL and from 16ms to 34ms for the UL direction.

V. CONCLUSIONS

The paper focuses on WiMAX networks enhanced by decentrally controlled RS (DC-RS). DC-RSs are practically independent on the BS and serve its own subordinate stations on its own. System resources are allocated to DC-RS by means of scheduling services such as UGS, rtPS or ertPS. The paper analyzes suitability of these services for MuS's allocation.

Simulation results show that the conventional UGS is not suitable as a lot of radio resources are wasted, thus the overall system capacity heavily decreases. The performance can be partially improved by proper distribution of resources. The best results are obtained when using the ertPS service. In case of ertPS utilization, the system throughput is significantly higher while signaling overhead and packet delay are sufficiently small. The advantage of ertPS remains as long as RS's requirements don't change too often. In the worst case scenario, when the RS asks for MuS size modification every frame, the ertPS and rtPS performance is similar.

In our next work we want to investigate performance of individual scheduling service when OFDMA (instead simple OFDM) is implemented.

ACKNOWLEDGMENT

This work has been performed in the framework of the FP7 project ROCKET IST-215282 STP, which is funded by the EC. The Authors would like to acknowledge the contributions of their colleagues from ROCKET Consortium (<http://www.ict-rocket.eu>).

REFERENCES

- [1] IEEE Std 802.16-2004, "IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Fixed Broadband Wireless Access Systems", 2004.
- [2] IEEE Std 802.16e-2005, "IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems", 2006.
- [3] K. Wongthavarawat and A. Ganz, "Packet Scheduling for QoS support in IEEE 802.16 Broadband Wireless Access Systems", International Journal of Communication Systems, Vol. 16, p.81-96, Feb. 2003.
- [4] M. Hayajneh, N.A. Ali and I. Khalil, "Uplink and Downlink Scheduling for Point to Multipoint WiMAX Networks", Wireless Communication and Mobile Computing, 2008.
- [5] V. Singh and V. Sharma, "Efficient and Fair Scheduling of Uplink and Downlink in IEEE 802.16 OFDMA networks", IEEE Wireless Communications and Networking Conference (WCNC), p.984-990, 2006.
- [6] IEEE: "IEEE Baseline Document 802.16j-2006, Air Interface for Fixed and Mobile Broadband Wireless Access Systems: Multihop Relay Specification", 2007.
- [7] Z. Tao, A. Li, K.H. Teo and J. Zhang, "Frame Structure Design for IEEE 802.16j Mobile Multihop Relay (MMR) Networks", Global Telecommunications Conference (GLOBECOM), p. 4301-4306, 2007.
- [8] X. Leng et al., "Non-Transparent Relay Structure Extension for Multi-hop (>2 hops) Support", IEEE C802.16j-07/145, 2007.
- [9] D.H. Ahn et al., "Multiframe structure consistent to 802.16e for MR networks", IEEE C802.16-07/162r5, 2007.
- [10] Ch. Hoymann and K. Klagges, "MAC Frame Concepts to Support Multi-hop Communication in IEEE 802.16 Networks", Wireless World Research Forum, 2005.
- [11] Mach, P. and Bestak, R., "WiMAX Throughput Evaluation of Conventional Relaying," Personal Wireless Communications. New York: Springer, 2007, p. 75-86.
- [12] IEEE 802.16j, Multi-hop Relay System Evaluation Methodology (Channel Model and Performance Metric) paper No. 06/013r3, 2007.