

# RLAM: A dynamic and efficient reinforcement learning-based adaptive mapping scheme in mobile WiMAX networks

M. Louta<sup>a,\*</sup>, P. Sarigiannidis<sup>a</sup>, S. Misra<sup>b</sup>, P. Nicopolitidis<sup>c</sup> and G. Papadimitriou<sup>c</sup>

<sup>a</sup>*Department of Informatics and Telecommunications Engineering, University of Western Macedonia, Kozani, Greece*

<sup>b</sup>*School of Information Technology, Indian Institute of Technology, Kharagpur, West Bengal, India*

<sup>c</sup>*Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece*

**Abstract.** WiMAX (Worldwide Interoperability for Microwave Access) constitutes a candidate networking technology towards the 4G vision realization. By adopting the Orthogonal Frequency Division Multiple Access (OFDMA) technique, the latest IEEE 802.16x amendments manage to provide QoS-aware access services with full mobility support. A number of interesting scheduling and mapping schemes have been proposed in research literature. However, they neglect a considerable asset of the OFDMA-based wireless systems: the dynamic adjustment of the downlink-to-uplink width ratio. In order to fully exploit the supported mobile WiMAX features, we design, develop, and evaluate a rigorous adaptive model, which inherits its main aspects from the reinforcement learning field. The model proposed endeavours to efficiently determine the downlink-to-uplink width ratio, on a frame-by-frame basis, taking into account both the downlink and uplink traffic in the Base Station (BS). Extensive evaluation results indicate that the model proposed succeeds in providing quite accurate estimations, keeping the average error rate below 15% with respect to the optimal sub-frame configurations. Additionally, it presents improved performance compared to other learning methods (e.g., learning automata) and notable improvements compared to static schemes that maintain a fixed predefined ratio in terms of service ratio and resource utilization.

Keywords: IEEE 802.16, WiMAX, OFDMA, mapping, channel allocation ratio, learning

## 1. Introduction

Worldwide interoperability for microwave access (WiMAX) constitutes one of the most promising broadband access technologies in the next generation networking environments, supporting high capacity, long-distance communication and users' mobility. WiMAX is based on wireless metropolitan area networking (WMAN) standards developed by the IEEE 802.16 group. IEEE 802.16e standard, referred to as Mobile WiMAX, provides nomadic and terminal mobility support.

---

\*Corresponding author: M. Louta, Department of Informatics and Telecommunications Engineering, University of Western Macedonia, 50100, Kozani, Greece. Tel.: +30 2461056566; mobile: +306977973322; Fax: +30 2461056501; E-mail: louta@uowm.gr.

Mobile WiMAX adopts orthogonal frequency division multiple access (OFDMA), enabling users to utilize concurrently different bandwidth regions in the time and frequency domains. To this end, OFDMA combines time and frequency division multiple access, providing multiple timeslots at different frequencies to multiple users. OFDMA forms strong bonds between the physical (PHY) and the medium access control (MAC) layer, since it is responsible for allocating PHY resources to MAC requests. To this end, the term “slot” is defined as the minimum PHY resource unit that can be allocated to a single subscriber in the time and frequency domains [1].

Full duplex communication is achieved by dividing the entire MAC frame into two sub-frames, the downlink sub-frame carrying data from a base station (BS) towards mobile stations (MSs) and the uplink sub-frame transmitting data from MSs towards the BS [2]. Frequency division duplexing (FDD) or time division duplexing (TDD) technique may be applied for the transmission of the two sub-frames. The TDD is favored by the majority of applications mainly due to its flexibility and simplicity. In the current study, the authors adopt the TDD technique for bi-directional communication as well.

Considering a TDD-based frame, its duration is fixed and pre-defined; in turn the downlink sub-frame is followed by the uplink sub-frame after a specific guard time interval, which aims at avoiding interference between the downlink and uplink traffic. After a short control period used for synchronization and channel estimation at the beginning of a frame and downlink control information providing frame and bandwidth allocation configuration details, downlink allocations follow. The uplink sub-frame comprises control channels such as a ranging channel for performing frequency and power measurements as well as the MSs’ bandwidth requests.

Acquiring an optimal or a near-optimal bandwidth allocation (i.e., optimal assignment of slots to the subscribers’ requests in both the downlink and uplink directions) is not a trivial problem. The target is to exploit the available bandwidth to the extent possible [3]. Additionally, bandwidth assignment should follow certain rules and comply with constraints imposed by the OFDMA technique [4]. Specifically, the available bandwidth is structured as a two-dimensional rectangular shaped allocation bin, having one dimension associated with frequency (height) and the other with time (width) [5]. Moreover, each downlink request (or a set of requests that share common PHY characteristics, referred to as burst [6]) must comply with the rectangular shape. This rule does not apply to the uplink requests, simplifying, thus, the uplink sub-frame allocation process. Hence, in this context, a problem that should be addressed is the accommodation of traffic requests to both downlink and uplink sub-frames.

The BS is responsible for accommodating all MSs’ requests to both uplink and downlink sub-frames. For the downlink sub-frame, the BS decision is based on the needs and characteristics of the incoming traffic, while, for the uplink sub-frame, allocations are based on requests originating from the MSs, involving several mechanisms and polling schemes [7]. The network elements that accomplish this task are the scheduler and the mapper. Specifically, the scheduler receives the requests originating from all MSs and determines an initial bandwidth distribution, taking into account both subscribers’ needs and potential service quality requirements so as to support quality of service (QoS) provisioning [8]. The final bandwidth allocation takes place into the mapper. The mapping process collects the requests from the scheduler and constructs, for each connected MS, a complete transmission program (both downlink and uplink), defining the exact time and frequency of transmitting (uplink) and receiving (downlink) data.

Efficiently utilizing the available bandwidth through the OFDMA technique constitutes an open and active area of research due to the following two reasons: a) the efficiency of scheduling and mapping operations significantly affect network performance and b) the standard does not include specific algorithms to be adopted and the problem’s solution is left open to alternate WiMAX implementations.

In the light of the aforementioned aspects, the mapping process is of paramount importance in the OFDMA-based WiMAX. Recognizing this issue, the standard supports flexible and dynamic adjustment of the downlink-to-uplink ratio, which may be varied from 3:1 to 1:1. This feature could lead to significant performance improvement with respect to the mapping operation, considering different traffic profiles such as real-time and non-real time traffic. However, most existing mapping techniques proposed in related research literature are inflexibly designed, defining static allocations that do not take into account subscribers' demands. Thus, they suffer from either inefficient use of bandwidth in case of low traffic conditions or high number of unserved requests under high traffic.

This paper endeavors to address this weakness by proposing a novel scheme, hereafter referred to as reinforcement learning based adaptive mapping (RLAM), for dynamically adjusting the downlink-to-uplink ratio, by exploiting learning from experience techniques in order to adequately follow traffic dynamics. Specifically, the proposed scheme is based on reinforcement learning, introducing rewards/penalties in order to favor/punish previous actions, depending on the feedback produced from the environment as a response to the actions taken. Thus, progressively, the scheme is enabled to acquire the most appropriate selection of the forthcoming frame's downlink-to-uplink ratio.

The scheme's performance efficiency is examined in terms of OFDMA resources utilization and service ratio, considering also fixed schemes that maintain downlink and uplink sub-frames' configurations static, defined in advance. Additionally, in comparison to a learning automata based solution presented in [9], it exhibits improved performance with respect to the downlink-to-uplink ratio determination accuracy.

The remainder of this paper is organized as follows. Section 2 presents a related research literature overview, briefly describing the mapping techniques proposed. Section 3 provides the fundamental concepts underlying the RLAM scheme. Section 4 presents in detail the dynamic adaptation scheme designed. Section 5 evaluates its performance, having conducted extensive simulation experiments. Finally, Section 6 concludes this paper and highlights our future plans.

## 2. Related research overview

Various scheduling and mapping schemes have been proposed for both uplink and downlink data streams in the related research literature. Downlink bandwidth distribution constitutes a challenging issue that has attracted the attention of the researchers due to the imposed strict shaping restriction (i.e., the requests must be accommodated in the allocation bin forming a two-dimensional rectangular). Thus, research efforts on downlink mapping are more frequently met than in the uplink domain, where this restriction does not apply. Indicatively, the reader could refer to [10–18].

Simple packing algorithm (SPA) [10] is one of the earliest efforts in downlink mapping, designed to accommodate the incoming bandwidth requests within the downlink sub-frame. The scheme involves a top to bottom and left to right slot allocation, accommodating symbols (columns) and sub-channels (rows) for each request in a first in first out (FIFO) discipline, until the requested number of slots is met. If this number is not an integer multiple of the frequency or the time dimension respectively, the remaining unallocated space remains idle. The authors in [11] apply a persistent mapping technique using a binary full search tree, but the final result indicates complex operation, limited to eight users. Other attempts involve, as a first step, an initial request sorting in terms of the number of requested slots and, as a second step, either bucket definition and accommodation [12,13], or heuristic packing algorithms. Concerning bucket definition and accommodation, the combined bursts define buckets [12,13] that are accommodated into the allocation bin in a column by column basis. Heuristic packet algorithms allocate

the incoming requests in a two step procedure, conducting first a horizontal mapping and then vertical accommodation [14]. Adaptive horizon burst mapping (AHBM) algorithm for the downlink sub-frame is presented in [15]. It applies horizon-based allocation, creating initial pilots for the forthcoming requests. Large requests are accommodated first, leaving minimum remaining idle space, while pilots are formed in a right-to-left and bottom-to-top manner. In the sequel, the remaining requests are mapped based on the pilots. In [18] the authors present a burst mapping algorithm for downlink Mobile WiMAX systems based on a tree-map visualization algorithm. However, the so-called sqTM algorithm suffers from high operation running time, restricting its usability only in the 5 ms frame configuration. AHBM presents improved performance with respect to other leading schemes [10,14]. Thus, it is adopted as the main downlink mapping technique for the rest of our study.

The operation of the uplink mapping is much simpler than the downlink one, due to the absence of the rectangular shaping restriction. The simplest and most effective way for accommodating the uplink requests lies on a row-by-row basis. One after the other, uplink requests are accommodated into the uplink sub-frame, without occurrence of row cuts. Upon the accommodation of an uplink request, the following one is sided directly next to it, without leaving gaps (i.e., idle slots). In this manner, the set of uplink requests fill uniformly the allocation bin, until either all requests are mapped or the bin comes full. This fixed and simple uplink mapping technique is also adopted in this study.

### 3. RLAM fundamentals

#### 3.1. Basic concepts, novelty and contribution

The scope of this paper falls within the OFDMA mapping process in mobile WiMAX wireless networks. The efficiency of the mapping operation is of paramount importance, as requests that fail to be mapped are returned to the scheduler delaying their transmission for at least one entire frame, inducing thus network performance degradation in terms of the overall network delay, network throughput and consumption of resources. Efficiently utilizing the available bandwidth through the OFDMA technique has attracted the attention of the research community, since the IEEE 802.16e standard does not comprise specific mapping algorithms, leaving the solution of the mapping problem open to different WiMAX implementations. In this paper, we endeavor to accommodate all downlink and uplink requests to both downlink and uplink sub-frames, utilizing as much as possible the allocation space given, subject to the rules and constraints imposed by the OFDMA technique.

IEEE 802.16e standard supports flexible and dynamic adjustment of the downlink-to-uplink ratio, which can be varied from 3:1 to 1:1, in order to efficiently handle traffic variations. However, to the best of our knowledge, existing mapping approaches do not exploit the aforementioned aspect, defining static allocations that do not take into account subscribers' demands. Thus, they may suffer from either inefficient use of bandwidth in case of low traffic conditions or high number of unserved requests under high traffic. Considering the fact that a wireless network may experience intense traffic changes over time, dynamic adjustment of the downlink-to-uplink ratio constitutes a key mechanism of the mapping process design in order to efficiently and effectively adapt to network traffic differentiation in highly dynamic and unpredictable environments.

Our first attempt to dynamically adjust the downlink-to-uplink width ratio is found in our previous work [15]. Specifically, AHBM involves a prediction tool based on hidden markov chains in order to re-define the length of the downlink sub-frame in accordance with downlink traffic profiles. As a next step,

we exploit a learning automaton [17] in order to sense the traffic conditions and accordingly determine the most appropriate length ratio of both sub-frames in order to maximize the network performance.

In this study, we design, model and evaluate RLAM scheme that applies a reinforcement learning technique in order to efficiently adjust the downlink-to-uplink ratio in accordance with the subscribers' traffic, subject to OFDMA rules and constraints. Specifically, RLAM is responsible for increasing the downlink/uplink sub-frame's width size in terms of symbols (i.e., columns of the allocation bin) in case the requests demand for more capacity. In the same manner, RLAM reduces the width size, if it senses that downlink/uplink traffic tends to lower levels. This task is accomplished by efficiently exploiting feedback received on the efficiency of the mapping operation and forming a reward/penalty evaluative signal that drives RLAM operation in order to favor/punish previous actions (i.e., previously selected sub-frame width sizes). The feedback comprises two important parameters: a) the unserved slots, which refer to the cumulative number of requests that fail to get accommodated at each sub-frame and b) the unused slots, which show the total number of wasted slots within each sub-frame. The combination of these parameters indicate the conditions of both uplink and downlink traffic flows; specifically, more unserved slots indicate more bandwidth needs, while the larger the number of unused slots is the larger bandwidth wastage occurs, which, in essence, means that the corresponding direction could be operational enough even with less network resources. At the same time, the following rules specified by the standard should be satisfied: a) the summation of the downlink and uplink sub-frame width values should be equal to the total frame width and b) the determination of both sub-frames width should not violate the permissible downlink-to-uplink ratios (i.e., 1:1 to 3:1).

RLAM operates on a frame-by-frame basis, re-determining the values of both uplink and downlink sub-frames in the context of the current frame for the forthcoming frame. RLAM accomplishes its task in less than the frame's length (i.e., less than 10ms for the simulation environment considered). Thus, no time restrictions are imposed for RLAM's application to real-time scenarios.

RLAM model is presented in a quite detailed manner, accompanied with analytical equations and expressions, giving emphasis to its operation on handling the input and the feedback produced by the environment and deciding on the best course of action concerning the forthcoming frame. RLAM formulation is quite generic in nature and it can be easily applied to any OFDMA wireless network that enables the dynamic reconfiguration of the downlink-to-uplink width ratio, as is the case with long term evolution (LTE)-advanced, one of the most challenging technologies for the 4G era.

Extensive experiments were conducted in order to evaluate the performance of the proposed scheme. In this study, we obtained and used real WiMAX multimedia traces concerning both the downlink and uplink as input to our custom made simulator. The obtained results indicate the superiority of the proposed scheme in comparison to the learning automata based solution presented in [9] with respect to the downlink-to-uplink ratio determination accuracy.

In a nutshell, the contribution of this paper lies in the following areas. First, the design, definition and mathematical formulation of a reinforcement learning-based adaptive mapping scheme in order to efficiently and effectively re-configure the downlink-to-uplink width ratio on a frame-by-frame basis, subject to OFDMA rules and constraints, in mobile WiMAX wireless networks. Second, RLAM formulation is quite generic in nature and it can be readily applied to any OFDMA wireless network that enables the dynamic reconfiguration of the downlink-to-uplink width ratio, as is the case with LTE-Advanced, one of the most promising technologies for the 4G era. Third, the provision of indicative evidence on the performance and the superiority of the proposed scheme, conducting extensive simulation experiments and using real WiMAX multimedia traces as input to our custom made simulator.

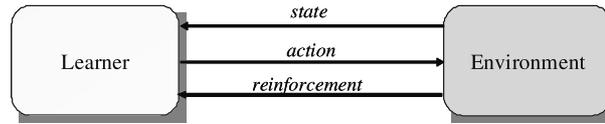


Fig. 1. General structure of a reinforcement learning system.

### 3.2. Reinforcement learning overview

Learning from experience constitutes a vital property of every component presenting a robust, autonomous and adaptive behavior over long periods of time. Learning refers to an entity's ability to take whatever information it has gained about the stochastic environment it is deployed in, and use it in order to modify/improve its own behavior. Incorporating learning capabilities in an entity's intelligence may lead to the development of new improved behavior and avoidance of repeating previous mistaken actions [17]. Reinforcement learning techniques have proved to be a valuable asset for the optimization and performance improvement of various mobile systems and wireless networks. Numerous paradigms exploiting rigorous learning techniques in the context of wireless communication networks could be found in the literature [19–26].

In this study, we define a mechanism for updating the downlink-to-uplink width ratio based on the key concepts that apply to a general reinforcement learning system. The general structure of a learning system based on evaluative reinforcement signals is depicted in Fig. 1. There is a bidirectional information exchange between the environment and the learner. Specifically, the learner acts at each state by selecting a specific action from a pool of possible actions and the environment reacts to the action taken by producing a feedback. The feedback, which controls the learner's behavior, is received by the learner and it is appropriately processed so as to define the next action. The underlying idea is to strengthen the action that produced favorable results in the past, and weaken it otherwise.

Having as a basis the general structure of the reinforcement learning model, we may correspond the learner system to the mapper, which, based on history records, should at each state take the action that would result in the reception of the maximum reinforcement signal. The environment corresponds to the OFDMA resource allocation process (i.e., the mapping process), considering the current users' bandwidth requests in the context of each frame, the state corresponds to the size of the downlink and uplink sub-frames, while the pool of actions is expressed by the available downlink-to-uplink width values in accordance with the adopted coding and sub-channelization techniques. The feedback produced from the environment leads to a reward/penalty evaluative signal (reinforcement signal) that is fed to the mapper, supporting the learning process towards the optimal action (i.e., determination of the most appropriate downlink-to-uplink ratio). The reward/penalty signal considers both the portion of the wasted bandwidth as well as the amount of the bandwidth requests that fail to be served due to lack of physical resources in the context of each single frame, concerning both the uplink as well as the downlink sub-frames. Thus, each decision-action will be evaluated by means of the reward/penalty signal returned from the environment. These analogies are given in Table 1.

## 4. RLAM formulation

Each frame has a fixed duration, hereafter denoted as  $FS$ . Each frame has a fixed height denoted by  $H$ . Let  $DSFS_k$  and  $USFS_k$  denote the size (width) of the downlink and uplink sub-frames  $DSF_k$  and  $USF_k$  in the context of frame ( $k \geq 1$ ). The range of acceptable values for the downlink and uplink sub-frames

Table 1  
Analogies to reinforcement learning model

| Learner              | Mapper   |
|----------------------|--|
| Environment          | OFDMA allocation process considering the current users' bandwidth requests in the context of each frame  |
| State                | Size of the downlink and uplink sub-frame  |
| Action               | Decide on the most appropriate downlink – to – uplink ratio  |
| Reinforcement signal | Reward/Penalty Signal-Function of both the portion of the wasted bandwidth as well as the amount of the bandwidth requests that fail to be served due to lack of physical resources concerning both downlink and uplink sub-frames |

size is denoted as  $[DSFS_{min}, DSFS_{max}]$  and  $[USFS_{min}, USFS_{max}]$ , respectively. At this point it should be noted that both the downlink and uplink sub-frame acceptable size limits are defined in accordance with [28]. Hence, assuming a frame structure comprising 42 symbols (i.e., 42 allocation columns), the range of acceptable values for the downlink and uplink sub-frame is  $[DSFS_{min} = 21, DSFS_{max} = 33]$  and  $[USFS_{min} = 9, USFS_{max} = 21]$ , respectively. For each frame  $k$ , it stands that  $DFSF_k + USFS_k = FS$ .

#### 4.1. Determination of the downlink-to-uplink sub-frame sizes

Concerning the formation of the downlink/uplink sub-frame's size  $DSFS_k/USFS_k$  in the context of frame,  $k > 1$ , the following Eq. (1) may be adopted. At this point it should be noted that as following equations are applied for the determination of the size of both sub-frames; however, for notation simplicity in the following we do not refer to a specific sub-frame (down-link/uplink); instead we denote the sub-frame under consideration as  $SFS_k$ .

$$NSFS_k = NSFS_{k-1} + m_r \cdot l(NSFS_{k-1}) \cdot \{rr_{k-1} - E[rr_{k-1}]\} \tag{1}$$

where  $NSFS_k$  and  $NSFS_{k-1}$  are the normalized downlink/uplink sub-frame sizes in the context of frame  $k$  and  $k - 1$ , after and before the updating procedure, respectively. The  $NSFS_k$  value for all frames belongs to  $[0, 1]$  range and is given as follows:

$$NSFS_k = \frac{(SFS_k - SFS_{min})}{(SFS_{max} - SFS_{min})} \tag{2}$$

It may be observed that the  $NSFS_k$  value for  $SFS_{min}$  equals to 0 and the  $NSFS_k$  value for  $SFS_{max}$  equals to 1. For  $k = 1$ ,  $NSFS_k$  for the downlink sub-frame assumes the normalized minimum acceptable value (that is equal to 0), while, due to the fact that the sum of the downlink and uplink sub-frames should be equal to  $FS$ ,  $NSFS_k$  for the uplink assumes the normalized maximum acceptable value (that is equal to 1).  $rr_{k-1}$  is a (reward/penalty) function reflecting whether the downlink/uplink sub-frame size is adequate for servicing current requests, without however wasting unallocated slots. Thus, in the context of this study, reward function  $rr_{k-1}$  is dependent on two factors, that is slots unused (denoted as  $UnUS$ ) and slots unserved (denoted as  $UnSS$ ) during the previous frame  $k - 1$ , i.e.,  $UnUS_{k-1}$  and  $UnSS_{k-1}$ , respectively. The  $rr_{k-1}$  function may be implemented in several ways. Without loss of generality, it was assumed that the  $rr_{k-1}$  values range from  $-1$  to  $1$ .

$$-1 \leq rr_{k-1} \leq 1 \tag{3}$$

The mathematical description of the  $rr_{k-1}$  function will be given subsequently. In general, the larger the  $rr_{k-1}$  value, the worse the performance of the downlink/uplink sub-frame is with respect to the unserved

and unused slots and therefore the bigger (positive or negative) the influence on the next downlink/uplink sub-frame size.

$E[rr_{k-1}]$  is the mean (expected) value of the  $rr_{k-1}$  variable. Factor  $m_r$  ( $m_r \in (0, 1]$ ) determines the relative significance of the new outcome with respect to the old one. In essence, this value determines the memory of the system. Small  $m_r$  values mean that the memory of the system is large.  $l(NSFS_{k-1})$  is a function of the size of the normalized downlink/uplink sub-frame in the context of frame  $k-1$  and is introduced in order to keep the normalized downlink/uplink sub-frame size within the range  $[0, 1]$  (the respective downlink/uplink sub-frame sizes belong within the range  $[SFS_{min}, SFS_{max}]$ ). In the current version of this study,

$$l(NSFS_{k-1}) = [1 - \exp(1 - NSFS_{k-1})] / (1 - e) \quad (4)$$

for which it stands  $l(NSFS_{k-1}) \rightarrow 1$  and  $l(NSFS_{k-1}) \rightarrow 0$ .

It should be noted that bad performance (or at least deterioration of previous performance) of a specific downlink/uplink sub-frame leads to an alternate sub-frame size (bigger or smaller) due to the modification introduced by the  $\{rr_{k-1} - E[rr_{k-1}]\}$  factor.

#### 4.2. Mathematical description of the reward function

As already mentioned, reward function  $rr_{k-1}$  is dependent on two factors, slots unused  $UnUS_{k-1}$  and slots un-served  $UnSS_{k-1}$  in the context of the downlink/uplink sub-frames during the previous frame  $k-1$ , respectively. We consider the following weighted additive expression for estimating the reward function  $rr_{k-1}$ :

$$rr_{k-1}(UnUS_{k-1}, UnSS_{k-1}) = w_{unused} \cdot rr_{k-1}(UnUS_{k-1}) + w_{unserved} \cdot rr_{k-1}(UnSS_{k-1}) \quad (5)$$

where  $rr_{k-1}(UnUS_{k-1})$  is the part of the reward function that depends on the slots unused and  $rr_{k-1}(UnSS_{k-1})$  is the part of the reward function dependent on the slots unserved in the context of downlink/uplink sub-frames of frame  $k-1$ , respectively. A wide range of functions may be defined simply by introducing different specifications of the  $rr_{k-1}(UnUS_{k-1})$  and  $rr_{k-1}(UnSS_{k-1})$  functions. However,  $rr_{k-1}(UnUS_{k-1})$  and  $rr_{k-1}(UnSS_{k-1})$  functions should satisfy the following constraint:

$$-1 \leq rr_{k-1}(UnUS_{k-1}) \leq 0 \leq rr_{k-1}(UnSS_{k-1}) \leq 1 \quad (6)$$

Function  $rr_{k-1}(UnUS_{k-1})$  should be decremental, while the following constraints should also be satisfied:

$$rr_{k-1}(UnUS_{k-1} = UnUS_{min} = 0) = 0 \quad (7)$$

$$rr_{k-1}(UnUS_{k-1} = UnUS_{max} = SFS_{max} \cdot H) = -1 \quad (8)$$

In a similar manner, function  $rr_{k-1}(UnSS_{k-1})$  should be incremental, while the following constraints should also be satisfied:

$$rr_{k-1}(UnSS_{k-1} = UnSS_{min} = 0) = 0 \quad (9)$$

$$rr_{k-1}(UnSS_{k-1} = UnSS_{max} = SFS_{max} \cdot H) = 1 \quad (10)$$

#### 4.2.1. Formulation of the unused slots-related reward family of functions

We restrict our attention to two families of functions that satisfy the constraints presented in the previous section: exponential and polynomial. Other functions could be defined as well.

$$rr_{k-1}(UnUS_{k-1}) = - \left( \frac{UnUS_{k-1}}{UnUS_{max}} \right)^{1/\theta} \quad (11)$$

$$rr_{k-1}(UnUS_{k-1}) = - \frac{1}{1-e} \cdot \left[ 1 - \exp \left( \frac{UnUS_{k-1}}{UnUS_{max}} \right)^{1/\theta} \right] \quad (12)$$

Expressions (11) and (12) provide a formal model of the exponential and polynomial families of functions, respectively. These families of functions represent an infinite number of different reward/penalty policies, one for each value of  $\theta$ . Parameter  $\theta$  has been included in order to highlight the different patterns of reward/penalty modification with respect to the adopted rate of concession. For example, Boulware policy [29] may be adopted, according to which the reward function is almost the same and equal to 0, until the unused slots reach their maximum value  $UnUS_{max} = SFS_{max} \cdot H$ , whereupon concedes to the reservation value ( $-1$ ). Otherwise, Conceder policy [30] may be exploited, according to which reward function assumes its reservation value ( $-1$ ) even with small increase in the number of unused slots. It should be mentioned that function  $rr_{k-1}(UnUS_{k-1})$  expressed by Eqs (11) and (12) satisfy constraints (7) and (8).

#### 4.2.2. Formulation of the unserved slots-related reward family of functions

The unserved slots-related reward family of functions may be modeled in a similar manner to the unused slots-related reward family functions. The downlink/uplink sub-frame size is expected to become progressively bigger in case the number of unserved slots increases, reaching its maximum value  $SFS_{max}$  (in case only the number of unserved slots is considered) when the  $UnSS_{k-1}$  factor assumes its maximum value (that is  $UnSS_{k-1} = UnSS_{max} = SFS_{max} \cdot H$ ). Expressions (13) and (14) provide a formal model of reward/penalty function  $rr_{k-1}(UnSS_{k-1})$ , for polynomial and exponential families of functions, respectively.

$$rr_{k-1}(UnSS_{k-1}) = \left( \frac{UnSS_{k-1}}{UnSS_{max}} \right)^{1/\theta} \quad (13)$$

$$rr_{k-1}(UnSS_{k-1}) = \frac{1}{1-e} \left[ 1 - \exp \left( \frac{UnSS_{k-1}}{UnSS_{max}} \right)^{1/\theta} \right] \quad (14)$$

It should be mentioned that function  $rr_{k-1}(UnSS_{k-1})$  expressed by Eqs (13) and (14) satisfy constraints (9) and (10). Hereafter, the expressions (11) and (13) are adopted in order to express the formulation of the unused and the unserved slots-related feedback functions.

#### 4.2.3. Definition of the weighting factors

The weighting factors  $w_{unused}$  and  $w_{unserved}$  provide the relative significance of the two factors (i.e., slots unused  $UnUS_{k-1}$  and slots unserved  $UnSS_{k-1}$  in the context of the downlink/uplink sub-frames of frame  $k-1$ , respectively) to the determination of the reward factor  $rr_{k-1}(UnUS_{k-1}, UnSS_{k-1})$ .

The number of unused slots  $UnUS_{k-1}$  represents the wasted bandwidth and tends to reduce the downlink/uplink sub-frame size, while the number of unserved slots  $UnSS_{k-1}$  symbolizes the additional bandwidth that downlink/uplink sub-frame needs and tends to increase the downlink/uplink sub-frame size of the next frame.

The values of the weighting factors belong in the  $[0, 1]$  range, while their sum is equal to 1 (i.e.,  $w_{unused} + w_{unserved} = 1$ ). For fairness reasons, we may assume that equal significance is attributed to both factors (that is  $w_{unused} + w_{unserved} = 0.5$ ). In the more general case, any value in the  $[0, 1]$  range could be assumed. However, in the definition of the weighting factors, the mapping algorithm deficiencies should be taken into account. Specifically, the mapping algorithm may spend a larger portion of the allocation space in order to accommodate a specific number of requested slots, leading, thus, to unused slots. This phenomenon could be attributed to the rectangular shaping restriction that the standard applies. For example, let us assume that the total downlink requests are 450 slots and the mapping algorithm binds 500 slots within the downlink sub-frame. The extra 50 slots are wasted in order to accommodate the requested 450 slots in a rectangular shape. The number of unused slots that are provided due to inefficiencies of the mapping process should not drive for decreasing the sub-frame size (or at least not with the same significance).

Hence, in the light of the aforementioned aspects, we define the utilization factor of the mapping process concerning both directions (downlink and uplink), denoted hereafter as  $uf$ , as follows:

$$uf = \frac{TotalRequestedSlots}{TotalBindedSlots} \quad (15)$$

Considering  $uf = 0.8$  for the downlink sub-frame, the downlink mapping process is 80% effective (i.e., it needs 100 slots to accommodate 80 requested slots). Consequently, the utilization factor expresses the number of unused slots that are needed for serving a specific number of unserved slots in average.

Finally, the weighting factors incorporating the utilization factor are defined as follows:

$$w_{unused} = \frac{0.5 + uf/2}{2} \quad (16)$$

$$w_{unserved} = \frac{0.5 + 1 - uf/2}{2}. \quad (17)$$

As it may be observed, a small utilization factor (i.e., a large number of unused slots as a result of an inefficient mapping process) leads to a decrease in the significance attributed to the number of unused slots, lessening, thus, the impact of the number of the unused slots to the formation of the sub-frame sizes in the context of the next frame.

#### 4.3. Re-adjustment of the downlink-uplink sub-frame sized

Up to this point, the downlink and uplink sub-frame sizes ( $DSFS_k$  and  $USFS_k$ , respectively) are determined in the context of each frame  $k$  independently, applying Eqs (1)–(17). The range of their acceptable values is  $[DSFS_{min}, DSFS_{max}]$  and  $[USFS_{min}, USFS_{max}]$  for the downlink and uplink sub-frames, respectively. However, as already noted, for each frame  $k$ , the following constraint should be satisfied  $DSFS_k + USFS_k = SF$ . Considering  $DSFS_k$  and  $USFS_k$ , we may identify three distinct cases. According to the first case, the summation of the downlink and uplink sub-frame sizes is equal to the frame size, thus, the aforementioned constraint holds. In this case, our scheme forms the frame  $k$  in accordance with the downlink  $DSFS_k$  and uplink  $USFS_k$ . sub-frame sizes, as determined in Sub-section 4.1. In the second

Table 2  
Example: Ten random requests

|                  |           |
|------------------|-----------|
| MS <sub>1</sub>  | 134 slots |
| MS <sub>2</sub>  | 121 slots |
| MS <sub>3</sub>  | 42 slots  |
| MS <sub>4</sub>  | 99 slots  |
| MS <sub>5</sub>  | 151 slots |
| MS <sub>6</sub>  | 15 slots  |
| MS <sub>7</sub>  | 80 slots  |
| MS <sub>8</sub>  | 112 slots |
| MS <sub>9</sub>  | 33 slots  |
| MS <sub>10</sub> | 22 slots  |

and third case, the summation of the downlink and uplink sub-frame sizes is not equal to the frame size (being bigger and less, respectively), thus, in both cases, the two sub-frames need to be re-adjusted in order to meet the frame requirements. Our scheme provides for a balanced re-adjustment of both sub-frames for fairness reasons, while, when necessitated, it takes into account the traffic requests originated in both directions. The final downlink and uplink sub-frame sizes for the formation of the frame  $k$  are the output of the re-adjustment phase. The re-adjustment algorithm applied is described as follows:

### Re-adjustment Algorithm

Step 1: **IF**  $DSFS_k + USFS_k \neq SF$

Step 2: **IF**  $DSFS_k + USFS_k < SF$

Step 3:  $DSFS_k = DSFS_k + \frac{SF - (DSFS_k + USFS_k)}{2}$  and  $USFS_k = USFS_k + \frac{SF - (DSFS_k + USFS_k)}{2}$

Step 4: **IF**  $\text{mod}((SF - (DSFS_k + USFS_k)), 2) \neq 0$

Step 5: **IF**  $UnSS_{k-1,Downlink} > UnSS_{k-1,Uplink}$

Step 6:  $DSFS_k = DSFS_k + 1$

Step 7: **ELSE**

Step 8:  $USFS_k = USFS_k + 1$

Step 9: **END\_IF**

Step 10: **END\_IF**

Step 11: **ELSE\_IF**  $(DSFS_k + USFS_k > SF)$

Step 12:  $DSFS_k = DSFS_k + \frac{SF - (DSFS_k + USFS_k)}{2}$  and  $USFS_k = USFS_k + \frac{SF - (DSFS_k + USFS_k)}{2}$

Step 13: **IF**  $\text{mod}((SF - (DSFS_k + USFS_k)), 2) \neq 0$

Step 14: **IF**  $(UnSS_{k-1,Downlink} > UnSS_{k-1,Uplink})$

Step 15:  $USFS_k = USFS_k - 1$

Step 16: **ELSE**

Step 17:  $DSFS_k = DSFS_k - 1$

Step 18: **END\_IF**

Step 19: **END\_IF**

Step 20: **END\_IF**

Re-adjustment algorithm defines the final downlink and uplink sub-frame sizes for each frame  $k$ . Initially, it examines if the summation of the two sub-frames (as specified in accordance with Eqs (1) and (2)) exceeds or is less than the frame size. In the former case, it adds the half difference to each direction, while in the latter case it reduces the half difference from each direction in order for the final allocation to yield the required frame size. Additionally, the algorithm checks whether the difference is an odd number, in which case it adds the unity to the direction with the most unserved slots considering

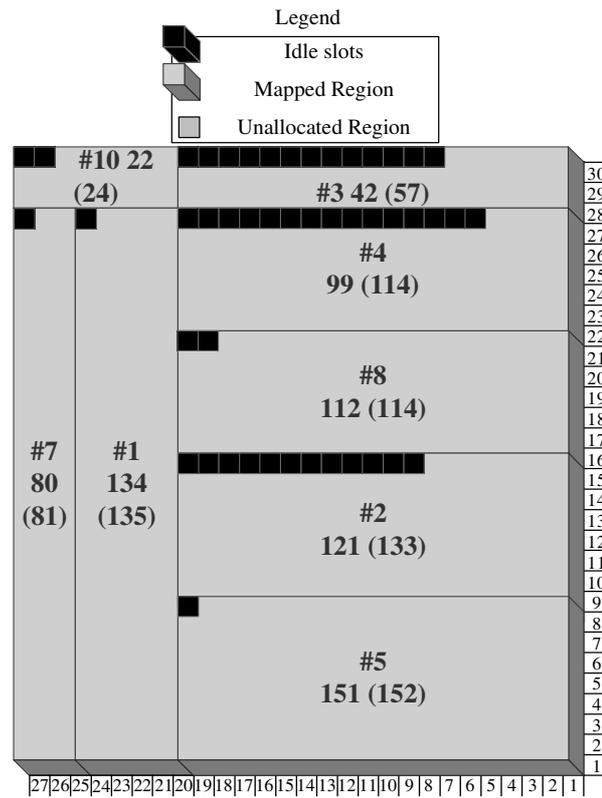


Fig. 2. The final form of the allocation bin.

the former case or subtracts the unity from the direction with the less unserved slots considering the latter case.

#### 4.4. Numerical example

RLAM, in the context of each frame, re-determines the length size of each sub-frame for the forthcoming frame. To be more specific, for each frame, the most appropriate downlink and uplink sub-frame sizes are defined by selecting one of the states that correspond to the available downlink and uplink width values (i.e., assuming frame’s length equal to 10 ms, the available states are 13). The transition to the most appropriate state for the forthcoming frame is decided by the learner module on the basis of a reward/penalty evaluative signal (reinforcement signal) produced by the learner, which takes into account the feedback returned from the environment concerning the output of the mapping operation for the current frame. Considering that our aim is to accommodate all downlink/uplink requests utilizing as much as possible the allocation space given, feedback comprises both the portion of the wasted bandwidth (i.e., unused slots) as well as the amount of bandwidth that fails to be served due to lack of physical resources (i.e., unserved slots) in the context of the current frame, concerning both the uplink as well as the downlink sub-frames, for the currently selected downlink and uplink sub-frame sizes.

In order to help the reader better comprehend all aspects of RLAM operation, in the following we provide an illustrative example of the application of the proposed scheme. Due to space limitations, we

restrict our attention to the application of RLAM to the downlink sub-frame; RLAM is applied in the same manner to the uplink sub-frame.

Let us suppose that ten MSs are connected to the BS. The scheduler informs the mapper about the allocation requests of each MS. Let us consider frame duration equal to 10 ms. In the context of frame  $k$ , Table 2 shows ten randomly chosen requests to be accommodated into an  $30 \times 27$  downlink sub-frame (810 slots), where the downlink bin's height is 30 and the width is 27. Thus, for the frame  $k$  it stands that the downlink sub-frame's size is 27 columns and the uplink sub-frame's size is 15 columns:  $DSFS_k = 27$  and  $USFS_k = 15$ .

The final downlink allocation bin's form is depicted in Fig. 2. AHBM algorithm [15] has been adopted for the accommodation of requests in the downlink sub-frame. It should be noted that the standard uplink mapping scheme is considered regarding the uplink sub-frame.

As a first step, the reward/penalty reinforcement signal should be estimated. According to Eq. (5), the reward signal is a function of two important parameters: the number of unserved slots and the number of unused slots that indicate the conditions of the downlink traffic flow; more unserved slots indicate more bandwidth needs, while the larger the number of unused slots is bandwidth wastage occurs. As it may be observed from Fig. 2, the final downlink allocation results in 47 unused (idle) slots (i.e., 1 slot due to the accommodation of MS<sub>5</sub>, 12 slots due to the accommodation of MS<sub>2</sub>, 2 slots due to the accommodation of MS<sub>8</sub>, 15 slots as a result of the accommodation of MS<sub>4</sub>, 13 slots as a result of the accommodation of MS<sub>3</sub>, 1 slot due to the accommodation of MS<sub>7</sub>, 1 slot due to the accommodation of MS<sub>1</sub>, 2 slots due to the accommodation of MS<sub>10</sub>), while the requests of MS<sub>6</sub> and MS<sub>9</sub> could not be mapped in the current downlink sub-frame due to the lack of physical resources, resulting, thus, in 48 unserved slots. Thus,  $UnUS_k = 47$  and  $UnSS_k = 48$ . At this point it should be noted that all unused slots in the final downlink allocation's bin result due to inefficiencies of the mapping process, attributed to the OFDMA rectangular shaping restriction that the standard applies to the downlink sub-frame. In the specific example considered, there are no unused slots due to unallocated region.

In accordance with Eqs (11) and (13), respectively, we estimate the reward factor  $rr_k(UnUS_k)$  concerning the unused slots and the reward factor  $rr_k(UnSS_k)$  concerning the unserved slots. Specifically,  $rr_k(UnUS_k) = -47/810 = -0.058$  and  $rr_k(UnSS_k) = 48/810 = 0.059$ .  $\theta$  parameter is taken equal to 1. Concerning the determination of the weighting factors  $w_{unused}(UnUS_k)$  and  $w_{unserved}(UnSS_k)$  that provide the relative significance of the two reward factors  $rr_k(UnUS_k)$  and  $rr_k(UnSS_k)$ , respectively, we first estimate the utilization factor  $uf$  Eq. (15). Hence:

$$uf = \frac{TotalRequestedSlots}{TotalBindedSlots} = \frac{763}{810} = 0,942$$

Then, in accordance with Eqs (16) and (17), respectively, the weighting factors  $w_{unused}$  and  $w_{unserved}$  are estimated. Specifically,  $w_{unused} = 0,4855$  and  $w_{unserved} = 0,5145$ . Applying now Eq. (5) we estimate reward factor  $rr_k(UnUS_k, UnSS_k) = 0,0022$ .

Finally, in order to determine the downlink sub-frame's size for the forthcoming frame  $k + 1$ , we apply Eq. (1), where:

$$NSFS_k = \frac{(SFS_k - SFS_{min})}{(SFS_{max} - SFS_{min})} = \frac{(27 - 21)}{(33 - 21)} = 0,5$$

$$m_r = 0,9, l(NSFS_k) = 0,378.$$

Thus,  $NSFS_k = 0,501$  and  $DSFS_k = 27$ . As it may be observed, in the current example, RLAM maintains the same size for the downlink sub-frame of the forthcoming frame  $k + 1$  (i.e., 27 symbols).

This is attributed to the fact that only small difference exist between the two parameters that drive the operation of RLAM: the number of unserved and the number of un-used slots.

RLAM following exactly the same procedure estimates the size of the uplink sub-frame for the forthcoming frame  $k + 1$ . As a final step, RLAM checks whether the summation of the downlink and uplink sub-frame width values (rule that is imposed by the standard) is equal to the total frame width (i.e.,  $FS = DSFS_k + USFS_k = 42$ ). In case this rule is not satisfied, the re-adjustment algorithm is applied, which for fairness reasons it adds/subtracts half the difference to each direction in order for the final allocation to acquire the required frame size. In case the difference is odd number, the direction with the most unserved slots is favored.

## 5. Evaluation results

In this section, our aim is to provide evidence on the efficiency of the proposed adaptive scheme. To this respect, we conducted extensive simulation experiments considering real WiMAX multimedia traffic traces in order to evaluate its performance. In the following, as a first step, we describe in a detailed manner the simulation environment built for evaluation purposes. In the sequel, the performance of the proposed RLAM scheme on acquiring the most appropriate downlink-to-uplink width ratio is estimated. Additionally, RLAM's performance with respect to the downlink and uplink sub-frames' size configuration is compared with an automata-enhanced scheme presented in [9]. Finally, we study the RLAM's effectiveness in comparison to static schemes which operate by setting a predefined and fixed width ratio.

### 5.1. Evaluation environment

The performance evaluation experiments have been conducted in a simulation environment that has been designed and implemented in Matlab. The implemented scenario, i.e., the WiMAX testbed where the packet traces were captured, includes a Mobile WiMAX access network comprising a BS and multiple MSs. The physical characteristics of the testbed wireless network are in accordance with the IEEE 802.16e network specifications. The well-known partially used sub-channelization (PUSC) mode (the most common frequency diversity mode for practical mobile communications environments) is adopted. According to this technique, 30 different channels are associated to the downlink and the uplink sub-frames. Moreover, the duration of the whole frame is set equal to 10 ms, which correspond to 45 OFDMA symbols. Three symbols are reserved for control purposes (one symbol for Preamble, and two symbols for the MAP and FCH fields) and are excluded from the available for allocation purposes bandwidth slots.

The WiMAX BS was configured to support a 10 ms frame length. Considering that the downlink-to-uplink width ratio may vary from 1:1 to 3:1, for a frame of 10 ms, it stands that 21 to 33 symbols are assigned to the downlink sub-frame and 21 to 9 symbols are destined to the uplink sub-frame, considering a 1:1 and a 3:1 downlink-to-uplink width ratio, respectively. Accordingly, the downlink sub-frame defines a rectangular allocation space of 630 ( $30 \times 21$ ) slots, under 1:1 ratio, to 990 ( $30 \times 33$ ) slots, under 3:1 ratio, while the uplink sub-frame defines a rectangular allocation space of 630 ( $30 \times 21$ ) slots, under 1:1 ratio, to 270 ( $30 \times 9$ ) slots, under 3:1 ratio, since the allocation bin is constructed by 30 channels. In Fig. 3, the learner – state – action model is graphically illustrated, for the considered simulation environment.

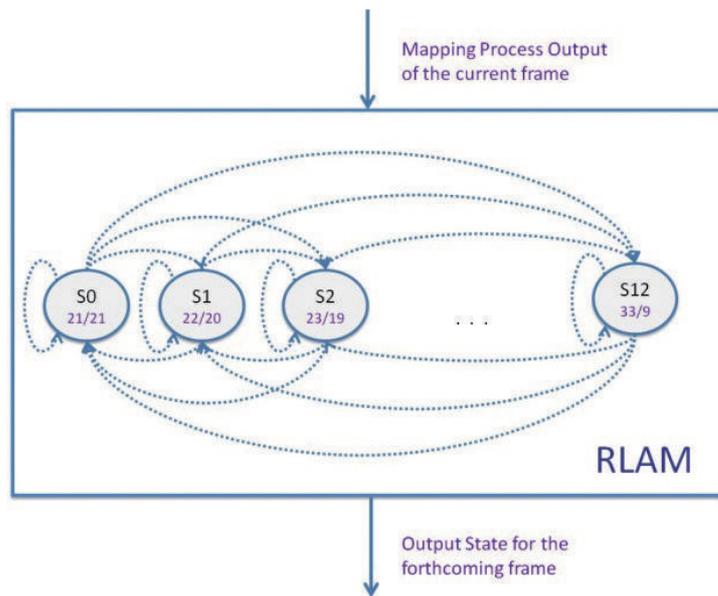


Fig. 3. The learner-state-action model.

Real WiMAX traffic traces, captured from both downlink and uplink data flows, were obtained and provided as input to our custom-made simulator. An aggregate traffic of 10 minutes was captured, using the Wireshark capture tool [31]. Specifically, four different traffic sessions were captured, with each session generating a unique packet stream. The first stream belongs to a Real Media Streaming application, running over multiple MSs via a single BS, using the transmission control protocol (TCP). The session lasted for 3 minutes. Afterwards, a file upload stream from a MS to the BS with 2 minutes duration took place using the file transfer protocol (FTP). The third stream represents a VoIP session between the BS and a MS via the user datagram protocol (UDP) and lasts for 3 minutes. Lastly, the well-known peer-to-peer software utorrent [32] was used to capture a file download from the BS to a connected MS. The stream duration was 2 minutes. Hence, a total connection time of 10 minutes was realized, whereas four different services and protocols were considered.

Regarding the streaming application, the downstream row generated an average traffic of 0.038 Mbit/sec per MS and the average packet size was 1372 Bytes, while the upstream direction produced an average traffic of 0.001 Mbit/sec per MS and the average packet size was 68 Bytes. The FTP session lasted for 2 minutes, it generated an average upstream (downstream) traffic of 10 Mbit/sec (0.7 Mbit/sec), and the average upstream (downstream) packet size was 1380 Bytes (67 Bytes). Concerning the VoIP application, each participating MS (upstream direction) produced an average traffic of 0.068 Mbit/sec and the average packet size was equal to 162 Bytes. The downstream direction, generated by the BS, produced an average traffic of 0.047 Mbit/sec, while the average packet size was equal to 126 Bytes. Finally, the utorrent download session lasted for 2 minutes, it generated an average downstream (upstream) traffic of 8 Mbit/sec (0.5 Mbit/sec), and the average downstream (upstream) packet size was 1360 Bytes (67 Bytes).

It is considered that a request (per frame) length may vary from 1 to 267 slots, which approximately corresponds to 1600 Bytes under QPSK-1/2 status (minimum operational channel conditions), regarding the downlink and the uplink direction. The queue length for each direction was set equal to 100 MB. The BS comprises a logical queue, which stores the traffic requests generated by each MS.

Table 3  
Simulation assumptions

| Channel mode                | PUSC                                |
|-----------------------------|-------------------------------------|
| Frame Length                | 10 ms                               |
| Preamble Size               | 1 Symbol                            |
| MAP, FCH Sizes              | 2 Symbols                           |
| Downlink sub-frame Symbols  | 21 to 33 (1:1 to 3:1 ratio)         |
| Uplink sub-frame Symbols    | 21 to 9 (1:1 to 3:1 ratio)          |
| Downlink sub-frame capacity | 630 to 990 slots (1:1 to 3:1 ratio) |
| Uplink sub-frame capacity   | 630 to 270 slots (1:1 to 3:1 ratio) |
| Downlink Queue Length       | 100 MB                              |
| Uplink Queue Length         | 100 MB                              |

Table 4  
Wireless channel parameters

| Probability | Modulation and coding | Bits per slot |
|-------------|-----------------------|---------------|
| 0.05        | Outage                | 0             |
| 0.15        | QPSK-1/2              | 48            |
| 0.2         | QPSK-3/4              | 72            |
| 0.3         | 16QAM-1/2             | 96            |
| 0.3         | 16QAM-3/4             | 144           |

Bandwidth (in bytes) is associated with OFDMA slots in accordance with the selected modulation and coding schemes. WiMAX supports adaptive modulation and coding, taking into account current radio channel conditions. Specifically, the BS determines the most appropriate modulation and coding schemes to be applied based on the ranging signal power, which is indicative of the current channel conditions experienced by the MS. In the following, for each frame, the MSs' requests in bytes are associated with slots based on specific probabilities [33,34]. For example, it is assumed that a MS receives QPSK-1/2 channel status with probability equal to 0.15. In this way, since each request for the downlink or the uplink direction is randomized according to the specific probabilities, mobility is taken into account in our scenarios, with the MSs continuously changing their position with respect to the BS.

Additionally, on stimulating the realism of the simulation environment, we consider that a MS may experience outage if its channel condition is too bad, with probability equal to 0.05. Table 3 depicts the main simulation assumptions, Table 4 shows the adopted wireless channel parameters, and Table 5 summarizes the traffic captured and service details.

For evaluation purposes the following schemes have been implemented: A) the proposed adaptive RLAM scheme, B) the Learning – Automata based scheme presented in [9], C) the Fixed 1:1 scheme, keeping the downlink-to-uplink ratio static to 1:1, D) the Fixed 2:1 scheme, maintaining the downlink-to-uplink ratio equal to 2:1, and E) the Fixed 3:1 scheme, defining a static 3:1 ratio. For all implemented schemes, both the downlink and uplink directions are considered, while the AHBM algorithm [15] is adopted as the mapping algorithm concerning the downlink sub-frame and the standard uplink mapping scheme is considered regarding the uplink sub-frame, as described in Section 2.

## 5.2. RLAM performance evaluation

In this sub-section, extensive experiments are performed in order to provide evidence on the performance of the proposed scheme. The performance metric considered is the average error rate with respect to the optimal size of the downlink and uplink sub-frames in the context of each frame, which is defined

Table 5  
Traffic and service details

| Incorporated service | Duration | Rate  | Average packet size                               |
|----------------------|----------|---|---|
| Real media Streaming | 3 sec    | 0.038 Mbit/sec in Downstream and 0.001 Mbit/sec in Upstream | 1372 Bytes in Downstream and 68 Bytes in Upstream |
| FTP uploading        | 2 sec    | 0.7 Mbit/sec in Downstream and 10 Mbit/sec in Upstream      | 67 Bytes in Downstream and 1380 Bytes in Upstream |
| VoIP                 | 3 sec    | 0.047 Mbit/sec in Downstream and 0.068 Mbit/sec in Upstream | 162 Bytes in Downstream and 126 Bytes in Upstream |
| uTorrent Downloading | 2 sec    | 8 Mbit/sec in Downstream and 0.5 Mbit/sec in Upstream       | 1360 Bytes in Downstream and 67 Bytes in Upstream |

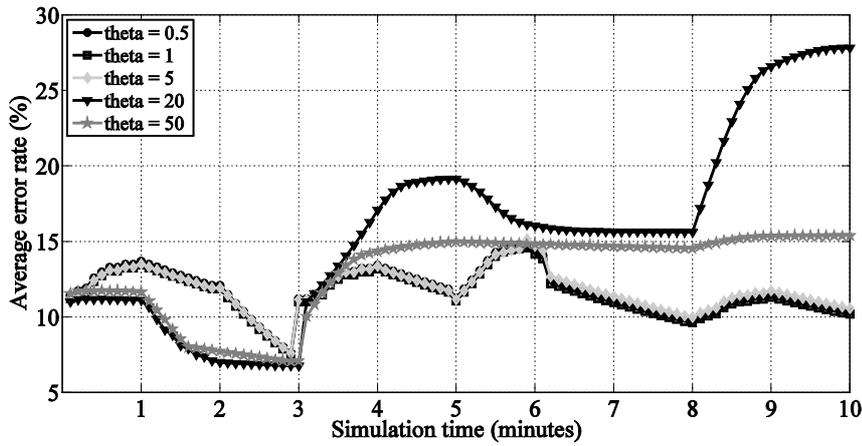


Fig. 4. The impact of the  $\theta$  (theta) parameter to the downlink sub-frame width determination

as follows.

$$average\_error\_rate = \frac{\sum_{i=1}^{Total\_Frame} \frac{|SFS_i - Optimal_i|}{SFS_{max} - SFS_{min}}}{Total\_Frames}$$

where the parameter  $Total\_Frames$  represents the total number of simulated frames and the parameter  $Optimal_i$  expresses the optimal downlink/uplink width for frame  $i$ , where  $1 \leq i \leq Total\_Frames$ . The optimal value,  $Optimal_i$ , is defined upon the completion of the mapping process concerning the frame  $i$ , knowing at that time the bandwidth demands for both directions. Since the simulation experiments last for 10 min and the frame duration is equal to 10 ms, it may be easily concluded that the parameter  $Total\_Frames$  is equal to 60000 frames.

The first set of experiments aims at studying the impact of parameter  $\theta$  (theta) to the RLAM’s performance. To this respect, different values are considered for the parameter  $\theta$  (i.e.,  $\theta = \{0.5, 1.5, 20, 50\}$ ), while the respective results are illustrated in Fig. 4. In this experiment, the memory factor  $m$  is fixed, equal to 0.9 and the number of connected MSs is considered static and equal to 10.

As it may be observed from Fig. 4,  $\theta$  values larger than 2 negatively influence the RLAM’s performance. Specifically, considering  $\theta = 5, 20, and 50$  results in sub-frames’ width estimations that include a significant error element. This is attributed to the fact that when a large  $\theta$  value is assumed, the model is less sensitive to the influence of the  $UnUS_{k-1}$  and  $UnSS_{k-1}$  parameters, as minor modifications to

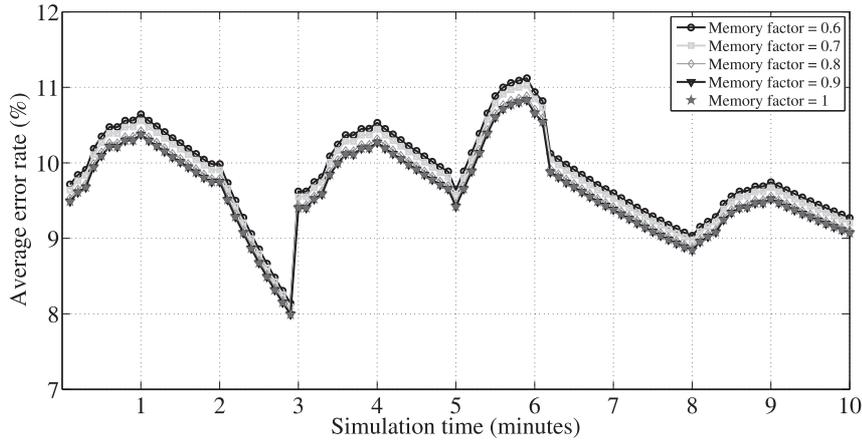


Fig. 5. The impact of the memory factor to the downlink sub-frame width determination.

their values drive the reward parameter towards quickly approaching its reservation value (i.e.,  $rr$  tends to zero). To be more specific, large  $\theta$  values decrease the  $rr$  in case it is larger than zero (i.e.,  $rr_2$  is larger than  $rr_1$ ), and it increase the  $rr$  in case it is negative (i.e.,  $rr_1$  is larger than  $rr_2$ ). This ‘senseless’ behavior produces estimations not accurate enough and degrades the network performance. On the other hand, the linear selection of the  $\theta$  parameter ( $\theta = 1$ ) seems to be the most appropriate one, achieving better performance than  $\theta = 0.5$  and  $\theta = 2$ , because the  $rr_1$  and the  $rr_2$  parameters are linearly (i.e., smoothly) varied offering more accurate width estimations, adequately following traffic variations.

The second set of experiments (Fig. 5) studies the impact of the memory factor  $m_r$ . Five different values are considered (i.e.,  $m_r = \{0.6, 0.7, 0.8, 0.9, 1\}$ ). Since the memory factor expresses the significance attributed to the latest downlink-to-uplink width adjustment, as the parameter  $m_r$  decreases, the memory of the system is strengthened. In such a case, due to the small value of  $m_r$ , the algorithm takes into consideration the previous sub-frames’ width value to a greater extend. From Fig. 5, it may be observed that the adaptive algorithm seems to be capable of instantly adapting to the networks dynamics, independently of the considered value of the memory factor, since all variations present almost the same performance. However, the values of 0.9 and 1 seem to be the most appropriate ones for the determination of the memory factor. At this point it should be noted that due to space limitations the results depicted in Figs 4 and 5 concern the downlink sub-frame. The uplink sub-frame related results are easily inferred since the two sub-frame sizes are complement and are readily available.

From the obtained results, we reach to the following conclusion. The proposed RLAM scheme is quite accurate presenting an error rate below 15%. The error level seems to increase upon changing the bandwidth demands of the two directions, since the model necessitates some time to adapt to the new traffic situation, which is the case for the time periods between the 3<sup>rd</sup> and the 4<sup>th</sup> minutes, between the 5<sup>th</sup> and the 6<sup>th</sup> minutes, and between the 8<sup>th</sup> and the 9<sup>th</sup> minutes. Upon adapting to the traffic change (e.g., between the second and the third minutes), the model presents a quite accurate behavior, reducing the mean error rate up to 5%.

Consequently, the model’s error rate occurring only under extreme traffic conditions may be considered acceptable.

### 5.3. Learning performance

In this section, we compare the RLAM’s performance against the learning-automata based model presented in [9]. The learning automaton operation involves the following tuning parameters. First, the  $L$

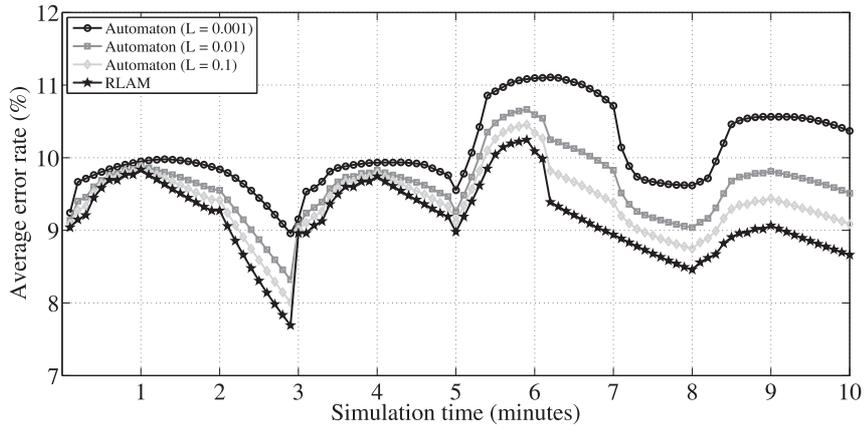


Fig. 6. Performance comparison of RLAM and various Automaton versions in terms of average error rate.

parameter, which governs the convergence speed of the learning process, and second, the  $\alpha$  parameter, which prevents the state probabilities from receiving the zero value. We thoroughly compare the performance of the two schemes by implementing and experimenting on three versions of the automaton. The first version assumes  $L = 0.001$ , the second one considers  $L = 0.01$ , and the third version sets  $L = 0.1$ . The parameter  $\alpha$  is stable and taken equal to  $10^{-4}$  for all considered versions. Our model has been evaluated assuming  $\theta = 1$ , and  $m_r = 0.9$ . Furthermore, the number of connected MSs is taken equal to 10 and the simulation time is 10 min (i.e., 60000 frames).

Figure 6 graphically illustrates the obtained average error rate with respect to the optimal sub-frames width value for both schemes. Initially, and during the first three minutes, the error rate remains low, as both models provide a good solution in a quite straightforward manner; the optimal downlink-to-uplink width ratio is obvious and equal to 2:1 (27 symbols for downlink, opposed to 15 symbols for uplink direction), since the downlink flow is much more demanding than the uplink. After the 3<sup>rd</sup> minute, the error rate increases. This is attributed to the fact that the ftp file upload service demands much more bandwidth for the uplink direction. As a consequence, the learning process necessitates some time to reach to the most appropriate decision and adapt to the current traffic conditions. The next error rate fluctuation happens after the 5<sup>th</sup> minute, where the VoIP application initiates. Here, the optimal ratio is the 1:1, granting 21 symbols for each direction. Both techniques, i.e., the proposed RLAM and the automaton, suffers from a temporary error rate increase, however the proposed model keeps it at minimum levels. Once, the traffic relation tilts in favor of the downlink flow, i.e., upon the beginning of the 8th minute, the error rate decreases; the optimal ratio becomes now 3:1, i.e., 33 symbols for the downlink sub-frame and 9 for the uplink one and the most appropriate sub-frames configuration is easily acquired. From the results obtained considering all schemes, it may easily be observed that the RLAM exceeds the automaton in estimation accuracy, since it presents reduced average error rate compared to all versions of the automaton scheme, i.e., regardless of the parameter  $L$ . The superiority of the RLAM scheme lies in its sophisticated adaptive nature; it takes into account sub-frames' configuration history values, attributing to them however different significance in accordance with the memory factor, and exploits the utilization factor, which expresses the number of unused slots that are needed for serving a specific number of unserved slots in average, in order to address the inefficiencies of the mapping process. Thus, the estimation procedure becomes more precise.

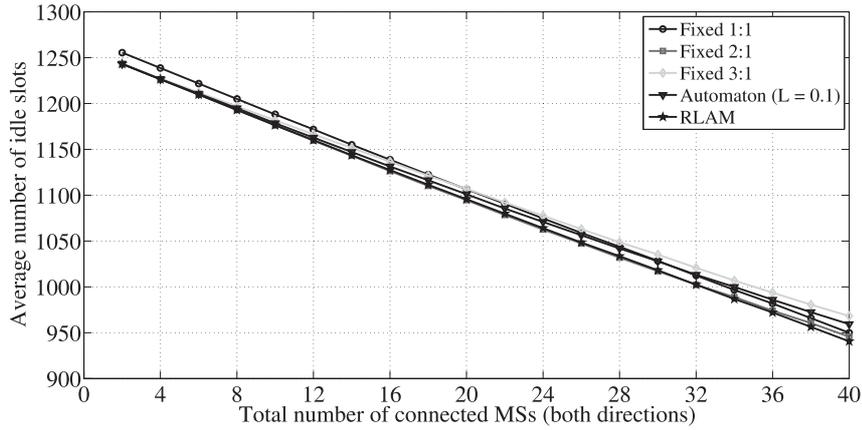


Fig. 7. Average number of idle slots per frame with respect to the total number of connected MSs (both directions).

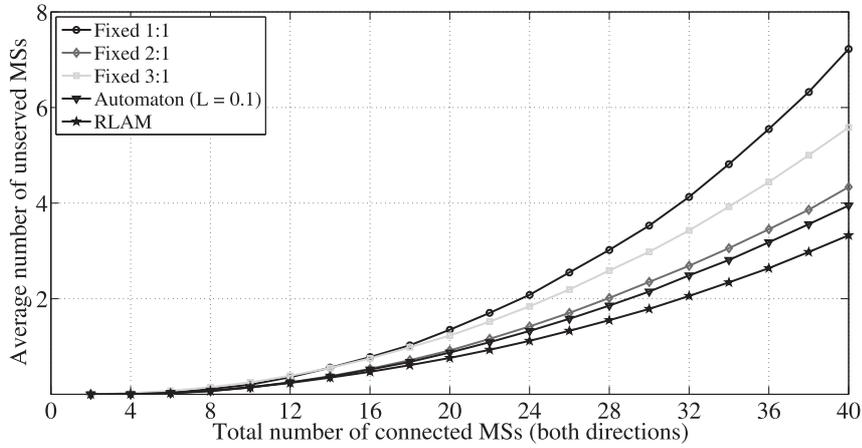


Fig. 8. Average number of unserved MSs per frame with respect to the total number of connected MSs (both directions).

#### 5.4. Network performance

In this sub-section, we study the effectiveness of the model designed compared against three schemes that keep static the downlink-to-uplink width sub-frame ratio to a predefined value. Furthermore, the performance of the scheme proposed is examined against the automaton efficiency assuming the most effective  $L$ . To this respect, five different schemes have been implemented and compared in the simulation experiments conducted: a) RLAM, the adaptive model proposed, which is capable of adjusting the ratio from 3:1 to 1:1 on a frame-by-frame basis, b) the Fixed 1:1, which keeps the ratio stable and equal to 1:1, c) the Fixed 2:1, which maintains a fixed ratio of 2:1, d) the Fixed 3:1, which defines a predefined and static ratio of 3:1, and e) the learning automaton considering  $L = 0.1$  and  $\alpha = 10^{-4}$ . Our model has been evaluated assuming  $\theta = 1$ , and  $m_r = 0.9$ .

Three main performance metrics have been considered: a) the average number of idle slots per frame, which indicates the utilization and efficiency of the mapping process, b) the average number of unserved MSs, which expresses the portion of MSs that fail to be accommodated concerning both the downlink and uplink sub-frames due to lack of resources, and c) the average number of unserved slots, which

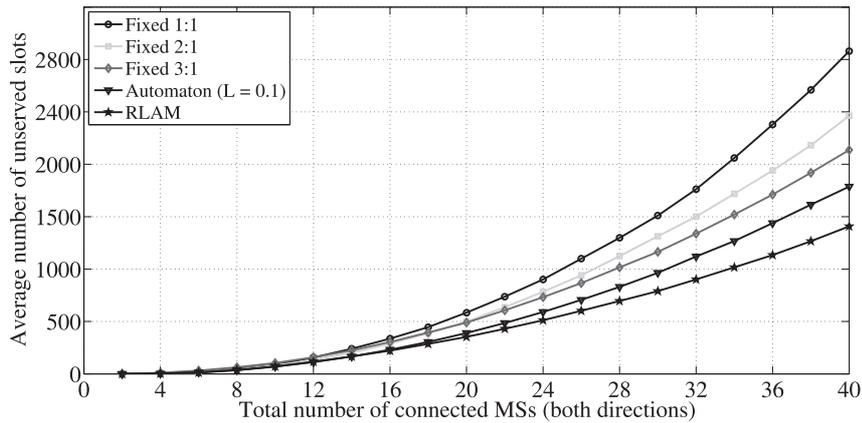


Fig. 9. Average number of unserved slots per frame with respect to the total number of connected MSs (both directions).

denotes the total number of slots that fail to find allocation space in both sub-frames due to lack of resources.

The results are illustrated in Figs 7, 8, and 9, which depict the average number of idle slots, the average number of unserved MSs, and the average number of unserved slots, respectively. From the obtained results, it may easily be observed that our model is capable of a) keeping the portion of the wasted (frame) allocation space low, in terms of slots, b) reducing the mean number of the MSs that fail to find accommodation space per frame, and c) increasing the service ratio of the system by reducing the portion of bandwidth requests that return to the scheduler. To be more specific, as depicted in Fig. 7, our model presents the lowest mean number of idle slots compared to all other fixed schemes up to ten connected MSs, by effectively exploiting the available bandwidth, even though the differences between the various algorithms are marginal. This is attributed to our scheme’s efficiency concerning the determination of the most appropriate downlink-to-uplink ratio and, thus, to the efficient distribution of the available OFDMA columns to the direction that needs them more. However, when the number of the connected MSs increases, the RLAM sacrifices bandwidth to efficiently support the subscribers’ requests. In other words, more bandwidth is assigned to the downlink direction where the shaping rule applies, in order to increase the number of served MSs per frame, a performance metric that is the most important in a wireless access network. This argument is verified by the results obtained in Fig. 8. It is clearly depicted that the model proposed achieves a reduced number of unserved MSs as the number of the connected MSs increases. Hence, it provides improved service ratio compared to the considered fixed schemes, which sacrifice the service capability to save bandwidth. Similar results are graphically illustrated in Fig. 9; Adopting RLAM, more bandwidth requests in terms of slots are successfully allocated per frame compared to the fixed schemes. As expected, the novel adaptive model allows for a more flexible bandwidth management, offering multiple benefits to the mobile WiMAX network. Contrary to the static schemes that maintain a predetermined and fixed downlink-to-uplink sub-frame width ratio, the adaptive model is able to adjust the ratio in accordance with the traffic dynamics. Moreover, the learning mechanism of the RLAM presents notable improvements in the network performance compared to the automaton’s learning procedure, fact that indicates the superiority of the scheme proposed concerning the enhanced learning framework.

In a nutshell, the obtained simulation results indicate the superiority of the adaptive scheme proposed under realistic conditions including real WiMAX traces and multiple services; it is the most efficient one with respect to estimation accuracy, network resources utilization, and bandwidth management

compared against the considered static schemes. Finally, RLAM succeeds in considerably reducing the portion of the requests that are returned back to the scheduler in terms of slots, increasing, thus, the network service rate.

## 6. Conclusions

In this paper, RLAM, a novel adaptive model has been presented, efficiently determining the downlink-to-uplink width ratio on a frame by frame basis, following in quite accurate manner the traffic dynamics in mobile WiMAX access networks. The model exploits a rigorous reinforcement learning based model in order to effectively estimate the most appropriate sub-frames size configuration in the context of each frame. The learning module designed takes into account past width values and adjusts the forthcoming downlink and uplink sub-frame width, considering the portion of the wasted bandwidth as well as the amount of the bandwidth requests that fail to be served due to lack of physical resources. Extensive simulation results show that the model proposed offers an accurate learning framework, keeping the average error rate below 15%. Additionally, it exhibits improved performance with respect to other learning schemes, e.g., learning automata, and it provides significant benefits in the wireless network performance in terms of service ratio and resources utilization compared to fixed mapping schemes, which maintain a predefined and fixed ratio. Our future plans involve the experimentation of the proposed scheme in LTE wireless networks.

## References

- [1] M. Katz and F. Fitzek, *WiMAX Evolution: Emerging Technologies and Applications*, Wiley Publishing, 2009.
- [2] J. Andrews, A. Ghosh and R. Muhamed, *Fundamentals of WiMAX, Understanding Broadband Wireless Networking*, Prentice Hall, 2007.
- [3] D. Pareit, M. Deruyck, E. Tanghe, W. Joseph, I. Moerman, L. Martens and P. Demeester, Detailed Modeling of MAC Throughput and Ranges for Mobile WiMAX, *IEEE Communications Letters* **15** (2011), 839–841.
- [4] I. Ahmad and D. Habibi, Improving throughput in WiMAX communication at vehicular speeds, *Journal of High Speed Networks* **18** (2012), 157–171, IOS Press.
- [5] T.-H. Lee, C.-H. Liu, A.S.G. Campbell and Y.-W. Kuo, A Data Mapping Algorithm for Two-Level Requests in WiMAX Systems, *In Proc. of the 75th IEEE Vehicular Technology Conference (VTC Spring)*, (2012), 1–5.
- [6] O. Sharon, Y. Alpert and G. Tabajah, Efficient coupled PHY and MAC use of physical Bursts in WiMAX/IEEE 802.16e networks, *Elsevier, Physical Communication*, Available online 26 September 2012.
- [7] U. Farooq, S. Majumdar and E. Parsons, High performance publish/subscribe middleware for mobile wireless networks, *Mobile Information Systems* **3** (2007), 107–132, IOS Press.
- [8] T.-M. Bohnert, M. Castrucci, N. Ciulli, G. Landi, I. Marchetti, C. Nardini, B. Sousa, P. Neves and P. Simoes, QoS management and control for an all-IP WiMAX network architecture: Design, implementation and evaluation, *Mobile Information Systems* **4** (2008), 253–271, IOS Press.
- [9] A. Sarigiannidis, P. Nikipolitis, G. Papadimitriou, P. Sarigiannidis and M. Louta, Using Learning Automata for Adaptively Adjusting the Downlink-Uplink Ratio in IEEE 802.16e Wireless Networks, *In Proc. of the 2011 IEEE Symposium on Computers and Communications (ISCC'11)*, (2011), 353–358.
- [10] Y. Ben-Shimol, I. Kitroser and Y. Dinitz, Two-Dimensional Mapping for Wireless OFDMA Systems, *IEEE Transactions on Broadcasting* **52** (2006), 388–396.
- [11] C. Desset, E.B. de Lima Filho and G. Lenoir, WiMAX Downlink OFDMA Burst Placement for Optimized Receiver Duty-Cycling, *In Proc. of the 2007 IEEE International Conference on Communications (ICC 2007)*, (2007), 5149–5154.
- [12] T. Ohseki, M. Morita and T. Inoue, Burst Construction and Packet Mapping Scheme for OFDMA Downlinks in IEEE 802.16 Systems, *In Proc. of the 2007 IEEE Global Telecommunications Conference (GLOBECOM 2007)*, (2007), 4307–4311.
- [13] T. Wang, H. Feng and B. Hu, Two-Dimensional Resource Allocation for OFDMA System, *In Proc. of the IEEE International Conference on Communications Workshops*, (2008), 1–5.

- [14] C. So-In, R. Jain and A. Al-Tamimi, eOCSA: An Algorithm for Burst Mapping with Strict QoS requirements in IEEE 802.16e Mobile WiMAX Networks, *In Proc. of the 2nd IFIP Wireless Days (WD 2009)*, (2009), 1–5.
- [15] P.G. Sarigiannidis, G.I. Papadimitriou, P. Nicopolitidis, M.S. Obaidat and A. Pomportsis, A Novel Adaptive Mapping Scheme for IEEE 802.16 Mobile Downlink Framing, *In Proc. of the 2010 IEEE GLOBECOM Conference (GLOBECOM 2010)*, (2010), 1–5.
- [16] S. Misra, B. Banerjee and B.E. Wolfinger, A learning automata-based uplink scheduler for supporting real-time multimedia interactive traffic in IEEE 802.16 WiMAX networks, *Elsevier, Computer Communications* **35** (2012), 1871–1881.
- [17] Z.-H. Liu and J.-C. Chen, Design and Analysis of the Gateway Relocation and Admission Control Algorithm in Mobile WiMAX Networks, *IEEE Transactions on Mobile Computing* **11** (2012), 5–18.
- [18] J. Vanderpypen and L. Schumacher, Treemap-Based Burst Mapping Algorithm for Downlink Mobile WiMAX Systems, *In Proc. of the 2011 IEEE Vehicular Technology Conference (VTC 2011 Fall)*, (2011), 1–5.
- [19] S. Misra, P.V. Krishna and V. Saritha, An Efficient Approach for Distributed Channel Allocation With Learning Automata-Based Reservation in Cellular Networks, *SIMULATION: Transactions of the Modeling and Simulation International* **88** (2012), 1166–1179.
- [20] G. Naddafzadeh-Shirazi, P.-Y. Kong and C.-K. Tham, Distributed Reinforcement Learning Frameworks for Cooperative Retransmission in Wireless Networks, *IEEE Transactions on Vehicular Technology* **59** (2010), 4157–4162.
- [21] P. Nicopolitidis, K. Christidis, G.I. Papadimitriou, P.G. Sarigiannidis and A.S. Pomportsis, Performance evaluation of acoustic underwater data broadcasting exploiting the bandwidth-distance relationship, *Mobile Information Systems* **7** (2011), 285–298, IOS Press.
- [22] K.-L. Alvin Yau, P. Komisarczuk and P.D. Teal, Reinforcement learning for context awareness and intelligence in wireless networks: Review, new features and open issues, *Elsevier, Journal of Network and Computer Applications* **35** (2012), 253–267.
- [23] H. Tabrizi, G. Farhadi and J. Cioffi, A Learning-Based Network Selection Method in Heterogeneous Wireless Systems, *In Proc. of the 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, (2011), pp. 1–5.
- [24] N. Mastrorade and M. van der Schaar, Fast Reinforcement Learning for Energy-Efficient Wireless Communication, *IEEE Transactions on Signal Processing* **59** (2011), 6262–6266.
- [25] S. Misra, T.I. Ghosh and M.S. Obaidat, Routing Bandwidth Guaranteed Paths for Traffic Engineering in WiMAX Mesh Networks, *International Journal of Communication Systems (Wiley)*, (2013), DOI: 10.1002/dac.2518.
- [26] S. Misra, P.V. Krishna and V. Saritha, LACAV: An Energy-Efficient Channel Assignment Mechanism for Vehicular Ad hoc Networks, *The Journal of Supercomputing (Springer)* **62** (2012), 1241–1262.
- [27] R.S. Sutton and A.G. Barto, Reinforcement Learning: An Introduction, *Cambridge University Press*, 1998.
- [28] IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, *IEEE Std. 802.16-2009*, 2009.
- [29] H. Raiffa, The Art and Science of Negotiation, *Harvard University Press*, Cambridge, USA, 1982.
- [30] D.G. Pruitt, Negotiation Behaviour, *Academic Press*, 1981.
- [31] <http://www.wireshark.org/>.
- [32] <http://www.utorrent.com/>.
- [33] S. Doirieux, B. Baynat, M. Maqbool and M. Coupechoux, An efficient analytical model for the dimensioning of WiMAX networks supporting multi-profile best effort traffic, *Elsevier, Computer Communications* **33** (2010), 1162–1179.
- [34] B. Baynat, G. Nogueira, M. Maqbool and M. Coupechoux, An Efficient Analytical Model for the Dimensioning of WiMAX Networks, *Lecture Notes in Computer Science* **5550** (2009), 512–534.

---

**Malamati D. Louta** is Assistant Professor at the Department of Informatics and Telecommunications Engineering, University of Western Macedonia, Greece. She holds M.Eng. (1997) and Ph.D. (2000) degrees in Electrical and Computer Engineering and M.B.A. degree (2004) from the National Technical University of Athens. Dr Louta was a member of the staff of Public Power Corporation (1999–2005), Assistant Professor at the Department of Business Administration, T.E.I. of Western Macedonia (2005–2008) and Lecturer at the Department of Informatics and Telematics of Harokopio University of Athens (2008–2010). Her research interests include telecommunication networks and advanced services engineering. She is the author of over 60 peer-reviewed publications. Dr Louta serves as an associate editor, as a technical program committee member and as a reviewer in a number of international conferences and journals. She is member of the IEEE, the ACM and the Technical Chamber of Greece.

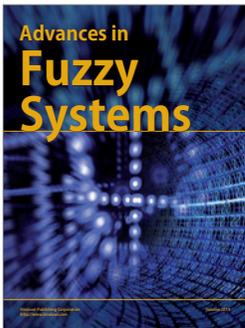
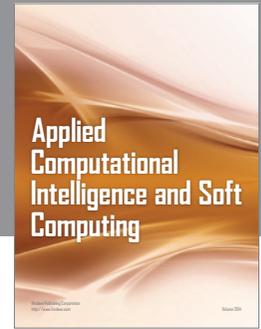
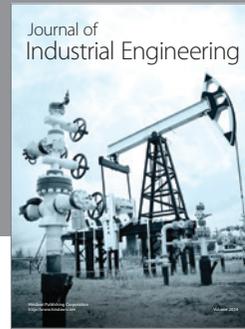
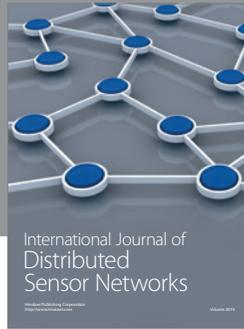
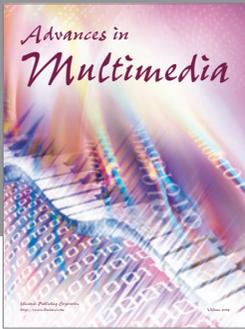
**Panagiotis G. Sarigiannidis** received the B.Sc. and Ph.D. degrees in computer science from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2001 and 2007, respectively. He is currently a Lecturer with the University of Western Macedonia, Kozani, Greece. His research interests include medium access protocols in optical networks, dynamic bandwidth allocation schemes in passive optical networks, scheduling policies in IEEE 802.16 wireless networks, wireless push systems design and optimization, quality of service provisioning in optical and wireless networks, traffic estimation and prediction via

numerical analysis, and design of burst allocation for optical burst switching networks. He has published over 45 papers in international journals, conferences, and book chapters.

**Dr. Sudip Misra** is an Assistant Professor in the School of Information Technology at the Indian Institute of Technology Kharagpur. Prior to this he was associated with Cornell University (USA), Yale University (USA), Nortel Networks (Canada) and the Government of Ontario (Canada). He received his Ph.D. degree in Computer Science from Carleton University, in Ottawa, Canada, and the masters and bachelors degrees respectively from the University of New Brunswick, Fredericton, Canada, and the Indian Institute of Technology, Kharagpur, India. His current research interests include algorithm design for emerging communication networks. Dr. Misra is the author of over 140 scholarly research papers. He has won six research paper awards in different conferences. Recently, he has been awarded the IEEE ComSoc Asia Pacific Outstanding Young Researcher Award at IEEE GLOBECOM 2012, Anaheim, California, USA. He was also the recipient of several academic awards and fellowships such as the Young Scientist Award (National Academy of Sciences, India), Young Systems Scientist Award (Systems Society of India), Young Engineers Award (Institution of Engineers, India), (Canadian) Governor General's Academic Gold Medal at Carleton University, the University Outstanding Graduate Student Award in the Doctoral level at Carleton University and the National Academy of Sciences, India – Swarna Jayanti Puraskar (Golden Jubilee Award). He was also awarded the Canadian Government's prestigious NSERC Post Doctoral Fellowship and the Humboldt Research Fellowship in Germany.

**Petros Nicopolitidis** (M'08) received the B.S. and Ph.D. degrees in computer science from the Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1998 and 2002, respectively. From 2004 to 2009, he was a Lecturer at the Department of Informatics, Aristotle University of Thessaloniki, where since 2009, he has been an Assistant Professor. He has published more than 60 papers in international refereed journals and conferences. He is coauthor of the book *Wireless Networks* (New York, NY: Wiley, 2003). His research interests are in the areas of wireless networks and mobile communications. Dr. Nicopolitidis has been an Associate Editor for the *International Journal of Communication Systems* since 2007.

**Prof. Papadimitriou Georgios** (M'89, SM'02) received the Diploma and Ph.D. degrees in Computer Engineering from the University of Patras, Greece in 1989 and 1994 respectively. In 1997 he joined the faculty of the Department of Informatics, Aristotle University of Thessaloniki, Greece, where he is currently serving as Professor. His main research interests include optical networks and wireless networks. Prof. Papadimitriou is Associate Editor of four IEEE journals. He is co-author of three international books published by Wiley. He is author or coauthor of more than 100 journal and 100 conference papers. He is a Senior Member of the IEEE.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

