Computer Networks 56 (2012) 244–259

Contents lists available at SciVerse ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

# Traffic-prediction-assisted dynamic bandwidth assignment for hybrid optical wireless networks

Maysam Mirahmadi, Abdallah Shami *

*Department of Electrical and Computer Engineering, The University of Western Ontario, Canada*

ARTICLE INFO

*Article history:*
Received 24 January 2011
Received in revised form 16 June 2011
Accepted 12 August 2011
Available online 16 September 2011

*Keywords:*
Access networks
Dynamic bandwidth allocation
Optical/wireless access networks

ABSTRACT

Hybrid optical-wireless networks provide the inexpensive broadband bandwidth, vital for modern applications, as well as mobility, and scalability required for an access network. However, in order to provide satisfactory Quality of Service (QoS) on such a non-homogeneous network, innovative designs are required.

This paper proposes a novel scheduling mechanism to significantly improve the delay guarantee, while maintaining high-level throughput, by predicting the incoming traffic to optical network units (ONU). The proposed scheduler managed to exploit the available information in hybrid optical-wireless networks, to enhance the ONU scheduler. This results in accurate prediction of incoming traffic, which leads to intelligent and traffic-aware, scheduling and dynamic bandwidth assignment (DBA).

Based on the proposed architecture, two DBA algorithms are proposed and their performance is evaluated by extensive simulations. Moreover, the maximum throughput of such network is analyzed. The results show that by using the proposed algorithms, the delay bound of delay-sensitive traffic classes can be decreased by a factor of two, without any adverse effect on the throughput.



## 1. Introduction

In recent years, numerous bandwidth consuming applications have emerged. These applications like Internet Protocol TeleVision (IPTV) and Video-on-demand (VoD), require considerable bandwidth and are attracting widespread attention among consumers. Moreover, service providers intend to offer voice, video, data and even wireless services to their customers based on an all-IP shared network infrastructure. As a result, a broadband access network capable of ensuring Quality of Service (QoS) for different service types is required.

Optical fiber networks are capable of transferring massive amounts of data, satisfactory for the everyday growing demand of recent applications. Modern passive optical networks, such as Gigabit Ethernet Passive Optical Network (GEPON) or Wavelength Division Multiplexed Passive Optical Network (WDM-PON) can transfer several gigabits of data per second and yet they have less maintenance cost in comparison to old optical networks, like Synchronous Optical Network (SONET). Nevertheless, it is still quite expensive to lay fiber to every user's premise and build a purely optical network. It is impractical to build a true optical network, such as Fiber To The Home (FTTH), in a congested and built-up urban area, where considerable number of broadband users are.

On the other hand, wireless networks are relatively inexpensive and have the unique feature of delivering data to mobile users. The drawbacks are that they offer much less bandwidth and they are error-prone. Another disadvantage of wireless networks is that wireless spectrum is shared among many users, further limiting the bandwidth offered to each user.

The integration of optical and wireless networks offers an inexpensive broadband solution. In a hybrid network,

* Corresponding author.
*E-mail addresses:* mmirahma@uwo.ca (M. Mirahmadi), ashami@eng.uwo.ca (A. Shami).

1389-1286/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2011.08.018

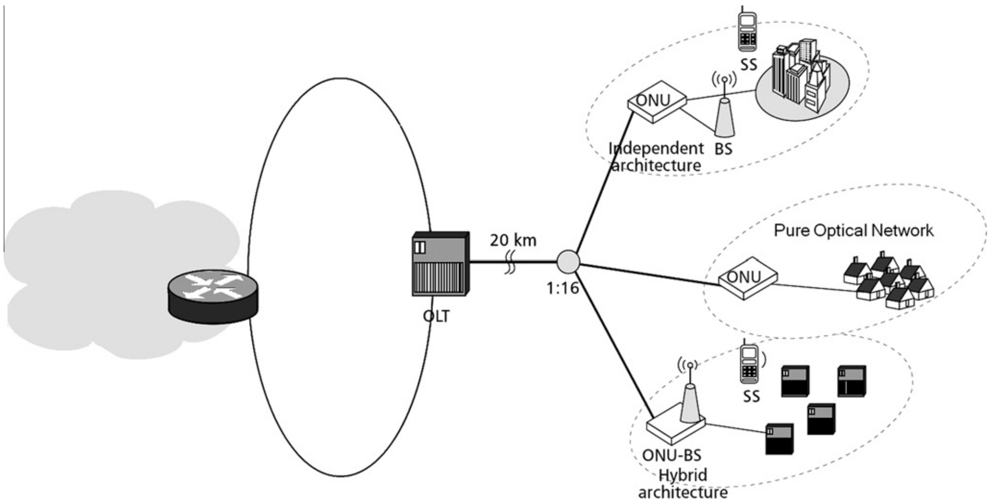



**Fig. 1.** Integrated EPON and WiMAX architectures [27] [moved accordingly].

such as the one shown in Fig. 1, the optical component is used to connect several wireless base stations as well as broadband customers to the central office. The abundant bandwidth of optical network is divided among the wireless base stations to serve their wireless users. In addition, hybrid optical-wireless networks offer several value-added features, such as redundancy, mobile access and cooperative communication which make it more appealing.

The hybrid optical-wireless networks seem to be the technology of future but some challenging issues still need to be addressed. Here, we focus on ensuring the QoS over a hybrid network, which is vital to offer a satisfactory service. The choice of how the scheduler assigns bandwidth to different flows has a tremendous impact on the QoS parameters. The scheduler typically relies on a dynamic bandwidth assignment (DBA) algorithm to divide the available bandwidth between different flows.

In this work, we propose a method to enhance the DBA algorithms in the optical part of hybrid optical-wireless networks. A prediction method, which uses the information in the wireless domain, helps decrease the delay bound of delay-sensitive flows, and hence results in improved QoS. The extensive simulation-based studies not only shed light on the factors, affecting the proposed algorithm, but also compare fixed and adaptive cycle length optical DBA algorithms side by side.

The content of this paper addresses the integration of Worldwide Interoperability for Microwave Access (WiMAX) and Ethernet Passive Optical Network (EPON) as successful technologies of their kind. However, most of the discussions can be applied to any hybrid optical-wireless network.

The paper is organized as follows. Section 2 introduces the system model. Section 3 compares advantages and disadvantages of integration in hybrid optical-wireless networks. The proposed algorithm is presented in Section 4 and the maximum throughput of optical-wireless network is analytically studied in Section 5. The traffic prediction assisted algorithms, used in simulation experiments, are explained in Section 6, while Section 7 defines the simulation platform. Section 8 analyzes the simulation results and finally Section 9 concludes the paper.

## 2. System model

The topology of hybrid optical-wireless network is commonly considered to be tree [2,5,14,16,27,30–32].

The root of the tree is the Optical Line Terminal (OLT) which is located at the central office (CO) and is directly connected to the core network. Each optical network unit (ONU), also called gateway [16], is connected to the OLT through one or several splitters as shown in Fig. 1. WiMAX Base Stations (BS) are connected to an ONU via a standard interface, like Ethernet, or even implemented in the same box. We refer to the latter case as ONU-BS. In practice, the WiMAX BSs share the optical bandwidth with other users, such as residential or corporate local area networks.

The WiMAX BSs provide wireless connectivity to their subscriber stations (SS) through either a single hop legacy wireless network or a relay-based wireless network with mesh topology. In the latter case a routing algorithm, like the one presented in [16], is necessary to effectively deliver the packets.

Due to the intrinsic property of passive optical networks, downstream data, transmitted from OLT to ONU, is broadcast to all ONUs. By contrast, upstream data, transmitted from the ONUs can only be received at the OLT. Thus, the downstream and upstream links are considered as Point-to-MultiPoint (PMP) and point-to-point, respectively.

In cases where the wireless base station and the OLT are collocated, the wireless and optical parts can be integrated, resulting in an intelligent network. The key point in the integrated architecture is that the BS and the ONU can share their internal data, such as detailed information about bandwidth requests, bandwidth allocation and packet scheduling, which can be used to improve the overall performance. An intelligent scheduler at the integrated

ONU-BS should leverage this extensive information to achieve better performance and QoS guarantee.

## 3. Integration concept and literature review

Researchers have proposed a number of methods to integrate wireless networks, such as WiMAX or LTE, with passive optical networks. The idea of using information already available at the wireless BS to assist the optical part is first introduced in [27]. But no specific methods or performance evaluation has been done. To the best of our knowledge, none of the researchers devised a comprehensive method to use the information to improve the performance of the optical domain, and hence improve the overall performance.

Some addressed hybrid optical-wireless network by proposing a single MAC frame and QoS structure on both domains [6,11] while others translated one technology's QoS parameters into another's [2,5,12,14,29,32]. Due to the differences between optical and wireless systems, such as their scalability requirements, it is too difficult to propose a single QoS structure for both technologies without sacrificing scalability.

Sarkar et al. presented an architecture called Wireless-Optical Broadband Access Network (WOBAN) [17–24]; This work proposed several approaches on the placement of ONU-BS nodes. The work presented in [6] addressed the integration of WiMAX and Optical Burst Switching (OBS) and evaluated the performance for BE and UGS traffic flows. Another architecture, namely MARIN, that incorporates multiple OLTs in the network was proposed in [26].

Ou et al. [14] devised a method to dynamically map the WiMAX classes to the Gigabit PON (GPON) classes. The design was implemented using commercially available equipments to deliver video on demand service to test users. The work was later enhanced in [32] and an Excess Distribution DBA scheme was adopted. Yan et al. proposed a DBA algorithm that takes channel condition into account [30]. Another channel aware algorithm was presented in [2]. The proposed scheduling algorithm considers wireless condition of the cell and head of line delay of rtPS queue as well as queue length. The authors of [5] proposed a two-step mapping that combines BE and nrtPS into a priority queue and rtPS and ertPS into another one and then maps them to EPON classes of traffic. Obele et al. [11,12] proposed an architecture based on a unified QoS infrastructure which is similar to IEEE 802.16a. The performance of their proposed architecture was analyzed under self-similar and long-reach-dependent traffic loads. Wang et al. [29] and Luo et al. [9,10] also used a similar unified QoS scheme throughout the network. The work described in [31] used the time of next EPON polling to control the admission of delay-sensitive traffics.

Amongst the various challenges that ensuring QoS on EPON poses, the dynamic bandwidth assignment of the upload channel attracts more research attention. Unlike the download channel where the bandwidth assigner has full knowledge of the bandwidth needs, the upload bandwidth assigner, which is located at the OLT is not aware of real time bandwidth requirements of each ONU. Therefore, a mechanism is required to report real time bandwidth requirements of each individual ONU to the OLT.

To exchange the bandwidth requirement information, a request/grant mechanism, called Multi-Point Control Protocol (MPCP), is used. In MPCP, ONUs send information regarding their bandwidth requirements to the OLT. For example, in a widely used method, ONUs send their queue length to the OLT at the end of the grant slot. OLT processes the results and sends the bandwidth assignments back to ONUs by grant packets. The problem is that MPCP introduces latency which degrades the performance of the bandwidth assigner, i.e., the bandwidth assigner in OLT can only know the ONU bandwidth requirements with some delay. The MPCP request is sent once in each EPON cycle. As a consequence, the packets that are received after sending the request remain in ONU buffer for an extra EPON cycle.

Using information about the incoming traffic can alleviate the degradation of the extra delay. Integrated WiMAX/EPON provides us with a tool to give this information to OLT. Prior to sending upstream data, the user requests the required wireless bandwidth from the WiMAX BS. The BS scheduler assigns the available bandwidth based on these requests and then sends the grant in UL map field of WiMAX frame. Therefore, the BS has full knowledge of the incoming data for at least the next coming frame. If this information becomes accessible to the OLT scheduler, it can be used to predict the incoming traffic which would result in a more effective bandwidth scheduling. However, in order to achieve optimal performance, an intelligent design that considers the EPON cycle and WiMAX frame length as well as the traffic load and its classes is required.

### *3.1. Effect of sharing knowledge of incoming traffic on the delay*

Predicting the incoming traffic makes it possible to request the required bandwidth for transmitting a packet that has not yet arrived. At the next EPON cycle, due to pre-assignment of the required bandwidth, the packets that have just arrived at the ONU, are transmitted to the OLT right away without experiencing any additional delay for exchanging MPCP request/grant messages.

Since the WiMAX BS assigns bandwidth to the Mobile Subscribers (*MSs*) at the beginning of each frame, it is aware of the quantity of data to be received during the following frame. Fig. 2 compares the process of sending a marked packet in normal and integrated scheduling. As it can be seen in the Fig. 2, predicting the packet reduces its queuing delay. This helps ensure QoS for delay sensitive applications.

In order to use knowledge of incoming traffic to optimize the bandwidth assignment in hybrid optical-wireless networks, the length of the WiMAX frame and EPON cycles should be carefully adjusted. The essential condition for predicting incoming traffic is that the WiMAX frame length should be longer than the EPON cycle length. For better illustration, consider Fig. 3(a) in which the WiMAX frame is considerably shorter than EPON cycles. Since the BS knows only the traffic of the current frame, it is not possible to accurately predict the amount of traffic that accumu-

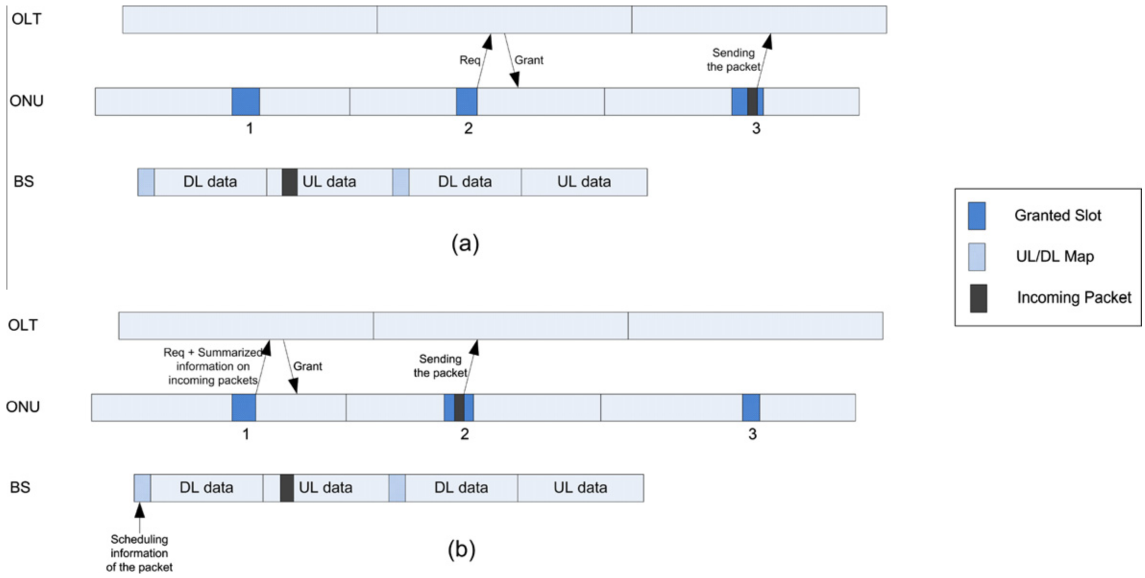



**Fig. 2.** Timing diagram of sending a marked packet (a) normal scheduling (b) integrated scheduling [moved accordingly].

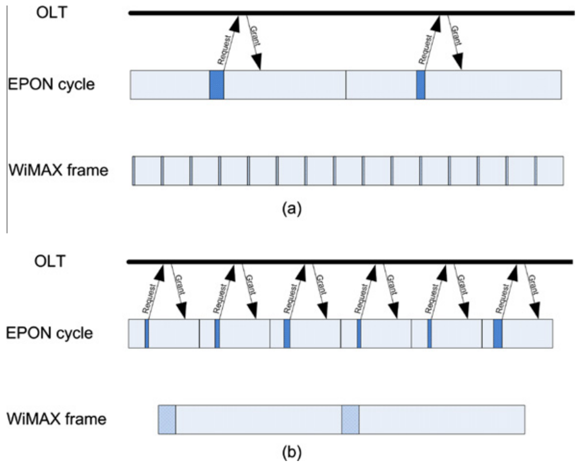



**Fig. 3.** EPON cycle in comparison with WiMAX frames.

lates until the next EPON cycle. Therefore, only a fraction of incoming traffic can be predicted. Hence, the performance improvement that could be gained from full knowledge recedes. By contrast, Fig. 3(b) illustrates the case in which WiMAX frame is longer than EPON cycle. Hence, the ONU-BS knows the quantity of data to be sent to the OLT at the next EPON cycle.

The key part of the proposed algorithm is estimating the amount of traffic that an ONU-BS will have to relay to the OLT in the next grant. Without lack of generality, we assume that all received data at the ONU-BS should be relayed to the OLT, i.e., there is no local traffic at ONU-BS. Given that the amount of incoming traffic is known, ONU-BS requests OLT for the amount of bandwidth, denoted by $B_r$ below.

$$B_r = Q + B_p(T_{nx}), \tag{1}$$

where $Q$ denotes the current length of the queue and $B_p(T_{nx})$ is the predicted incoming traffic that will be received up to beginning of the next grant, which is represented by $T_{nx}$.

It is essential to know the time of next transmission to estimate the bandwidth requirement. If EPON uses fixed length cycles, the next time of transmission is simply calculated as,

$$T_{nx} = T_g + C, \tag{2}$$

where $T_g$ is the grant window beginning time and $C$ is the cycle length. It is worth mentioning that this method does not account for the variations that occur because of different ONU grant sizes in each EPON cycle. A more accurate approach is to calculate $T_{nx}$ based on the internal data of OLT scheduler. This approach is applicable to both fixed-length cycle and adaptive-length cycle algorithms.

## 4. Algorithm description

The proposed algorithm consists of two phases. The core of the algorithm is the inter-ONU scheduling which is performed at OLT scheduler. This part of the algorithm does not consider different classes of traffic and assigns the bandwidth to each ONU based on the available bandwidth, ONU requests and traffic prediction. The details of inter-ONU scheduling is described in Section 4.1.

The second part of the algorithm, namely intra-ONU scheduling, is performed at each ONU to assign the granted bandwidth to each traffic class. The detail of this part of the algorithm is explained in Section 4.2.

The algorithm is independent of the WiMAX BS scheduler. Therefore, it does not impose any requirement on the WiMAX scheduler. The only modification needed at WiMAX BS is that its upload bandwidth assignments, i.e., UL map, should be accessible to the ONU scheduler.

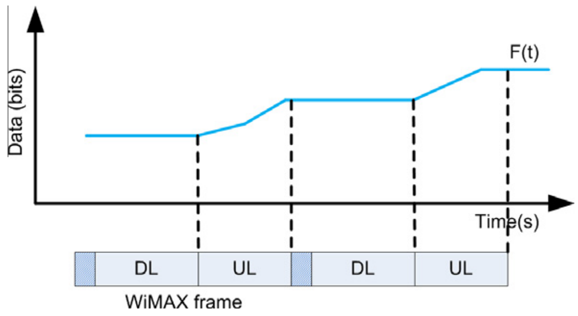



**Fig. 4.** Relationship between $F(t)$ and WiMAX frame.

Although choosing the WiMAX scheduler may affect the overall performance, the study of WiMAX scheduling algorithms and their impact is beyond the scope of this work. A comprehensive survey of WiMAX scheduling algorithms can be found in [28].

### 4.1. Inter-ONU scheduling

Information about incoming traffic in the next WiMAX frame is abstracted in the function $F(t)$, which is defined as the aggregated traffic at time $t$. $F(t)$ predicts the ONU queue length at time $t$. It is an increasing function of time that yields the amount of traffic that has been received up to time $t$. The relation between WiMAX frame and $F(t)$ is illustrated in Fig. 4.

Every time a UL map is generated at the BS, $F(t)$ is defined for the next WiMAX frame. Note that it models the worst case scenario and depending on the type of the scheduler used in BS, the real incoming traffic can be slightly less.

The summarized information, i.e., $F(t)$, is sent to the OLT along with bandwidth requests to be used in the scheduling process. In order to limit message exchanges and keep the OLT scheduler as simple as possible, the information regarding the incoming traffic class is not transmitted to the OLT. However, this can be simply implemented. $F(t)$ can be modeled by a linear piecewise function of time.

$$F(t) = \begin{cases} a_0t + b_0, & t_0 < t \leqslant t_1, \\ a_1t + b_1, & t_1 < t \leqslant t_2, \\ a_2t + b_2, & t_2 < t \leqslant t_3, \\ \vdots \\ a_nt + b_n, & t_n < t \leqslant t_{n+1}, \end{cases} \quad (3)$$

where $a_i$ and $b_i$, $i = 1, \ldots, n$ are coefficients. To simplify the calculations and prohibit time wrapping, time is typically measured from a common origin such as beginning of the cycle. Note that $F(t)$ is only defined in the prediction range $(t_0, t_{n+1}]$ which is discussed more in the next paragraphs.

The slope of $F(t)$ depends on the user's transmission modulation and coding. In scenarios where all of the MSs use the same modulation and coding technique and TDD duplexing is used, $F(t)$ could be simplified to

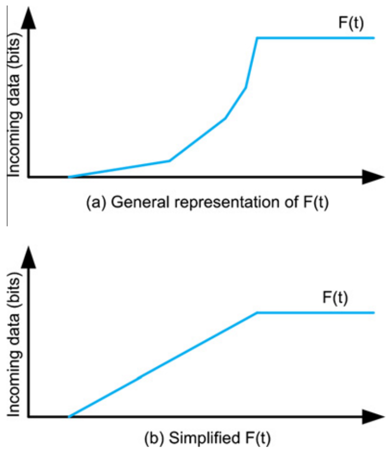



**Fig. 5.** A representation of $F(t)$: (a) General case, (b) simplified.

$$F(t) = \begin{cases} b_0, & t \leqslant t_1, \\ a_1t + b_1, & t_1 < t \leqslant t_2, \\ b_2, & t_2 < t \leqslant t_3. \end{cases} \quad (4)$$

This case is not unreal and it happens if the network operates at maximum capacity, where all users transmit at the maximum modulation and coding rate. An example of $F(t)$ is shown in Fig. 5.

The scheduler uses $F(t)$ to estimate the incoming traffic and grant the appropriate amount of bandwidth. Since it is assumed that TDD is used as the duplexing method, there are active and idle periods. Data packets are arriving during the active periods, where SSs are allowed to send. In WiMAX, the active period for upstream is the part of the UL subframe that is granted to the SSs. During idle periods users are not allowed to send and hence $F(t)$ remains constant.

$$F(t) = \begin{cases} b_0, & t_0 < t \leqslant t_s^1, \\ a_1t + b_1, & t_s^1 < t \leqslant t_e^1, \\ a_1t_e^1 + b_1, & t_e^1 < t \leqslant t_s^2, \\ \vdots \\ a_nt_e^n + b_n, & t_e^n < t \leqslant t_s^{n+1}, \end{cases} \quad (5)$$

where $t_s^i$ and $t_e^i$ are the beginning and end of $i$th activity period, respectively. From WiMAX TDD point of view, these are equivalent to beginning and end of the UL slot. Normally, the WiMAX scheduler puts together the grant slot of all users that transmit with the same modulation and coding. It also packs the grant at the beginning of the slot [13]. Therefore, it is not expected to have more than one activity period in the prediction range. It also makes the transmission of abstract data easier.

Considering the WiMAX rate, time division duplexing and request scheme, $F(t)$ is,

$$F(t) = \begin{cases} Q, & t_r < t \leqslant t_s^1, \\ R_Wt + Q - R_W \times min(t_s^1, t_r), & t_s^1 < t \leqslant t_e^1, \\ R_Wt_e^1 + Q - R_W \times min(t_s^1, t_r), & t_e^1 < t < t_s^2, \end{cases} \quad (6)$$

where $R_W$ is the WiMAX transmission rate, $Q$ is the request or equivalently, ONU's queue length at the time of transmitting request, and $t_r$ is the time of request transmission. $F(t)$ is

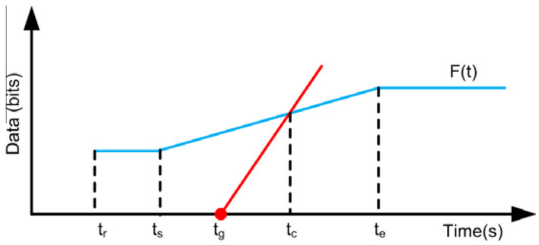



**Fig. 6.** Illustration of $F(t)$ and request corrections method.

illustrated in blue in Fig. 6. At the time of sending the request, $F(t)$ is initialized to $Q$. Then it increases with the WiMAX transmission rate when the UL period begins. The increase continues until the end of UL period and then remains constant.

The prediction information is updated regularly in a way that $t_r < t_s^2$, i.e., the updated prediction range overlaps with the previous one. Thus makes continuous prediction possible.

Since the WiMAX frame is supposed to be longer than the EPON cycle, it is possible that receiving data continues between the grant periods. The ideal solution is to consider these newly arrived packets and extend the grant to transmit them, if possible. The grant can be extended until the buffer becomes empty and all of the newly arrived packets are transmitted.

In our proposed algorithm, the newly arrived data is predicted with $F(t)$. In order to consider this prediction in the dynamic bandwidth assignment algorithm, it is required to update the requests, fed into the DBA, with the predictions. Since the correction depends on the DBA results, it is done iteratively. This way the DBA bandwidth assignments and request corrections are calculated jointly.

The DBA outcome without applying the request corrections is considered as the initial bandwidth assignment. To calculate the request correction, the granted bandwidth is modeled with the same approach as incoming data. This model, which we denote by $G(t)$, on the other hand, describes transmitted bits. It is modeled as a line segment which starts from $t_g$ with a slope of $R_E$. $t_g$ represents the start time of the granted window and $R_E$ represents the EPON transmission rate. These models are illustrated in Fig. 6.

$$G(t) = \begin{cases} 0, & t \geqslant t_g, \\ R_E \times (t - t_g), & t > t_g. \end{cases} \quad (7)$$

The reception prediction function $(F(t))$ and transmission function $(G(t))$ coincide with each other at time $t_c$. This is the time that ONU queue is expected to become empty if enough bandwidth is available.

$$F(t_c) = G(t_c). \quad (8)$$

By substituting $F(\cdot)$ and $G(\cdot)$ from (6) and (7), respectively and solving the equation for $t_c$, we have,

$$t_c = \begin{cases} t_g + \frac{Q}{R_E}, & t_g \leqslant t_s - \frac{Q}{R_E}, \\ \frac{R_W t_s - Q - R_E t_g}{R_W - R_E}, & t_s - \frac{Q}{R_E} < t_g < t_e - \frac{1}{R_E}(Q + R_W t_e - R_W t_s), \\ t_g + \frac{Q + R_W t_e - R_W t_s}{R_E}, & t_g \geqslant t_e - \frac{1}{R_E}(Q + R_W t_e - R_W t_1), \end{cases} \quad (9)$$

where $t_s$ and $t_e$ denote the beginning and the end of the next grant slot, respectively. The calculation should be done for all ONUs. Then the request correction of each ONU is,

$$R_{new}^k = R_E \times (t_c^k - t_g^k), \quad (10)$$

where $R_{new}^k$ is the new request for ONU $k$. $t_c^k$ and $t_g^k$ denote $t_c$ and $t_g$ for ONU $k$, respectively.

The correction process is repeated for each ONU until the granted bandwidths stall or a predetermined maximum iteration is reached. Then the granted bandwidth is sent to the ONUs along with the grant start times and predicted incoming traffic.

**Algorithm 1** (*Proposed Inter-ONU DBA pseudo-code*).

**Require**: $Q(i),\ i = 1, \ldots, N$
**for** $i := 1$ **to N do**
$R(i,1) \leftarrow Q(i)$
**end for**
for $k := 1$ to $Ite_{max}$ **do**
$R_{ex}^t \leftarrow 0$
$B_{ex}^t \leftarrow 0$
**for** $i := 1$ to N **do**
**if** $R(i,k) > W_{max}$ **then**
$R_{ex}(i) \leftarrow R(i,k) - W_{max}$
$B_{ex}^t \leftarrow B_{ex}^t + B_{ex}(i)$
**else**
$B_{ex}^t \leftarrow B_{ex}^t + W_{max} - (R(i,k))$
**end if**
**end for**
**for** $i := 1$ **to N do**
**if** $B_{ex}^t > R_{ex}^t$ **then**
$G(i) \leftarrow R(i,k) + \frac{B_{ex}^t - R_{ex}^t}{N}$
**else**
**if** $R(i,k) < W_{max}$ **then**
$G(i) \leftarrow R(i,k)$
**else**
$G(i) \leftarrow W_{max} + \frac{R_{ex}(i)}{R_{ex}^t} \times B_{ex}^t$
**end if**
**end if**
**end for**
Update $T_{sch}$
Update $P(i, G, T_{sch}, F_i(t))$
$R(i, k+1) \leftarrow R(i,k) + P(i, G, Tsch)$
**end for**
Update $W_{max}$

The pseudo-code of an inter-ONU DBA scheme with prediction is shown in Algorithm 1. The pseudo-code demonstrates how the prediction could be applied to a DBA algorithm.

Note that the time frame in each ONU and OLT is different due to propagation delay. For the sake of simplicity, this effect is not considered in the pseudo-code, but it has to be considered in simulations.

The notations used in the pseudo-code are explained in Table 1.

### 4.2. Intra-ONU scheduling

After the available bandwidth is divided between the ONUs, the intra scheduler at ONU has to divide the granted

**Table 1**
Notations used in algorithm.

| Notation | Description |
|---|---|
| $N$ | Number of ONUs |
| $R(i,k)$ | Request of ONU $i$ in iteration $k$ |
| $Q(i)$ | Queue length at the beginning of grant |
| $Ite_{max}$ | Maximum iteration for solving the bandwidth allocation |
| $R^t_{ex}$ | Total extra requested bandwidth |
| $B^t_{ex}$ | Total extra bandwidth available |
| $G(i)$ | Granted bandwidth to ONU $i$ |
| $P(i,G,T_{sch})$ | Prediction of incoming traffic for ONU $i$ according to granted time slots. Calculated according to Eq. 9 |

bandwidth between different traffic classes. Here, we consider three different classes of traffic, namely expedited forwarding (EF), assured forwarding (AF) and best effort (BE). Here, we adopt strict priority in the ONU scheduler. Therefore, the ONU scheduler serves EF, the highest priority class first. Then it serves AF and after that BE.

When strict priority is applied in the ONU scheduler, high priority packets that are recently received can consume the granted bandwidth that was originally requested by the lower priority packets. It causes the delay of lower priority packets to decrease significantly in light load ONUs. The phenomenon is called light load penalty. A remedy for this problem is suggested in [8]. The suggestion is to move the contents of queues in order of their priority to a separate transmit buffer at the time of sending the request. This buffer will be transmitted at the next grant regardless of the higher priority packets that may arrive after that. After the transmit buffer is emptied, other queues are served in order of their priority. In other words, the transmit buffer acts as the highest priority queue. This algorithm is called two-stage DBA.

We propose a variant of two-stage DBA that does not require another buffer. At the time of sending the request, the ONU records the length of each queue. The recorded request is then used to determine how many packets can be sent from each queue during the next grant.

The ONU scheduler works in two rounds. In the first round, it serves the queues in order of their priority up to the saved request for each queue. Since the queues are First-In-First-Out (FIFO), it means that only the packets that were in the queue at the time of request ($t_r$), are served. After that, the scheduler starts transmitting the next priority queue. The second round starts after all queues are swept once and all packets that were in the queues at request time are transmitted. In this cycle the ONU scheduler performs normal strict priority scheduling to serve the newly arrived packets. The pseudo-code of the algorithm is shown in Algorithm 2.

When DBA is enhanced with the prediction, it basically assigns bandwidth to incoming packets. But as it was described in the previous section, the class of the incoming traffic is not differentiated in order to keep the algorithm simple. To be able to use the bandwidth that is assigned to incoming packets, the accounted predicted traffic is also received along with the grant size. This predicted traffic granted bandwidth is represented with $G_p$. The scheduler adds $G_p$ to the highest priority traffic request. Since the incoming traffic at the highest priority is taken into account, the remaining of $G_p$ is used to transmit the newly received packets of the next priority queue. This continues to lower priority queues. After the first round, the second round is performed without any modification.

The pseudo-code of the algorithm is shown in Algorithm 2. In the pseudo-code, $R(i)$ denotes the length of the queue $i$ at time of request, $Q(i)$ denotes the length of queue at the beginning of grant, $G$ is the granted bandwidth and $P$ is the predicted coming traffic. $G_c(i)$ represents the granted bandwidth for class $i, N_c$ is the number of classes and $B_{rem}$ represents the remaining unused bandwidth.

**Algorithm 2** (*The intra-ONU scheduling algorithm pseudo code*).

**Require**: $R(i),\ i = 1,\ldots,N_c$
**Require**: $Q(i),\ i = 1,\ldots,N_c$
**Require**: $G,\ P$
  First round:
  $R(1) \leftarrow R(1) + G_p$
  **for** $i := 1$ **to** $N_c$ **do**
    $G_c(i) \leftarrow min\left(R(i), G - \sum_{k=1}^{i-1} G_c(k)\right)$
  **end for**
  Second round:
  **for** $i := 1$ to $N_c$ **do**
    $B_{rem} \leftarrow min\left(0, G - \sum_{k=1}^{N_c} G_c(k)\right)$
    $G_c(i) \leftarrow G_c(i) + min(Q(i) - R(i), B_{rem})$
  **end for**

## 5. Maximum throughput analysis

The maximum throughput is achieved when the whole granted bandwidth is used. Therefore in order to calculate the maximum throughput, it is assumed that all ONUs are backlogged. A direct result is consuming the whole granted bandwidth. In each EPON cycle, which is denoted by $C$, the factors that are explained in following paragraphs, waste the bandwidth.

The required gap between ONU's granted slots, known as inter-ONU guard time ($G_1$), is a contributing factor to bandwidth waste. In each cycle, there are $N_{onu}$ gaps between the grants; where $N_{onu}$ is the number of active ONUs in the network. Another contributing factor is the gap between two consecutive cycles, which we represent as $G_2$. The bandwidth that is used for sending the report messages also reduces user bandwidth. Again, there are $N_{onu}$ report messages in each EPON cycle.

Adding up all of the waste factors, it is possible to work out the maximum throughput. It is assumed that the DBA algorithm that is used and the round trip time of ONUs, let the grant slots to be assigned without any gap in between other than the described ones. The maximum throughput is,

$$R_{max} = 1 - \frac{N_{onu} \times G_1 - G_2 - N_{onu} \times \frac{L_{rep}}{R_E}}{C}, \tag{11}$$

where $R_{max}$ is the maximum throughput and $L_{rep}$ is the length of the report message in bits.

Some DBA algorithms such as the one presented in [3], require all ONU requests to process the bandwidth assignment. In these algorithms, the walk time between frames wastes a considerable amount of bandwidth. The walk time is basically composed of three different components; waiting time for reception of the last ONU request, the processing time of scheduling algorithm and the round trip time of first ONU. The last factor comes from the fact that the link remains idle for the duration of transmitting the first ONU's grant and receiving data from it. Supposing the processing time is negligible, it is equal to the first scheduled ONU's round trip time. In summary, walk time is defined as the time between reception of last report message and reception of the first data bit of the first ONU. In this period, the link is idle and the corresponding bandwidth is wasted. Taking the walk time into account the maximum throughput is worked out as,

$$R_{max} = 1 - \frac{1}{C} \times \left( N_{onu} \times G_1 - G_2 - N_{onu} \times \frac{L_{rep}}{R_E} - T_{rtt} - T_{proc} \right), \quad (12)$$

where $T_{rtt}$ is the round trip time of the ONU that its report is the last report and $T_{proc}$ is the processing time of DBA algorithm.

## 6. Traffic-prediction-assisted algorithms

The proposed traffic prediction mechanism can be applied to any type of DBA algorithm. Since it provides the DBA algorithms with real-time information of traffic, it generally improves the delay performance. There are two approaches among DBA algorithms to EPON cycle length which is a significant factor in the performance of the proposed method. Some DBA algorithms, such as the well-known Interleaved Polling with Adaptive Cycle Timing (IPACT) [7], use dynamic cycle length while others, such as Excess Distribution [3], adapt fixed cycle length. In order to investigate the proposed method's performance in both categories, the proposed method is applied to a DBA algorithm from each category.

The first DBA algorithm is Excess Distribution which is first proposed in [3] and later improved in [4]. This DBA scheme first divides the available bandwidth equally among the ONUs to define the threshold window. All of the ONU's requests that are less than the threshold are granted. Then, the unused remaining bandwidth is distributed equally between the unsatisfied ONUs. The variant used in this study adapts a two-stage scheduling algorithm, as described in Section 4.2, to allocate bandwidth to each of the traffic classes. This algorithm is referred to as Excess Distribution or ED, while the traffic-prediction-assisted version of it, is referred to as Integrated ED. The psuedo-code of the integrated ED is shown in Algorithm 1 as an example of applying the proposed traffic prediction on a DBA algorithm.

Unlike the first algorithm that maintains the fixed EPON cycle length, the second algorithm, namely CB-IPACT, adjusts EPON cycle length according to the introduced traffic load. Another difference is that CB-IPACT grants the requests on-the-fly, i.e., it processes each request independent of others. CB-IPACT grants the request if it is less than a predefined maximum grant. Otherwise, it grants the maximum grant.

$$Grant = min\{Request, \ MaxGrant\}. \quad (13)$$

Then the two-stage approach, explained in Section 4.2, is employed to distribute the granted bandwidth to different traffic classes. Employing the two-stage approach helps satisfy the QoS requirement for each of the classes. Applying traffic prediction to this algorithm by the method, explained in Section 4.1, is fairly simple. The traffic-prediction-assisted CB-IPACT is referred to as Integrated-CBIPACT.

## 7. Simulation platform

The performance of the proposed algorithm is extensively studied using simulations setup under different scenarios. Simulation experiments are conducted with the OPnet simulation package for wireless internetworking V.15 [1]. Custom models for OLT, ONU-BS as well as link are developed to ensure utmost similarity to the real equipment.

WiMAX standard supports several physical layers. This includes Orthogonal Frequency-Division Multiple Access (OFDMA), Orthogonal Frequency-Division Multiplexing (OFDM) and Single-Carrier Time Division Multiple Access (SC-TDMA); each supports Time Division Duplexing (TDD) or Frequency Division Duplexing (FDD). In this study, without lack of generality, we consider SC-TDMA with TDD.

The simulation model is depicted in Fig. 7. Unless mentioned otherwise, the scenario given in Fig. 7 is used as the simulation topology. This simulation model consists of one OLT which is connected to 32 ONU-BSs. The performance of the system is analyzed by simulating it under different DBA algorithms.

The proposed traffic prediction assisted bandwidth assignment method is applied to CB-IPACT and Excess Distribution. The performance of the enhanced algorithms as well as the original algorithms is studied extensively through simulations.

The Ethernet and WiMAX workstations are responsible for generating different classes of traffic. Since we focus on the performance of uplink, all of the traffic load is destined toward a server located near the OLT. The default simulation parameters are shown in Table 2, unless otherwise stated.

To simulate the WiMAX part of the network, a modified version of standard OPnet WiMAX models is used. The added modifications enable these models to share some of their internal information with the developed EPON models, which is vital to implement the proposed DBA algorithms for hybrid optical-wireless networks.

The EPON part of the network is custom designed to implement the point to multipoint aspects of the EPON as well as its Ethernet frame architecture. The EPON models consist of OLT MAC layer, ONU MAC layer and link. These models are accessible from [1].

Optical networks typically performs as the backbone infrastructure. In real implementations, they connect corporate networks as well as wireless base stations to the

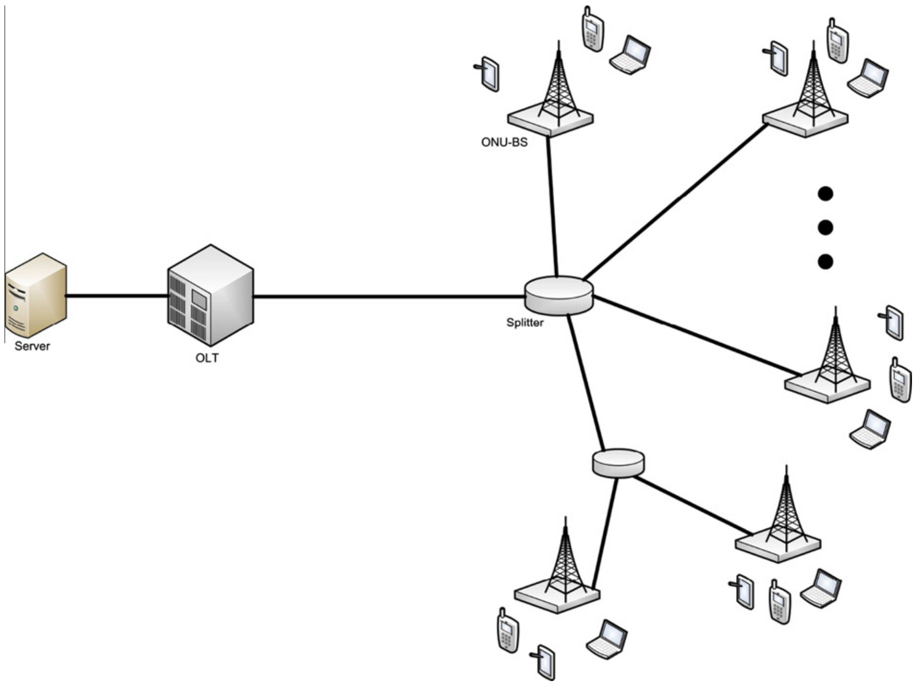



**Fig. 7.** Simulation scenario [replaced by a more meaningful diagram].

**Table 2**
Simulation parameters.

| Simulation parameters | Value |
|---|---|
| EPON cycle length | 2 ms |
| WiMAX phy | TDD<br>SC |
| WiMAX frame size | 5 ms |
| Inter ONU guard time | 1 µs |
| Propagation delay | 10 µs |

**Table 3**
Traffic loads in realistic scenario [added].

| | Low load (Mbps) | High load (Mbps) |
|---|---|---|
| ONU | 12.5 | 150 |
| ONUBS | 4.5 | 50 |

operator's core. This realistic scenario has been studied Section 8.4.

In the realistic scenario, half of the ONUBSs have been replaced by the legacy ONUs, that serve the corporate users, to study their effect on the performance of the proposed algorithm. The legacy ONUs cannot predict the traffic and has to request for bandwidth.

In addition, in a real network the real-time load on the users are not equal. To model this effect, in the scenario, 80% of the total traffic load is generated in 25% high load ONUs and ONUBSs. The remaining load is then generated by low load ONUs and ONUBSs. The traffic load on the ONUs and ONUBSs are shown in Table 3.

### *7.1. Traffic types*

Three traffic types are considered in the simulation experiments. The highest priority traffic is the Expedited Forwarding (EF) which models the voice or video traffic. The Quality of Service requirements of this class are defined by maximum tolerable delay, maximum jitter, maximum loss and throughput.

The second priority traffic is the Assured Forwarding (AF). Multimedia streaming traffic falls into this category. This class of traffic is delay-constrained and typically its delivery is assured as long as its throughput remains within predefined limits.

The third, and most basic traffic class is the Best Effort (BE) traffic that models file transfers, email and web traffic. This class has no specific requirement in terms of maximum delay or jitter. However, the throughput plays an important role for this class and hence, some researchers define a minimum throughput requirement to prohibit BE traffic from being cut-off from the network in highly loaded conditions.

### *7.2. Traffic generation*

WiMAX users (Fig. 7) are responsible for generating traffic with the required distribution. The EF traffic load is generated by a Poisson process while AF and BE traffic classes are generated by self-similar models to better model real traffic in the network. To model the traffic more realistically, EF traffic is composed of packets with fixed length equal to 100 bytes, while AF and BE packet length is uniformly distributed between 100 and 1500 bytes. The total traffic load is composed of BE, AF and EF; each of them shares 0.4, 0.4 and 0.2 of the total traffic, respectively [25].

A simple way to generate self-similar traffic is to multiplex several traffic flows with a distribution that has long-range dependence (LRD) property, like Pareto distribution [33]. The density function of Pareto distribution is,

$$f(x) = \frac{\alpha}{k}\left(\frac{k}{x}\right)^{\alpha+1}; \quad x > k, \ \alpha > 0. \tag{14}$$

Since a traffic with heavy-tailed distribution is desired, the shape parameter is bounded ($1 < \alpha < 2$). The degree of self-similarity which is normally expressed as Hurst parameter ($H$) is related to shape parameter by,

$$H = \frac{3-\alpha}{2}, \quad 1 < \alpha < 2. \tag{15}$$

The model used in simulation experiments, is a superposition of Fractal Point Processes (Sup-FRP). The Hurst parameter is set to 0.8 which is suitable for web traffic as well as file transfer [15]. The fractal onset time scale is set to 0.1 s.

The simulation scenario consists of three users per each ONU-BS. Each user is responsible of generating one class of traffic. This gives us the flexibility of modifying each traffic class and its corresponding parameters without affecting other classes of traffic.

## 8. Simulation results

In this section, the performance of the proposed algorithm is investigated under different conditions. Unless stated otherwise, the cycle length and the frame length are equal to 2 ms and 5 ms, respectively. Moreover, the load is sufficient enough to saturate the network.

The remainder of this section compares the performance of the CB-IPACT, Excess Distribution, Integrated-CBIPACT and Integrated ED, under different load conditions. For the sake of comparison, some of the performance criteria are also measured for static scheduling. The effect of modifying key parameters is also investigated.

### *8.1. Load analysis*

In the following paragraphs, the behavior of the proposed algorithms under different loads is investigated. Under light and moderate loads, the network throughput, shown in Fig. 8, is the same for all algorithms. This is an obvious result of introducing less load than network capacity which all of the DBA algorithms are capable of delivering. The throughput performance curves separate when the introduced load approaches the link capacity.

Fig. 9 compares the delay of the EF flow for different algorithms as the load increases. As expected, an increase in load results in more delay. Under lower loads, the gradual increase in the delay of CB-IPACT is mainly because of the increase in cycle length. The cycle length of CB-IPACT extends with load which in turn results in 1 ms increase in average delay. The probability distribution function of delay of EF flows (Fig. 11) clearly shows that increasing the load results in longer delays. In full load conditions, cycle length saturates to its maximum and the EF delay is increased by more than the maximum cycle length. The same discussion applies for AF flows with the minor difference that AF flows experience a faster increment due to their lower priority.

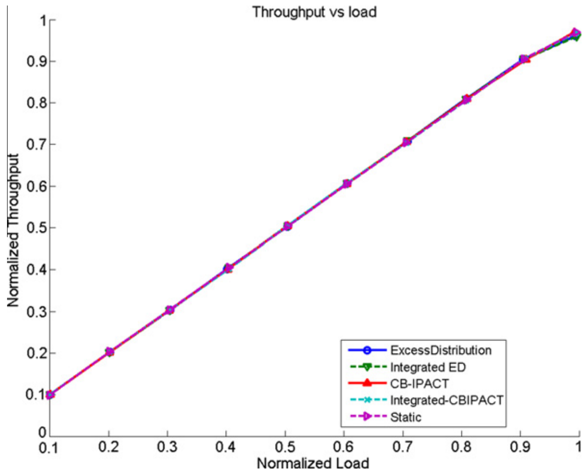



**Fig. 8.** Normalized throughput versus normalized load.

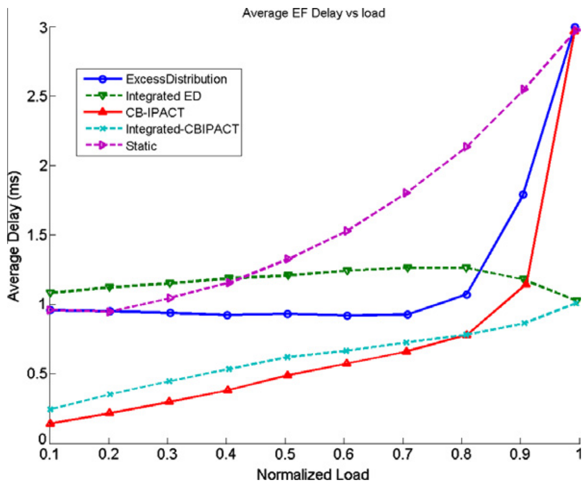



**Fig. 9.** Delay of EF flow in different loads.

Since incoming traffic prediction assumes that the traffic is continuous, the scheduler is not aware of the bandwidth requirement of each individual packet. Therefore, the scheduler may assign a portion of the required bandwidth to send the packet. Since fragmentation of packets is not permitted in EPON, this granted bandwidth would be wasted. The inaccuracy caused by partial bandwidth assignment slightly deteriorates the performance of the integrated algorithms and causes the delay to increase slightly. However, the reward of bearing a slightly longer delay in light loads is a significant delay decrease in high load, as shown in Figs. 9 and 10.

Both Excess Distribution and integrated ED algorithms fix a cycle length, equal to maximum cycle length of CB-IPACT and integrated-CBIPACT. As a result, their cycle length is longer than the average cycle length of CB-IPACT

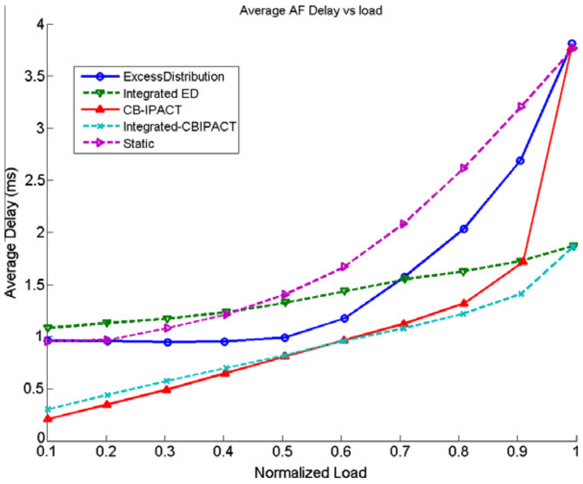



**Fig. 10.** Delay of AF flow in different loads.

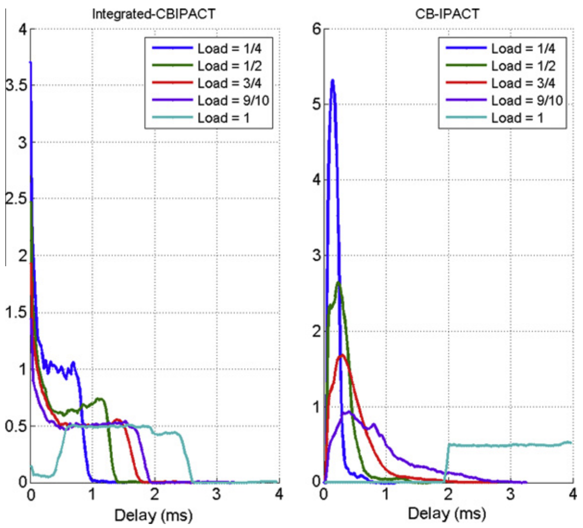



**Fig. 11.** Probability density function of EF delay with CB-IPACT and integrated-CBIPACT.

and integrated-CBIPACT in light and moderate loads. Therefore, their delay is generally greater than that of CB-IPACT and integrated-CBIPACT. ED can keep constant delay when load is up to 75% of link capacity for EF traffic and up to 50% of link capacity for AF traffic. This is because higher priority traffic can take advantage of the excess bandwidth which is distributed between ONUs, to send their packets. Again, Since AF has lower priority, it is more susceptible to increasing load and its effect can be seen sooner. The delay of both flows increases significantly at full load. However, if the proposed traffic prediction algorithm is applied, it counteracts the load effect and significantly decreases the delay. The effect of prediction can be best demonstrated with the probability distributed function of the delay that is shown in Fig. 12.

Fig. 9 shows an interesting phenomenon. The average delay of the EF flow decreases as load increases from 0.8 to 1. In moderate loads, there is an amount of excess bandwidth that is distributed among the ONUs. This excess bandwidth varies significantly for each cycle and as a result, the granted bandwidth also significantly varies. Therefore, the position of the granted slot in the cycle changes. This might increase the cycle length that is measured in some ONUs and as a result, increase their average delay. Not all of the ONUs are affected by the phenomenon. Since both ED and CB-IPACT serve ONUs in a round robin fashion, the variation in grants does not affect first and last ONUs. However, the full effect of variation can be seen at the ONUs that are granted in the middle of the cycle. When the load increases to full load, less excess bandwidth remains and the granted slots become equal. Therefore, the granted slot does not change its position and the phenomenon vanishes, which is shown in Fig. 9 as a slight decrease in average EF delay.

From QoS point of view, the throughput is almost the same for all four algorithms. The delay bound of the algorithms is the same in light lights. But in heavy loads, integrated algorithms, due to their unique prediction method, allow considerably lower delay bound, at the expense of an insignificant delay increase in light loads. A comparison of fixed and adaptive cycle length algorithms reveals that although both categories have the same delay bound, the adaptive cycle length algorithms perform better in terms of delay, in light load conditions. This holds true for integrated algorithms.

### *8.2. EPON cycle length analysis*

Since the ability to predict incoming traffic depends on the ratio between the lengths of the EPON cycle and WiMAX frame, as discussed in Section 3 and 4, changing the cycle length has a significant impact on the performance of the proposed algorithm. In saturated conditions, where all of the ONUs are backlogged, almost all DBAs assign the available bandwidth equally to the ONUs. Therefore, the difference in performance is mainly due to traffic prediction. By analyzing the full load behavior of the

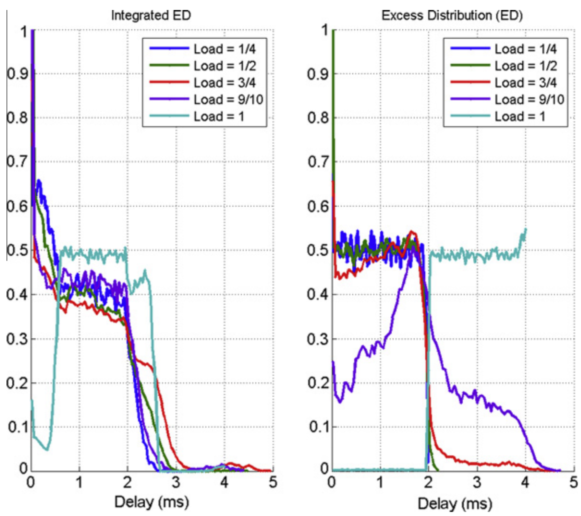



**Fig. 12.** Probability density function of EF delay with ED and integrated ED.

algorithm, it is possible to investigate the prediction performance in nearly isolated environment.

Fig. 13 compares the network throughput in different cycle lengths. The frame length is set to 5 ms in all of the simulations. The throughput is normalized to nominal capacity of the link which is 1 Gbps in EPON. The normalized throughput varies between 0.92 and 0.99. CB-IPACT assigns the bandwidth on the fly, i.e., as the scheduler receives each bandwidth request, it processes the request and sends the grant. On the other hand, ED processes all requests at the same time. Hence, walk time causes a reduction in ED's maximum throughput in comparison with CB-IPACT. Static bandwidth assignment does not perform the assigning based on requests and its maximum throughput is the same as CB-IPACT.

Maximum throughput increases with longer cycle lengths. This is mainly because there is less frequent request/grant exchanges which directly translates in reduced walk time. Therefore, the bandwidth consumed for the request/grant pairs and also the bandwidth wasted in walk times can be conserved.

Fig. 13 also indicates an insignificant throughput reduction in integrated versions of CB-IPACT and ED in comparison with the original algorithms. This is due to the overhead caused by extra message exchange to share the traffic information with the OLT in integrated versions.

The simulated throughput as well as the analytical throughput are shown in Fig. 14. The minor difference between simulation and analytic throughput is caused by the fragmentation effect, which we did not consider in analysis. Since fragmentation is not allowed in EPON, the assigned bandwidth is wasted if the packet is larger than the remaining granted bandwidth. This slightly decreases the throughput in simulation experiments. For shorter EPON cycles, this can happen more frequently and thus the effect is larger.

Fig. 15 shows the average delay of the EF stream with different cycle length. Generally, by increasing the cycle length, the average delay is also increased because there are less frequent grant slots available to the ONU. A packet may remain in queue longer in longer cycles. This justifies the linear increase in the delay shown in Fig. 15. By knowing the incoming traffic in advance, the proposed algorithms can save the time that is normally spent on exchanging request/grant. This results in almost a cycle length reduction in delay when full knowledge of incoming traffic exists. For example, assuming the frame length is 5 ms, the ONU has full knowledge of incoming traffic when the cycle length is equal to 1 ms or 2 ms (refer to Section 3 for details). It is backed by 1 ms and 2 ms delay reduction when cycle length is set to 1 ms and 2 ms, respectively.

By contrast, given that the WiMAX frame length is 5 ms, in 5 ms and 10 ms EPON cycle lengths, ONU has only partial knowledge on incoming traffic and hence a part of incoming traffic is handled in conventional manner, i.e., request/grant mechanism. Therefore it results in only a partial reduction in delay. This phenomenon can be best understood through careful investigation of PDF of the EF delay. The EF delay probability density function of ED and Integrated ED is shown in Fig. 16. When EPON cycle length is set to 2 ms, adding prediction to the algorithm clearly shifts the PDF leftward and decreases the delay. However, when EPON cycle length is 5 ms, it is not always possible to send the required information to the prediction module in time

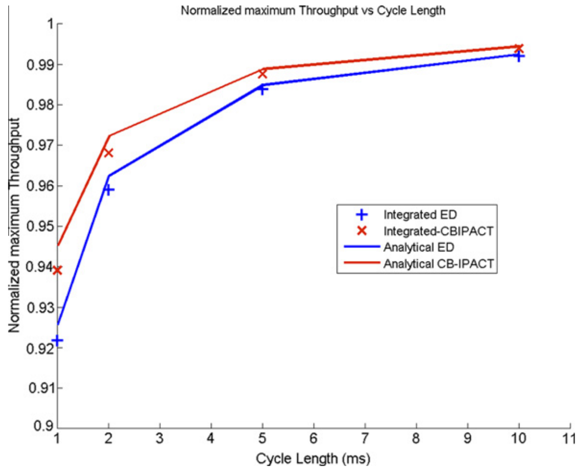



**Fig. 14.** Normalized analytic throughput versus cycle length.

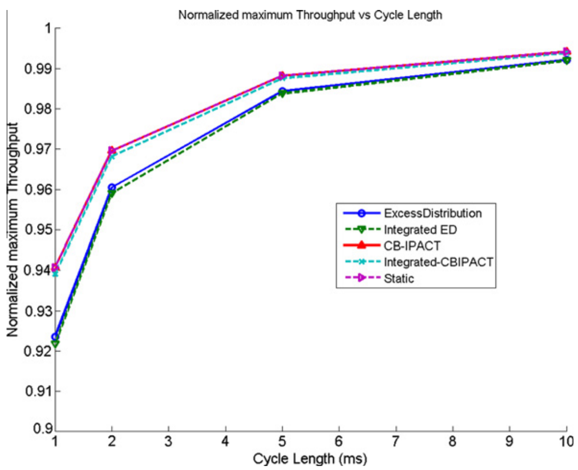



**Fig. 13.** Normalized throughput versus cycle length.

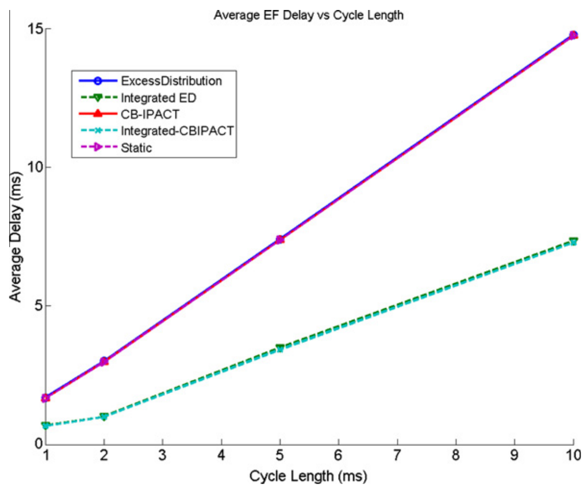



**Fig. 15.** EF stream delay versus EPON cycle length.

and information may be received when it is expired. This is demonstrated in Fig. 17. The required information for the prediction of incoming traffic which is extracted from the DL/UL map is transmitted to the OLT scheduler along with the request. In case of BS1 the information is updated after transmission of request and therefore, the incoming packet has to wait for the next request to be sent if there is no extra granted bandwidth available for sending it. On the other hand, BS2 updates the information before sending request and hence the incoming packet is predicted in the scheduler and the required bandwidth is assigned for transmitting it. The result is the PDF shown in Fig. 16.

The same effect can be seen in Fig. 18 for the AF streams. However, because the prediction module does not differentiate between traffic types, the bandwidth assigned to the predicted traffic is assigned to the highest priority flow, which is EF flow. This does not affect the performance when ONU has full knowledge of the coming traffic. But in case of partial knowledge, i.e., EPON cycle length is 5 ms or 10 ms, the unpredicted EF packets can take advantage of the bandwidth, originally assigned to AF packets. This deteriorates the performance of lower priority packets in case of partial knowledge.

In summary, throughput is not affected by augmenting the algorithms with the prediction mechanism. However the delay of high priority classes can be decreased by a factor of two. Although the method always decreases the delay, it performs best with full knowledge, i.e., when the EPON cycle length is less than WiMAX frame length.

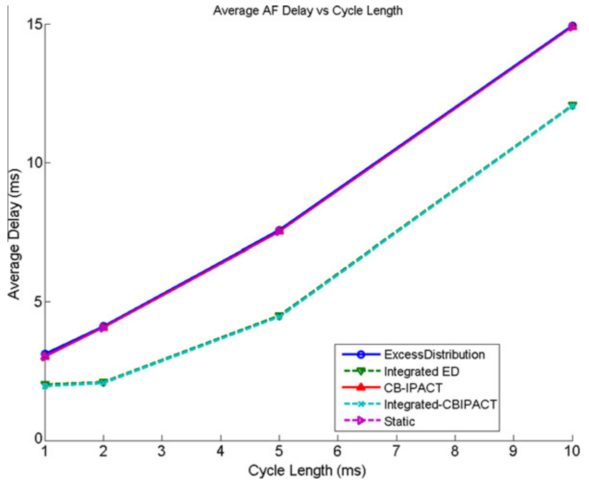


**Fig. 18.** AF stream delay versus EPON cycle length.

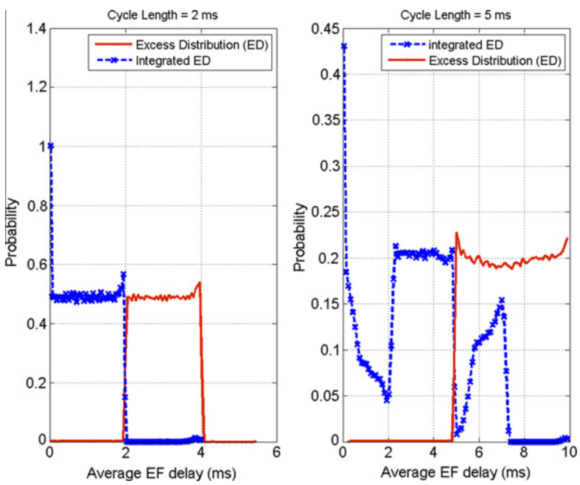


**Fig. 16.** Comparison of EF delay PDF for different EPON cycle length.

### 8.3. WiMAX frame length analysis

Changing the length of the WiMAX frame affects the performance metrics of the proposed algorithms. The only performance metric which is not affected is system throughput which is shown in Fig. 19. The EPON cycle length is fixed to 2 ms in all simulation experiments. Like the previous discussion, the ED's throughput is slightly less than CB-IPACT due to walk time.

The delay of EF and AF streams for different frame lengths is shown in Figs. 20 and 21, respectively. Although, changing the frame length does not affect the amount of traffic load, it changes the frequency and shape of the incoming traffic; i.e., when the frame length is extended, the frequency of incoming traffic decreases, however the amount of traffic delivered instantly to optical network increases.

The delay of the EF stream remains fixed for WiMAX frame lengths up to 5 ms and then it slightly increases, while the delay of AF streams is only fixed for WiMAX frame lengths up to 2 ms and then it increases dramatically afterward. Because WiMAX TDD is used as the physical layer of the wireless part of the network, the incoming traffic is bursty in nature. This does not affect the delay, if

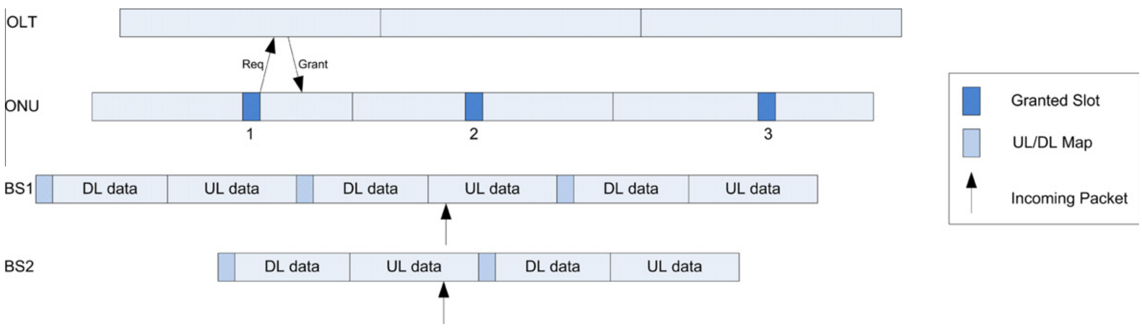


**Fig. 17.** Timing diagram for partial knowledge of incoming traffic.

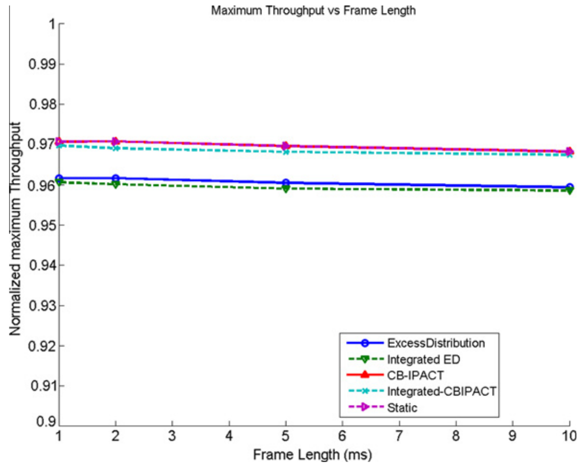



**Fig. 19.** Maximum throughput versus frame length.

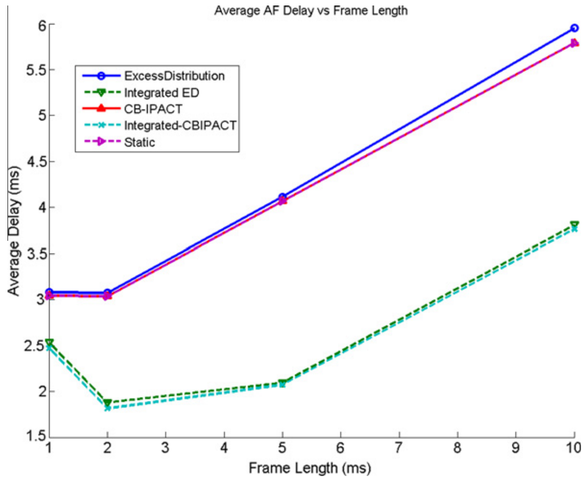



**Fig. 21.** AF stream delay versus frame length.

WiMAX frame length and EPON cycle length are set in a way that EPON sees continuous traffic, i.e., the amount of traffic introduced to EPON is almost the same in each EPON cycle. But as WiMAX frame length grows to more than that of the EPON cycle, the traffic becomes bursty and pulse-shaped. The amount of the traffic in each pulse also increases. In long frames, the amount of traffic delivered to the ONU is more than the maximum grant size. Thus it is divided among several EPON cycles, resulting in increased queuing delay. Since, the prediction algorithm does not provide the information about the amount of traffic in each class, the biggest share of the predicted bandwidth is allocated to the EF stream and hence it exhibits the increasing trend later than AF stream does. The phenomenon can be seen clearly in Figs. 20 and 21.

Extending the frame length leads to extending the range of the prediction, and in turn, decreases the mean EF delay. However, the decreasing trend stops at frame length equal to 5 ms and increases thereafter. As the WiMAX frame is extended, the amount of traffic aggregated in each frame grows. The aggregated traffic is delivered almost instantaneously to the ONU for transmission. In the short frames, the amount of traffic is little enough to be sent in a single EPON cycle. The lack of distinction between the traffic classes in prediction also acts in favor of high priority traffic. Since EF traffic can also use the granted bandwidth that is originally assigned to AF flows. But when the aggregated EF traffic is increased to more than the granted bandwidth, it spreads over two or more EPON cycles, which results in increased delay.

In summary, increasing the WiMAX frame length in comparison to EPON cycle length helps the prediction algorithm predict traffic for a longer range. On the other hand aggregating traffic, caused by longer WiMAX frames, can result in the loss of statistical multiplexing gain. Long WiMAX frame lengths also increases the queuing delay which contracts the delay reduction gained by traffic prediction. Therefore, it is vital to intelligently choose the EPON cycle length as well as the WiMAX frame length to take advantage of the prediction.

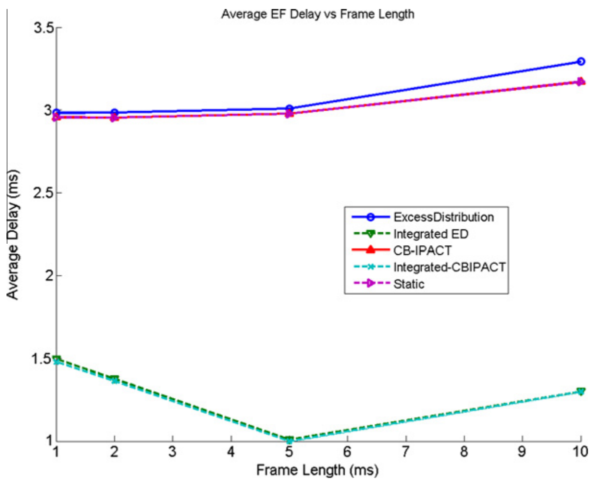



**Fig. 20.** EF stream delay versus frame length.

### 8.4. A realistic scenario

The delay of EF and AF traffic classes are shown in Fig. 22. It shows the delay of high and low load ONU and

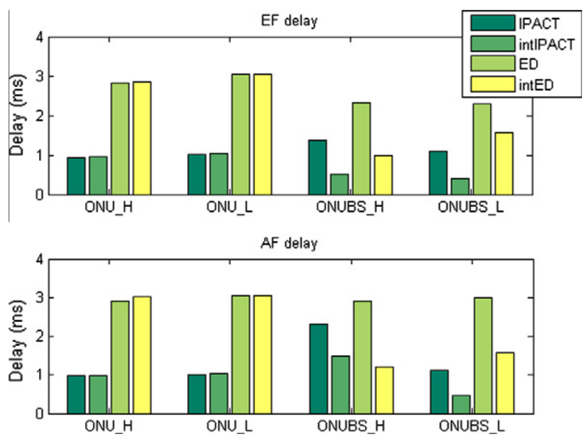



**Fig. 22.** The delay of EF and AF traffic classes in realistic scenario [added].

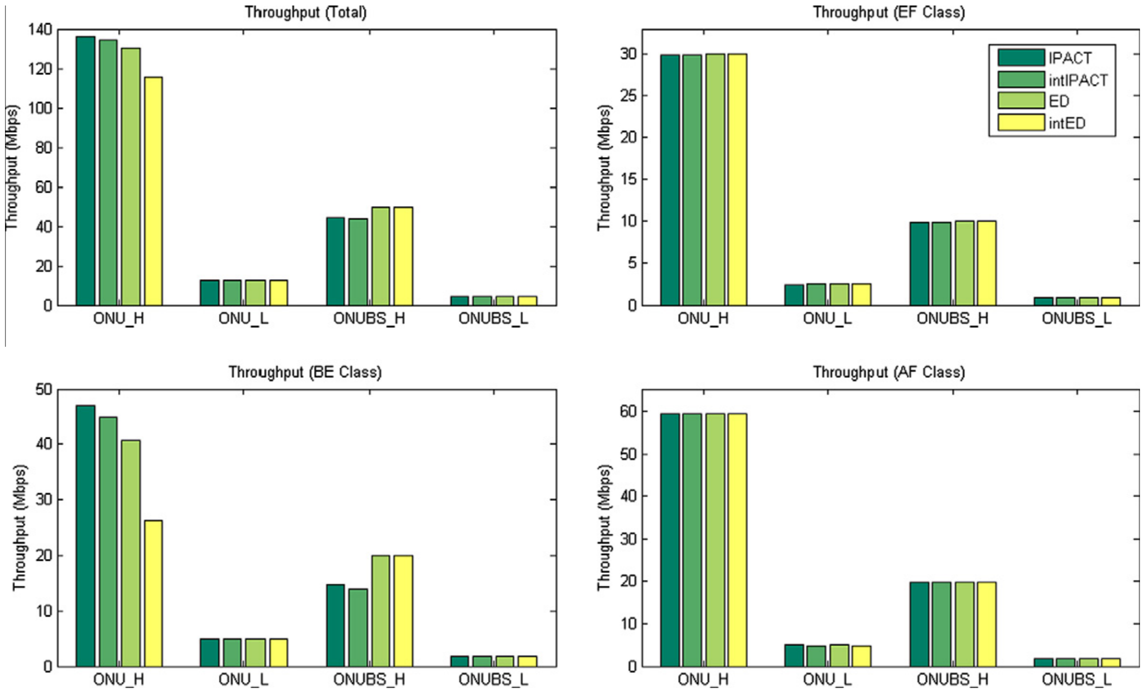



**Fig. 23.** ONU/ONUBS throughput in realistic scenario [added].

ONUBSs when various DBA algorithm are employed. As can be seen in the figure, applying the proposed prediction-based DBA algorithm does not affect the delay of legacy ONUs. However, it lowers the delay of ONUBSs' delay-sensitive packets dramatically. This is due the fact that in cases where the network as a whole is working with full capacity, the total requested bandwidth is greater than the total available bandwidth. Thus, the ONUs and ONUBSs are granted with the requested bandwidth or less which results in an extra delay for requesting bandwidth. Since the proposed scheme pre-grant the packets from ONUBSs, the algorithm benefits both of the high and low load ONUBSs. In summary, the proposed DBA algorithm is beneficial when the network is working in full capacity regardless of the traffic distribution.

Fig. 23 shows the throughput of ONU/ONUBSs. In this figure, High load ONU and ONUBS are represented by `ONU_H` and `ONUBS_H`, whereas `ONU_L` and `ONUBS_L` represent low load ONUs and ONUBSs. It is concluded from the figure that the delay decrement in delay-sensitive classes of traffic comes with a slight decrease of BE throughput in high load ONUs. On the other hand, the high load ONUBSs are granted more bandwidth to enable them to better serve the delay-sensitive classes. It is worth mentioning that the throughput of low load ONUs and ONUBSs are not affected. Also, applying the proposed algorithm does not harm the EF and AF throughput of high load ONUs.

## 9. Conclusion

Hybrid optical-wireless network is a promising technology that is relatively inexpensive and can address the huge bandwidth need of modern applications. In this work, a mechanism was proposed to enhance the performance of the optical scheduler algorithm by augmenting it with a novel incoming traffic prediction, which is extracted from the internal information of the wireless scheduler. The prediction method was described and investigated in detail. Based on the prediction method, two DBA algorithms were proposed. Extensive simulation experiments proved the performance of the algorithms and showed that due to the prediction method applied, the delay of both algorithms in high loads can be decreased by a factor of two, without affecting the throughput. The effects of changing the main parameters were also investigated. In addition to that, the performance of the proposed algorithm was proved in a real world scenario which consists of ONUs and ONUBSs.

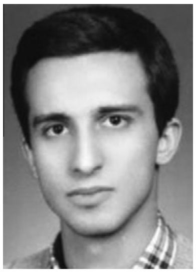
The results showed that predicting the traffic removes the significant delay increase which normally occurs in high load conditions. In fact, the delay of high priority traffic remains the same for high and light load conditions. This enables the service providers to establish a lower delay bound and a better Quality of Service for consumers, even in high load conditions.

Since the proposed algorithm is superior to conventional algorithms, only in the high load conditions, and also because the main part of the algorithms runs in the OLT, it is very simple to further tailor the algorithm to be activated just in high loads, where it is useful the most, and to remain idle otherwise. This results in a simpler DBA in light load conditions and a more efficient one in high load conditions. In future works, we will establish a proper load threshold to activate the proposed DBA algorithm.

## References

[1] Opnet simulation software. [Online]. <www.opnet.com>.

[2] F. Alsolami, P. Ho, A novel two-level hierarchical scheduling scheme for integrated WiMAX and EPON, in: ICS, 2009.
[3] C. Assi, Y. Ye, S. Dixit, M. Ali, Dynamic bandwidth allocation for quality-of-service over ethernet pons, IEEE Journal on Selected Areas in Communications 21 (9) (2003) 1467–1477.
[4] A. Dhaini, C. Assi, A. Shami, Dynamic bandwidth allocation schemes in hybrid tdm/wdm passive optical networks, in: 3rd IEEE Consumer Communications and Networking Conference, 2006, CCNC 2006, vol. 1, 2006, pp. 30–34.
[5] I.-S. Hwang, J.-Y. Lee, C.-W. Huang, Z.-D. Shyu, Advanced dynamic bandwidth allocation and scheduling scheme for the integrated architecture of epon and wimax, IEEE International Conference on Mobile Data Management 0 (2009) 655–660.
[6] K. Katrinis, A. Tzanakaki, S. Dweikat, S. Vassilaras, R. Nejabati, D. Simeonidou, G. Zervas, Backhauling wireless broadband traffic over an optical aggregation network: Wimax over obs, in: Sixth International Conference on Broadband Communications, Networks, and Systems, 2009, BROADNETS 2009, 2009, pp. 1–4.
[7] G. Kramer, B. Mukherjee, G. Pesavento, Ipact: a dynamic protocol for an ethernet pon (epon), IEEE Communications Magazine 40 (2) (2002) 74–80.
[8] G. Kramer, B. Mukherjee, S. Dixit, Y. Ye, R. Hirth, Supporting differentiated classes of service in ethernet passive optical networks, The Journal of Optical Networking 1 (8) (2002) 280–298.
[9] Y. Luo, S. Yin, T. Wang, Y. Suemura, S. Nakamura, N. Ansari, M. Cvijetic, QoS-aware scheduling over hybrid optical wireless networks, in: National Fiber Optic Engineers Conference, 2007.
[10] Y. Luo, T. Wang, S. Weinstein, M. Cvijetic, S. Nakamura, Integrating optical and wireless services in the access network, in: Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference, Optical Society of America, 2006. Paper NThG1.
[11] B.O. Obele, M. Iftikhar, S. Manipornsut, M. Kang, Analysis of the behavior of self-similar traffic in a qos-aware architecture for integrating wimax and gepon, Journal of Optical Communications and Networking 1 (4) (2009) 259–273.
[12] B.O. Obele, M. Kang, Fixed mobile convergence: a self-aware qos architecture for converging wimax and gepon access networks, International Conference on Next Generation Mobile Applications, Services and Technologies 0 (2008) 411–418.
[13] F. Ohrtman, WiMAX Handbook: Building 802.16 Wireless Networks, ser. McGraw-Hill Communications, McGraw-Hill, 2005.
[14] S. Ou, K. Yang, M. Farrera, C. Okonkwo, K. Guild, A control bridge to automate the convergence of passive optical networks and ieee 802.16 (wimax) wireless networks, in: Broadband Communications, Networks and Systems, 2008, BROADNETS 2008, September 2008, pp. 514–521.
[15] J. Potemans, B.V. den Broeck, Y. Guan, J. Theunis, E.V. Lil, A.V. de Capelle, Implementation of an advanced traffic model in opnet modeler, in: OPNETWORK 2003, Washington DC, USA, 2003.
[16] A. Reaz, V. Ramamurthi, S. Sarkar, D. Ghosal, S. Dixit, B. Mukherjee, CaDAR: an efficient routing algorithm for a wireless-optical broadband access network (WOBAN), IEEE/OSA Journal of Optical Communications and Networking 1 (5) (2009) 392–403.
[17] S. Sarkar, S. Dixit, B. Mukherjee, Hybrid wireless-optical broadband-access network (woban): a review of relevant challenges, Journal of Lightwave Technology 25 (11) (2007) 3329–3340.
[18] S. Sarkar, Y. Hong-Hsu, S. Dixit, B. Mukherjee, Hybrid wireless-optical broadband access network (woban): network planning using Lagrangean relaxation, IEEE/ACM Transactions on Networking 17 (4) (2009) 1094–1105.
[19] S. Sarkar, B. Mukherjee, S. Dixit, Optimum placement of multiple optical network units (onus) in optical-wireless hybrid access networks, in: Optical Fiber Communication Conference, 2006 and the 2006 National Fiber Optic Engineers Conference, OFC 2006, March 2006, p. 3.
[20] S. Sarkar, H.-H. Yen, S. Dixit, B. Mukherjee, Dara: delay-aware routing algorithm in a hybrid wireless-optical broadband access network (woban), in: IEEE International Conference on Communications, 2007, ICC'07, June 2007, pp. 2480–2484.
[21] S. Sarkar, H.-H. Yen, S. Dixit, B. Mukherjee, A mixed integer programming model for optimum placement of base stations and optical network units in a hybrid wireless-optical broadband access network (woban), in: Wireless Communications and Networking Conference, March 2007, WCNC 2007, IEEE, 2007, pp. 3907–3911.
[22] S. Sarkar, H.-H. Yen, S. Dixit, B. Mukherjee, Hybrid wireless-optical broadband access network (woban): network planning and setup, IEEE Journal on Selected Areas in Communications 26 (6) (2008) 12–21.
[23] S. Sarkar, H.-H. Yen, S. Dixit, B. Mukherjee, A novel delay-aware routing algorithm (dara) for a hybrid wireless-optical broadband access network (woban), IEEE Network 22 (3) (2008) 20–28.
[24] S. Sarkar, H.-H. Yen, S. Dixit, B. Mukherjee, Radar: Risk-and-delay aware routing algorithm in a hybrid wireless-optical broadband access network (woban), in: Conference on Optical Fiber Communication and the National Fiber Optic Engineers Conference, 2007, OFC/NFOEC 2007, March 2007, pp. 1–3.
[25] A. Shami, X. Bai, N. Ghani, C. Assi, H. Mouftah, Qos control schemes for two-stage ethernet passive optical access networks, IEEE Journal on Selected Areas in Communications 23 (8) (2005) 1467–1478.
[26] W.-T. Shaw, S.-W. Wong, N. Cheng, L. Kazovsky, Marin hybrid optical-wireless access network, in: Conference on Optical Fiber Communication and the National Fiber Optic Engineers Conference, 2007, OFC/NFOEC 2007, March 2007, pp. 1–3.
[27] G. Shen, R. Tucker, C.-J. Chae, Fixed mobile convergence architectures for broadband access: integration of epon and wimax [topics in optical communications], IEEE Communications Magazine 45 (8) (2007) 44–50.
[28] C. So-In, R. Jain, A.-K. Tamimi, Scheduling in ieee 802.16e mobile wimax networks: key issues and a survey, IEEE Journal on Selected Areas in Communications 27 (2) (2009) 156–171.
[29] S. Wang, H. Li, Y. Ji, Distribution qos scheme for a novel of hybrid optical wireless network, in: J. Wang, G.-K. Chang, Y. Itaya, H. Zech (Eds.), Proc. SPIE, vol. 6784, no. 1, SPIE, 2007, p. 67842N.
[30] Y. Yan, H. Yu, L. Dittmann, Wireless channel condition aware scheduling algorithm for hybrid optical/wireless networks, in: Access Nets: Third International Conference on Access Networks, AccessNets 2008, October 15–17, 2008, Springer, Las Vegas, NV, USA, 2009, p. 397. Revised papers.
[31] Y. Yan, H. Yu, H. Wang, L. Dittmann, Integration of epon and wimax networks: uplink scheduler design, in: SPIE, W. Hu, S.-K. Liu, K. ichi Sato, L. Wosinska (Eds.), vol. 7137, no. 1, 2008, p. 713704.
[32] K. Yang, S. Ou, K. Guild, H.-H. Chen, Convergence of ethernet pon and ieee 802.16 broadband access networks and its qos-aware dynamic bandwidth allocation scheme, IEEE Journal on Selected Areas in Communications 27 (2) (2009) 101–116.
[33] L. Ying, K. Feng-ju, Z. Lian-jiong, W. Zhi-Guang, A self-similar traffic generation method and application in the simulation of mobile ad-hoc network, in: International Colloquium on Computing, Communication, Control, and Management, 2009, CCCM 2009, ISECS, vol. 4, 8–9 2009, pp. 229–233.



**Maysam Mirahmadi** received the B.Sc. and M.Sc. Degree in Electrical Engineering from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2004 and 2007, respectively. From 2007 to 2009, he was a Senior Designer at Kavoshkam R&D Group, Tehran, Iran. Currently he is working toward his Ph.D. degree in Electrical Engineering in the Department of Electrical and Computer Engineering, University of Western Ontario, London, ON, Canada. His current research interests include wireless broadband access networks, quality of service in heterogeneous networks and wireless/optical integrated networks.

**Abdallah Shami** received the B.E. degree in Electrical and Computer Engineering from the Lebanese University, Beirut, Lebanon in 1997, and the Ph.D. Degree in Electrical Engineering from the Graduate School and University Center, City University of New York, New York, NY in September 2002. In September 2002, he joined the Department of Electrical Engineering at Lakehead University, Thunder Bay, ON, Canada as an Assistant Professor. Since July 2004, he has been with The University of Western Ontario, London, ON, Canada where he is currently an Associate Professor in the Department of Electrical and Computer Engineering. His current research interests are in the area of wireless/optical networking.