Dynamic Time Allocation and Wavelength Assignment in Next Generation Multi-rate Multi-wavelength Passive Optical Networks

Jingjing Zhang and Nirwan Ansari
Advanced Networking laboratory
Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102, USA
Email: {jz58, nirwan.ansari}@njit.edu

***Abstract*—Driven by emerging bandwidth-hungry applications, next generation passive optical networks (NG-PONs) provide higher bandwidth to users by using more wavelengths and increasing data rates of optical network units (ONUs). On the other hand, for smooth upgrading, NG-PON is desired to be backward compatible with the current TDM PONs where data rates of ONUs remain unchanged. Thus, both high-rate ONUs and low-rate ONUs may coexist in NG-PON. The key parameters of bandwidth allocation in this multi-rate multi-wavelength network include achieving fairness among all ONUs, encouraging low-rate ONUs to increase their data rates, and utilizing wavelength resources efficiently.**

This paper illustrates contributions in three main aspects. First, we define rate-dependent utilities for ONUs, which serve as the basis for bandwidth arbitration among low-rate and high-rate ONUs. Second, to achieve fairness among ONUs, we employ water-filling idea and formulate a utility max-min fair bandwidth allocation scheme. Third, to efficiently utilize the wavelengths, we map the resource allocation problem in multi-wavelength PON into a multi-processor scheduling problem and employ a heuristic algorithm to address the NP-hard wavelength assignment problem.

I. Introduction

To meet bandwidth requirements of a variety of bandwidth-hungry applications, including multimedia-oriented applications such as HDTV, next-generation passive optical networks (NG-PONs) provision broader bandwidths to optical network units (ONUs) in two main ways: supply more wavelengths and increase data rates of ONUs [1], [2]. For example, in next generation access stage 1 (NGA1) proposed by FSAN, the number of wavelengths can be up to four and the sum data rate in all wavelengths is expected to be 10.3 Gb/s [3]. On the other hand, for backward compatibility, ONUs in current GPON with 1.25 Gb/s are desired be accommodated in NG-PON. Then, NG-PONs will exhibit multi-rate property, where some ONUs upgrade their hardware devices with data rates being increased to 10.3 Gb/s, while some others still use their current hardware devices with data rates remaining at 1.25 Gb/s. Many technological advances in the physical layer foster the realization of multi-rate PONs [4], [5].

Besides the challenges in the physical layer, from the perspective of the MAC layer, time allocation and wavelength assignment schemes are typical challenging issues in multi-rate multi-wavelength PONs. Proper time allocation and wavelength assignment algorithms are desired to *ensure fairness among low-rate and high-rate ONUs, encourage low-rate ONUs to upgrade their devices, and utilize the wavelength resources efficiently.* For high throughput, the bandwidth is favored to be allocated to high-rate ONUs first. Then, high-rate ONUs receive better quality of service (QoS); low-rate ONUs are thus encouraged to upgrade their hardware devices with higher data rates. However, this will result in the starvation of low-rate ONUs. For fairness, all ONUs are expected to experience the same QoS. Formerly, in multi-rate WLAN, proper rate allocation and MAC parameter determination schemes have been proposed to achieve throughput-based fairness and time-based fairness, respectively [6], [7]. Throughput-based fairness tries to achieve equal throughput among users while time-based fairness aims at achieving equal channel occupying time among users in the long term. However, it is improper to treat either throughput-based fairness or time-based fairness as the resource allocation objective in multi-rate PON for the following reasons. Throughput-based fairness can ensure equal QoS among ONUs, but suffer from inefficiency. Time-based fairness improves efficiency, but it allocates equal time duration to ONUs irrespective of their reported traffic, thus potentially resulting in resource over- and under-allocation.

This paper considers utility-based fairness as the objective of time and wavelength allocation in PON. As compared to throughput-based and time-based fairness, utility-based fairness is a more general concept. When utilities are functions of throughput and homogeneous for all ONUs, utility-based fairness is equivalent to throughput-based fairness; when utilities are functions of time and homogeneous for all ONUs, utility-based fairness is equivalent to time-based fairness. In this paper, we set the utility of an ONU as a function of its data rate, its obtained bandwidth in a cycle, and its reported traffic in a cycle. To achieve the same utility, high-rate ONUs require higher bandwidth than low-rate ONUs do so that high-rate ONUs can enjoy higher QoS, and hence encouraging low-rate ONUs to upgrade their hardware devices.

Having defined ONU utilities, we then focus on addressing

978-1-4244-6404-3/10/$26.00 ©2010 IEEE

the utility max-min fair bandwidth allocation issue to achieve utility-based fairness. We investigated two network scenarios: multi-rate single-wavelength PONs and multi-rate multi-wavelength PONs. For the first scenario, we employ water-filling idea to obtain a utility max-min fair resource allocation scheme. For the second scenario, we map the problem into a multi-processor scheduling problem, and employ the first-fit decreasing algorithm to efficiently utilize the wavelength resources. These scheduling schemes as well as utility functions can ensure fairness among ONUs, encourage low-rate ONUs to increase their data rates, and utilize wavelength resources efficiently.

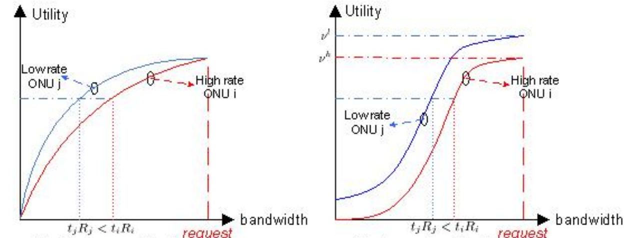
The remaining paper is organized as follows. Section II discusses the MAC control in WDM PON. Section III details the proposed dynamic time allocation and wavelength assignment schemes for multi-rate single-wavelength PON and multi-rate multi-wavelength PON, respectively. Section IV illustrates simulation results. Section V contains the concluding remark.

II. MAC CONTROL IN WDM PON

For backward compatibility, NG-PON inherits some characteristics of the MAC control protocol of current TDM PONs, such as EPON [8] and GPON, two major flavors of TDM PONs. As being specified mainly in IEEE 802.3av [9], 10G-EPON extends Multi-Point Control Protocol (MPCP) as its MAC control protocol. The ITU standard of NG-PON takes two stages: NGA1 and next generation access stage 2 (NGA2). The MAC protocol of NGA1 will mainly comply with that of GPON standardized in ITU-T Recommendation G.984.3. Then, in both 10G-EPON and NGA1, OLT collects queue requests, makes bandwidth allocation decisions, and then notifies ONUs when and on which wavelength channels they can transmit data packets and report queue lengths.

The key function of OLT is to decide when and how to allocate bandwidths [10]. For the problem of deciding when to allocate bandwidth, OLT in GPON determines bandwidth allocation after collecting requests from all ONUs in a cycle; OLT in EPON may allocate bandwidths to ONUs immediately after receiving their requests. However, in this case, the decision is based on only one ONU's request, thus potentially resulting in unfair bandwidth allocation. For fairness, OLT should better wait for the arrivals of more ONUs' requests [10], [11]; this is also the scenario we advocate. Then, in NG-PON, OLT needs to arbitrate bandwidths among multiple ONUs each time it makes decisions. The issue of how to arbitrate and allocate bandwidths is the focus of this paper. More specifically, we focus on addressing the bandwidth arbitration problem in high-load resource allocation cycles.

We refer to the high-load cycle as follows. In GPON, the bandwidth allocation cycle is fixed as 125 μs, whereas the cycle duration in EPON is dynamically adaptive to the incoming traffic with an upper bound of 2 ms. Let $\mathcal{T}$ be the upper bound of the cycle duration, $\mathcal{T} = 125\mu s$ in GPON and $\mathcal{T} = 2ms$ in EPON. Let $\mathcal{W}$ be the number of wavelengths, $\mathcal{N}$ be the number of ONUs with requests, $\mathbf{R} = [R_1, R_2, ..., R_{\mathcal{N}}]'$ be upstream data rates of ONUs, and $\mathbf{q} = [q_1, q_2, ..., q_{\mathcal{N}}]'$ be requests of ONUs. The sum of requested time periods of all ONUs is $\sum_{i=1}^{\mathcal{N}} q_i/R_i$. When $\sum_{i=1}^{\mathcal{N}} q_i/R_i \le \mathcal{W}\mathcal{T}$, all reported ONU requests can be accommodated in a cycle and this cycle is referred to as a low-load cycle; when $\sum_{i=1}^{\mathcal{N}} q_i/R_i > \mathcal{W}\mathcal{T}$, not all the requests can be satisfied, and this cycle is referred to as a high-load cycle.



(a) Concave utility function (b) General utility function

Fig. 1. Two examples of utilities.

III. DYNAMIC BANDWIDTH ALLOCATION IN MULTI-RATE MULTI-WAVELENGTH PONS

In this section, we first design rate-dependent utilities for ONUs, and then detail the bandwidth allocation algorithms in multi-rate single-wavelength PON and multi-rate multi-wavelength PON, respectively.

A. Rate-dependent utilities

Formerly, utilities, as functions of user bandwidth, have been investigated for VoIP, real-time IPTV, and elastic FTP and HTTP applications, respectively. With these utility functions, bandwidth are then arbitrated among queues with different applications [12], [13]. In this paper, we employ rate-dependent utility functions for bandwidth arbitration among ONUs with heterogeneous data rates.

Denote $\mathbf{u} = [u_1, u_2, ..., u_{\mathcal{N}}]'$ as utilities of ONUs. To enable high-rate ONUs receive higher QoS, and thus encourage low-rate ONUs increase their data rates, we set this the following rate-dependent property for utilities. To achieve the same utility, high-rate ONUs require larger bandwidth than low-rate ONUs do if they have the same requests, i.e., when $R_i > R_j$ and $q_i = q_j$, $t_iR_i > t_jR_j$ if $u_i(t_i) = u_j(t_j)$.

Fig. 1 shows two examples of utilities of ONUs with different data rates. In both examples, utilities of low-rate ONUs are above those of high-rate ONUs. Then, to achieve the same utility, low-rate ONUs require smaller bandwidth than high-rate ONUs do.

Utility functions constitute the key to the bandwidth arbitration among ONUs. Many existing algorithms can be interpreted as reversely solving a max-min utility problem with certain utility function settings. For example, when the range of utilities of high-rate ONUs is lower than that of utilities of low-rate ONUs as shown in Fig. 2 (a), to achieve near equal utilities of ONUs, requests of high-rate ONUs are always satisfied first. The algorithm is then equivalent to throughput-first algorithm in multi-rate WLAN. When utilities of high-rate ONUs equal to utilities of low-rate ONUs as shown in Fig. 2

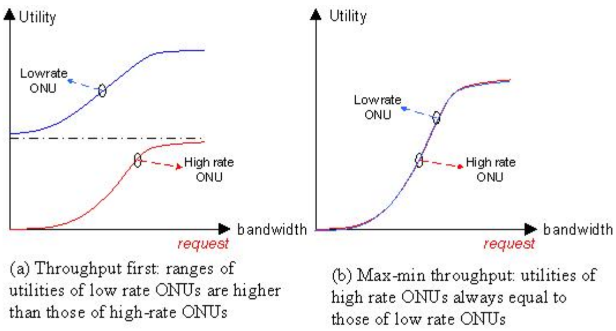


Fig. 2. Utilities for throughput-first algorithm and max-min throughput algorithm, respectively.

(b), to achieve equal utilities of ONUs, the bandwidth should be allocated proportionally to the requested time durations of ONUs. This algorithm is equivalent to max-min throughput fair allocation in multi-rate WLAN.

B. Scheduling of Multi-rate Single-wavelength PON

Here, we discuss the scheduling in multi-rate *single-wavelength* PON. Based on results presented in this section, we will next discuss the scheduling in multi-rate *multi-wavelength* PON in Section III-C.

For fairness, we try to achieve utility max-min fair allocation among ONUs by employing the water-filling idea. More specifically, we first derive the minimum utility among all ONUs, and then equally increase utilities of some ONUs using remaining resources if there exists any. This process is continued until all network resources are consumed.

Denote ϑ as the minimum utility among all ONUs, i.e., $\vartheta = \min_i u_i$. ϑ can be obtained by solving the following problem.

$$\text{maximize: } \nu \quad (1)$$
$$\text{subject to: } u_i \geq \nu, \forall i \quad (2)$$
$$t_i R_i \leq q_i \quad (3)$$
$$\sum_{i}^{N} t_i \leq \mathcal{T} \quad (4)$$

Objective (1) is to maximize the minimum utility among all ONUs. Constraint (2) states that u_i is no less than ν. The method of solving the max-min optimization problem will be discussed in the next section. Constraint (3) limits the bandwidth allocated to ONU i, $t_i R_i$, to be below its request q_i. Constraint (4) states that the total allocated time is no greater than $\mathcal{T}$ to avoid over-exploiting the cycle.

Note that utility function u_i may not be a convex function with respect to t_i. Consequently, the optimization problem may not be solved by convex optimization tools. However, utility u_i of user i is always a non-decreasing function with respect to the time duration t_i allocated to user i. Owing to the non-decreasing property, the decision version of the optimization problem can be easily solved. Hence, the optimization problem with any general utility function can be solved by using the bisection method [13].

When utilities of all ONUs are with the same range, as shown in Fig. 1 (a), the solution obtained for this problem consumes all the resources, and no additional resources can be allocated to improve utilities. When utilities of ONUs are with different ranges, as shown in Fig. 1 (b), the obtained solutions may not consume all the resources, and there may exist some additional resources to further increase utilities of some ONUs. Taken utilities in Fig. 1 (b) for example, let ν^h be the maximum utility of high-rate ONUs, and ν^l be the maximum utility of low-rate ONUs. $\nu^l > \nu^h$. Then, when $\vartheta = \nu^h$, utilities of high-rate ONUs reach the maximum value and cannot be further increased. However, utilities of low-rate ONUs can be increased if there are some remaining resources.

Denote ν_i^m as the upper bound of the utility of ONU i. ONUs are divided into two classes based on the relation between ν^m and ϑ. Without loss of generality, assume ONUs 1, 2, ..., k are with their maximum utilities being greater than ϑ, and the remaining ones are with their maximum utilities being no greater than ϑ. Then, there may have rooms to increase utilities of ONUs $k+1$, ..., $\mathcal{N}$. The problem of maximizing the minimum utility of ONUs $k+1$, ..., $\mathcal{N}$ can be formulated as follows.

$$\begin{aligned} &\text{maximize: } \nu \\ &\text{subject to: } u_i \geq \nu, k+1 \leq i \leq \mathcal{N} \\ &\qquad u_i = \nu_i^m, 1 \leq i \leq k \\ &\qquad t_i R_i \leq q_i \\ &\qquad \sum_{i}^{\mathcal{N}} t_i \leq \mathcal{T} \end{aligned} \quad (5)$$

Constraint (5) is to guarantee the maximum utility for ONUs with $\nu_j^m \leq \vartheta$.

By solving a set of max-min optimization problems, we can finally allocate all resources to ONUs. In the best case, only one max-min optimization problem needs to be solved. It happens when all utilities are of the same range or the minimum utility among all ONUs are small. In the worst case, $\mathcal{N}$ max-min optimization problems need to be solved. It happens when ONUs are with utility ranges different from each other, and there is enough resource to guarantee ν_i^m for every ONU i.

C. Scheduling of Multi-wavelength Multi-rate PON

To arbitrate bandwidth in the multi-wavelength multi-rate PON, our main idea is to first address the multi-wavelength scheduling issue, and then incorporate the above described multi-rate single-wavelength resource allocation algorithm into the scheduling.

Assume each ONU can support all wavelengths, and can only be scheduled by one wavelength in a cycle. All wavelengths are assumed to be available at the same time. The problem can be mapped into a multiprocessor scheduling problem with non-preemptable jobs, which is NP-complete.

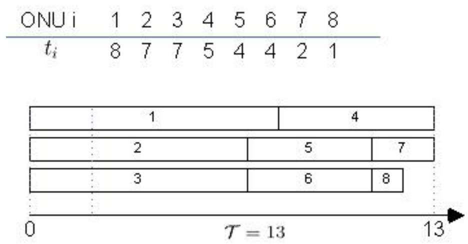


Fig. 3. An instance of the FFD algorithm.

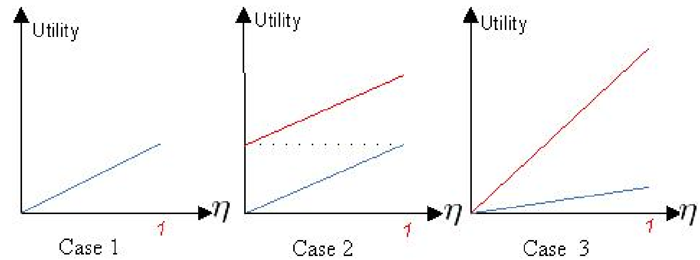


Fig. 4. Three cases of utility functions.

We employ first-fit decreasing (FFD) algorithm [14], as described in Algorithm 1. FFD first sorts ONU requests in the descending order of their sizes. Then, FFD searches from the first machine and tries to assign an ONU request with the first machine on which the ONU request fits.

Algorithm 1 FFD algorithm

```
Sort ONU requests in the descending order of their sizes
for each ONU request do
  k = 1
  while the request is not assigned yet do
    if wavelength k has enough bandwidth for the request
    then
      Assign wavelength k to the request from the earliest
      available time
    else
      k = k + 1
    end if
  end while
end for
```

Fig. 3 illustrates one instance of the FFD algorithm for 8 ONUs. The number of wavelengths is 3, and the upper bound of the cycle duration is 13. For this particular example, all ONU requests can be scheduled within this cycle.

FFD, as a heuristic algorithm to the NP-complete problem, may not find the feasible schedule when there exists one for a certain set of requests. Let $\mathcal{T}^{min}$ be the minimum makespan for these requests. It was proved that FFD can always find a feasible schedule when $\mathcal{T} > 72/61\mathcal{T}^{min}$ [14]. If all ONU requests can be scheduled within $\mathcal{T}$, then we obtain the solution; otherwise, we adjust ONU utilities, and rerun FFD.

IV. Simulations and Discussions

In the above, we have presented a scheme for designing rate-dependent utilities for ONUs, and allocating resources to achieve utility max-min fairness among ONUs. In the simulation part, our main objective is to observe the throughput and delay performances for three particular cases of utility functions. Simulation results are expected to draw the conclusion that certain performances can be achieved by setting utility functions properly. We only consider the single-wavelength scenario in the simulation. The multi-wavelength scenario adopts the same strategy for bandwidth arbitration among low-rate and high-rate ONUs. Owing to the limited space, we do not present the performance of joint wavelength assignment and time allocation scheme under the multi-wavelength scenario; this will be reported in the future.

Assume the PON consists of 32 ONUs, among which 16 ONUs are with data rate of 1.25 Gb/s and another 16 ONUs are with data rate of 10 Gb/s. Consider three cases of simple utility functions as follows. $\eta_i = t_i \cdot R_i/q_i$ ($0 \leq \eta_i \leq 1$) is defined as the unified throughput of ONU i.

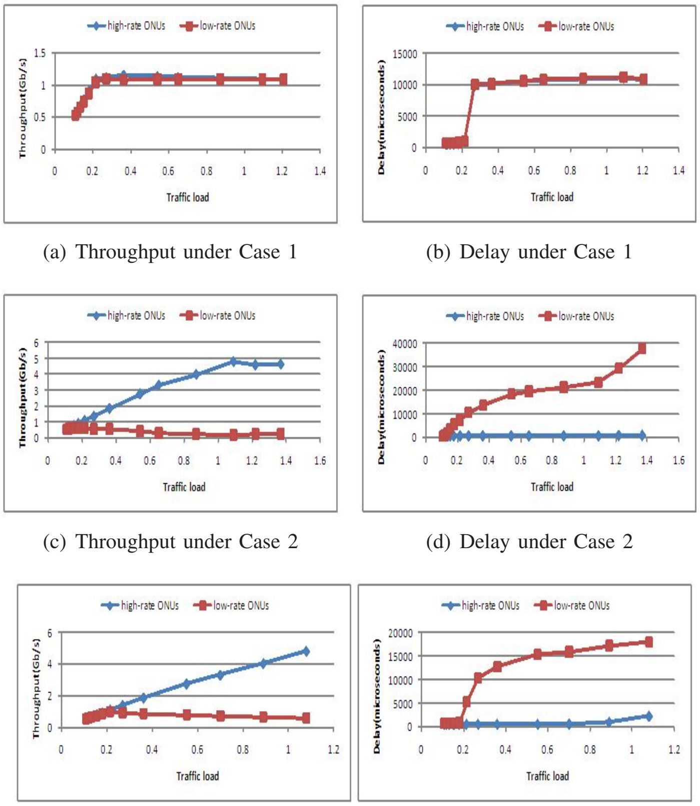
- Case 1: $u_i = \eta_i$ for all ONUs
- Case 2: $u_i = 1 + \eta_i$ for ONUs with rate of 1.25 Gb/s; $u_i = \eta_i$ for ONUs with rate of 10 Gb/s
- Case 3: $u_i = \eta_i/1.25$ for ONUs with 1.25 Gb/s; $u_i = \eta_i/10$ for ONUs with 10 Gb/s;

Fig. 4 illustrate the three cases. In Case 1, to achieve the same utility, low-rate ONUs and high-rate ONUs desire the same unified throughput. Hence, the time is allocated proportionally to the requested time durations of ONUs. In Case 2, utilities of high-rate ONUs are always lower than those of low-rate ONUs, and thus resulting in the higher priority of high-rate ONUs. In Case 3, ONUs are allocated with time durations proportionally to their requests. Intuitively, high-rate ONUs experience the best QoS under Case 2, and the worst QoS under Case 1. Next, we investigate performances of the three algorithms by using the OPNET simulator.

In the simulation, each ONU carries self-similar traffic with the same distribution. The traffic is generated by aggregating a sum of on-off sources with "on" and "off" periods exhibiting Pareto distribution. Denote X as a Pareto-distributed random variable, $P(X > x) = (x/x_m)^{-\alpha}, 1 < \alpha < 2, x > x_m$. The Hurst parameter is $(3-\alpha)/2$. We aggregate five on-off sources with α being set as 1.1 for both "on" and "off" periods, and vary x_m to obtain different traffic loads. With the assumption that all ONUs are with 10 Gb/s, the traffic load is defined as the ratio of the total required time duration of all ONUs over the cycle duration, i.e., $\sum_{i=1}^{32}(q_i/10^{10})/\mathcal{T}$.

Fig. 5 shows the throughput and delay performances of high-rate and low-rate ONUs under the above stated three cases of utility functions.

In Case 1, ONUs are equally treated and allocated with time periods proportionally to their requested time periods. Hence, all ONUs experience similar performances as shown in Fig. 5(a) and 5(b). It is also shown that throughput of ONUs stop increasing and delay of ONUs increases sharply when traffic load increases to around 0.21. This is because the total



(a) Throughput under Case 1 (b) Delay under Case 1

(c) Throughput under Case 2 (d) Delay under Case 2

(e) Throughput under Case 3 (f) Delay under Case 3

Fig. 5. Throughput and delay performances of high-rate and low-rate ONUs under three cases of utility functions.

requested "time'exceeds $\mathcal{T}$ under that traffic load. Assume all ONUs have the same traffic. Then, the time duration of high-rate ONUs is 1/8 of that of low-rate ONUs. Hence, the total requested "time" duration equals to the DBA cycle when the load is around 0.222.

In Case 2, high-rate ONUs are given higher priority over low rate ONUs. Fig. 5(c) and 5(d) show that, for high-rate ONUs, throughput increases linearly with the increase of the traffic load, and delay almost keeps constant when the traffic load is under 1. For low-rate ONUs, throughput drops and delay increases with the increase of the traffic load.

In Case 3, ONUs are granted with the time duration proportional to their requests. When ONUs are with the same traffic demand, they get equal share of the total cycle duration. Under low traffic load that the cycle duration can satisfy requirements from all ONUs, throughput increases with the increase of the traffic load. With the increase of the traffic load, the equal share of the cycle duration will exceed demands of low-rate ONUs first, and thus resulting in the decrease of throughput and increase of delay experienced by low-rate ONUs.

Comparing these three cases, high-rate ONUs experience the best performance under Case 1 and the worst performance under Case 2, and the contrary holds for low-rate ONUs. In addition, the performance in Case 3 always range between those in Case 1 and Case 2. These are consistent with utility functions in these three cases. Hence, utility functions determine performances of ONUs. The desired performances can be obtained by setting utility functions properly.

V. Conclusion

This paper presents a utility based max-min fair resource allocation scheme for next generation multi-rate multi-wavelength PON. To encourage low-rate ONUs to increase their data rates, we define rate-dependent ONU utilities with the property that high-rate ONUs require higher QoS than low-rate ONUs for the same utility. Then, we address the utility max-min fair resource allocation issue by recursively solving a set of max-min optimization problems. By utilizing some properties of utilities, the max-min programming problem can be solved by the bisection method. Besides solving the time allocation problem, we also propose to employ first-fit decreasing to address the NP-complete wavelength assignment issue in multi-wavelength PONs. With our proposed schemes, the resource allocation problem is transformed into the problem of designing utility functions. Simulation results show that the obtained performances generally agrees with those derived from utility functions.

References

[1] R. W. Heron, T. Pfeiffer, D. T. van Veen, J. Smith, and S. S. Patel, "Technology innovations and architecture solutions for the next-generation optical access network," Bell Labs Technical Journal, Tech. Rep. 1, 2008.

[2] J. Zhang and N. Ansari, "Design of WDM PON with tunable lasers: The upstream scenario," *IEEE/OSA Journal of Lightwave Technology*, vol. 28, no. 2, pp. 228–236, 2010.

[3] J. Zhang, N. Ansari, Y. Luo, F. Effenberger, and F. Ye, "Next-generation PONs: a performance investigation of candidate architectures for next-generation access stage 1," *IEEE Communications Magazine*, vol. 47, no. 8, pp. 49–57, August 2009.

[4] C.-J. Chae, "Multi-rate Ethernet PON based on secure optical CSMA/CD," in *28th European Conference on Optical Communication*, vol. 4, Sept. 2002, pp. 1–2.

[5] K. Hara, S. Kimura, H. Nakamura, K. Nishimura, M. Nakamura, N. Yoshimoto, and M. Tsubokawa, "1.25/10.3 Gbit/s dual-rate burst-mode receiver," *Electronics Letters*, vol. 44, no. 14, pp. 869–870, March 2008.

[6] B. Radunovic and J.-Y. L. Boudec, "Rate performance objectives of multihop wireless networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 334–349, 2004.

[7] A. Babu and L. Jacob, "Fairness analysis of ieee 802.11 multirate wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 3073–3088, Sept. 2007.

[8] G. Kramer and G. Pesavento, "Ethernet passive optical network (EPON): building a next-generation optical access network," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 66–73, Feb 2002.

[9] M. Hajduczenia, H. Da Silva, and P. Monteiro, "Development of 10 Gb/s EPON in IEEE 802.3av," *IEEE Communications Magazine*, vol. 46, no. 7, pp. 40–47, July 2008.

[10] M. P. McGarry, M. Reisslein, C. J. Colbourn, M. Maier, F. Aurzada, and M. Scheutzow, "Just-in-time scheduling for multichannel EPONs," *IEEE/OSA Journal of Lightwave Technology*, vol. 26, no. 10, pp. 1204–1216, 2008.

[11] A. R. Dhaini, C. M. Assi, M. Maier, and A. Shami, "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, no. 1, pp. 277–286, 2007.

[12] Z. Cao and E. Zegura, "Utility max-min: an application-oriented bandwidth allocation scheme," *IEEE INFOCOM*, vol. 2, pp. 793–801 vol.2, Mar 1999.

[13] J. Zhang and N. Ansari, "Utility max-min fair resource allocation for diversified applications in epon," in *AccessNets*, Hongkong, China, 2009.

[14] M. Pinedo, *Scheduling: theory, algorithms, and systems*. Prentice Hall, 2002.