
Volume 33, Issue 4, 1 March 2010 ISSN 0140-3664
ELSEVIER
computer
communications
Available online at www.sciencedirect.com
ScienceDirect

Computer Communications 33 (2010) 526–531

Contents lists available at ScienceDirect

Computer Communications

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

Bandwidth allocation with a particle swarm meta-heuristic for ethernet passive optical networks

Un Gi Joo [a], Alice E. Smith [b,*]

[a] Department of Industrial and Management Engineering, Sun Moon University, Asan, Chungnam 336-708, Republic of Korea
[b] Department of Industrial and Systems Engineering, Auburn University, Auburn, AL 36849-5346, USA

A R T I C L E   I N F O

Article history:
Received 25 July 2007
Received in revised form 23 October 2009
Accepted 27 October 2009
Available online 1 November 2009

Keywords:
Ethernet Passive Optical Network
Bandwidth allocation
Throughput
Weighted fairness
Particle Swarm Optimization (PSO)

A B S T R A C T

This paper considers the bandwidth allocation problem for an Ethernet Passive Optical Network (EPON). An EPON is one of the best options for high-speed access networks. This paper formulates the optimal bandwidth allocation problem with an analytical model to maximize throughput and weighted fairness simultaneously. First, the optimal solution under certain conditions is characterized. Then, two heuristic algorithms are devised which optimize the allocation problem under general conditions. One heuristic is a straightforward constructive one while the one use the Particle Swarm Optimization (PSO) meta-heuristic, the first known application of PSO to the EPON bandwidth allocation problem. The heuristics are tested and compared with previously published results. The computational experience shows that our algorithms are both effective and efficient in allocating bandwidth on an EPON.

1. Introduction

This paper considers the bandwidth allocation problem on an Ethernet Passive Optical Network (EPON), where there is one common fiber connecting several access points which share the bandwidth of the fiber. To provide service, two types of passive components are used. One is the Optical Line Terminal (OLT) and the other is the Optical Network Unit (ONU). The OLT provides connection between the backbone and access networks at a central office, and the OLT allocates the common bandwidth according to requested demands from the connected access points. The requests for bandwidth occur from ONUs located on the access side of the network. The ONU has buffer memory for incoming traffic from customers and for outgoing traffic to the OLT, and considers the transmission priority of the packets waiting in the buffers. To transmit the traffic from OLT to ONU, OLT broadcasts the traffic and then each ONU receives the packets. However, channels must be classified for transmission of the packets from ONUs to an OLT since there are many ONUs and there are conflicts when ONUs try to transmit packets simultaneously. EPON uses the Time Division Multiple Access (TDMA) method to resolve conflicts. The bandwidth allocation problem is to find a mechanism for sharing the common bandwidth without conflict.

Many studies have been published on how to allocate the common bandwidth to the ONUs. For example, Kramer et al. [1] and Bai et al. [2] developed bandwidth allocation algorithms of interleaved polling with adaptive cycle time (IPACT), and Kramer et al. [3] developed a two-stage buffer mechanism to reduce the light-penalty problem of IPACT. Assi et al. [4] suggested an Enhanced Dynamic Bandwidth Allocation (EDBA) algorithm. An et al. [5] developed a Hybrid Slot-Size/Rate (HSSR) algorithm, and Yang et al. [6] presented a burst polling algorithm for bandwidth allocation. A good survey on these methods is given by Zheng and Mouftah [7]. These papers all address finding good heuristic solutions with respect to some specific measure such as throughput, fairness or delay time. For example, Bai et al. [2] developed a weighted-based bandwidth algorithm (W-DBA) to improve fairness, delay time and link utilization. W-DBA assigns the excess bandwidth to each ONU proportionally to each relative weight. The authors showed that W-DBA is superior to another published algorithm (M-DBA) for randomly generated asymmetric traffic.

Our paper formulates the allocation problem as a nonlinear mathematical one and develops bandwidth allocation algorithms which maximize throughput and weighted fairness simultaneously. Characterizing the optimal allocation under certain conditions is one goal of this paper. The other goal is to develop an effective optimizer to the problem under general conditions. This paper is structured as follows: Section 2 describes and formulates the bandwidth allocation problem. The bandwidth allocation problem is characterized, and two heuristic algorithms are developed in

* Corresponding author. Tel.: +1 334 844 1400.
E-mail addresses: ugjoo@sunmoon.ac.kr (U.G. Joo), smithae@auburn.edu (A.E. Smith).

0140-3664/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2009.10.019

Section 3. These algorithms guarantee an optimal solution under a specific condition derived in Section 3. Section 4 carries out a number of numerical tests to evaluate the proposed algorithms by considering throughput, fairness, and computation time. Finally, Section 5 contains concluding remarks.

2. Problem description

This section defines the bandwidth allocation problem in detail. Let us define some notation for the formulation.

N: number of customers (ONUs).
r_i: demand of ONU i measured in bits per second, $i = 1, 2, \ldots, N$.
w_i: weight of ONU i for the fairness metric.
C: capacity of fiber (bits per second).
X_i: proportion of allocated bandwidth to demand of customer i, $0 \leqslant X_i \leqslant 1$ for all i.

According to the Multi-Point Control Protocol (MPCP) of EPON [8], the OLT allocates bandwidth and determines time of transmission to N ONUs according to their requests $\{r_i\}$. The allocation is performed periodically during each cycle time, and each ONU transmits the packets waiting in its buffer forward to the OLT as a predetermined amount and time of transmission along the common link. To transmit packets, setup is required between each transmission for laser on/off and guarding information, but this is independent of the ONU type and size of transmission and can be ignored. The time duration to transmit a unit packet also depends on the distance between the OLT and the ONU. However, we can set the unit time to 1 for all ONUs for simplicity without loss of generality.

Due to the capacity restriction C on the common link, only some waiting packets in each buffer will be serviced at the current cycle. The remaining will be serviced during succeeding cycles. Each ONU has one chance to transmit its packets at each cycle. Therefore, we can model the problem by considering just one cycle. There are several classes of service (CoS) packets in an ONU. Each ONU i has a weight w_i to represent its CoS according to the Service Level Agreement (SLA). EPON is required to utilize its capacity as much as possible and to provide fair service. Our paper considers the bandwidth allocation problem to maximize utilization (throughput) and weighted fairness simultaneously.

The bandwidth allocation problem is formulated as following:

$$\max \left(\sum_{i=1}^{N} r_i X_i\right)\left(\sum_{i=1}^{N} X_i \Big/ w_i\right)^2 \Big/ \left[N \cdot \sum_{i=1}^{N} (X_i/w_i)^2\right] \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} r_i X_i \leqslant C \tag{2}$$

$$0 \leqslant X_i \leqslant 1, \quad i = 1, 2, \ldots, N \tag{3}$$

Our aim is to find the optimal $\{X_i\}$ which maximizes the objective function value of Eq. (1) subject to Eqs. (2) and (3). The first term of Eq. (1) represents throughput, and the remaining terms represent fairness, where the fairness is a weighted variant of Jain's equal-weight version [9]. To maximize throughput, $\sum_{i=1}^{N} r_i X_i$, we need to increase the value of X_i as much as possible. If we set the value of X_i/w_i as evenly as possible, then the fairness term $\left(\sum_{i=1}^{N} X_i \Big/ w_i\right)^2 \Big/ \left[N \cdot \sum_{i=1}^{N} (X_i/w_i)^2\right]$ becomes close to its maximum value of 1. Eq. (2) implies that the total amount of assigned bandwidth cannot be larger than the capacity C. The decision variable X_i represents the proportion of demand allocated by OLT. We can obtain the allocated bandwidth $r_i X_i$ for ONU i by solving the bandwidth allocation problem.

3. Bandwidth allocation algorithms

This section formulates the optimal allocation to maximize throughput and fairness. We first consider the problem characterization under certain situations.

3.1. Problem characterization

First, as a special case, if all weights of ONUs are identical, the optimal allocation is determined as defined in Proposition 1.

Proposition 1. *If $w_i = w_0$ for all i, then optimal allocation is $X_i = \min\left\{C \Big/ \sum_{i=1}^{N} r_i, 1\right\}$ for all i and its optimal objective value is $\min\left\{\sum_{i=1}^{N} r_i, C\right\}$.*

Proof. If $w_i = w_0$ for all i, throughput is maximized when $X_i = 1$ for all i and fairness is maximized when $X_i = X_0$ for all i, where X_0 denotes a constant. We can easily show the optimal value of X_0 is $\min\left\{C \Big/ \sum_{i=1}^{N} r_i, 1\right\}$ and its objective value is $\min\left\{\sum_{i=1}^{N} r_i, C\right\}$. This completes the proof. □

If the weights of all ONUs are identical, we can easily obtain the optimal solution such that the objective value is C when $X_i = C \Big/ \sum_{i=1}^{N} r_i$ for all i if $\sum_{i=1}^{N} r_i \geqslant C$, and is $\sum_{i=1}^{N} r_i$ when $X_i = 1$ for all i otherwise.

However, if the weights are not identical, it is not straightforward to find the optimal solution. For the remainder of the paper, we will consider only the general weight problem. For the general problem, let $Y_o = \min\left\{C \Big/ \sum_{i=1}^{N} w_i r_i, \sum_{i=1}^{N} r_i \Big/ \sum_{i=1}^{N} w_i r_i\right\}$ be the total amount of required bandwidth divided by the total weighted amount of required bandwidth, where Y_o is determined by the requested bandwidth $\{r_i\}$. The relationship among Y_o, $\sum_{i=1}^{N} r_i$ and C is derived as follows.

Lemma 1. *If $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$, then $\sum_{i=1}^{N} r_i \geqslant C$.*

Proof. It is noticed that $Y_o \geqslant [\min_{1 \leqslant i \leqslant N}\{1/w_i\}]\left[\min\left\{C \Big/ \sum_{i=1}^{N} r_i, 1\right\}\right]$. Therefore, if $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$, then $C \Big/ \sum_{i=1}^{N} r_i \leqslant 1$. This completes the proof. □

According to Lemma 1, it is true that if $\sum_{i=1}^{N} r_i < C$ then $Y_o > \min_{1 \leqslant i \leqslant N}\{1/w_i\}$. However, the inverse of Lemma 1 is not true. For example, consider an EPON system with link capacity of C. Suppose that there are two ONUs with demand of $(r_1, r_2) = (100, 60)$. Assume that their weights are $(w_1, w_2) = (1, 2)$. Then, $Y_o = C/220 > \min_{1 \leqslant i \leqslant N}\{1/w_i\} = 0.5$ and $\sum_{i=1}^{N} r_i = 160 \geqslant C$ if the capacity C has a value of $110 \leqslant C \leqslant 160$. Therefore, we use the relationship $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$ instead of $\sum_{i=1}^{N} r_i \geqslant C$ for formulations of the optimal solution.

Proposition 2. *If $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$, the optimal allocation is obtained as $X_i = w_i Y_o$ and has the optimal objective value of C.*

Proof. By Lemma 1, if $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$ then $\sum_{i=1}^{N} r_i \geqslant C$. Therefore, the throughput term of Eq. (1) has maximal value C at $\left\{X_i \Big| \sum_{i=1}^{N} r_i X_i = C\right\}$ and the remaining terms of Eq. (1) have the maximal value of 1 when $X_i/w_i = Y_1$ for all i, where Y_1 is a constant. For the solution of $\sum_{i=1}^{N} r_i X_i = C$, let us substitute $w_i Y_1$ for X_i and simplify the equation; then we obtain the result of $Y_1 = C \Big/ \sum_{i=1}^{N} w_i r_i$ which equals Y_0 by Lemma 1. The solution $\{X_i\}$ also satisfies Eq. (3) since $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$. For the solution of $X_i = w_i Y_o$, the objective value becomes C since $\sum_{i=1}^{N} r_i \geqslant C$ by Lemma 1. This completes the proof. □

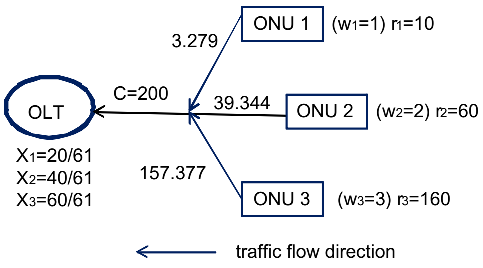

Fig. 1. Example EPON system.

For example, consider an EPON system with $C = 200$ (Fig. 1). Suppose that there are three ONUs with demand of $(r_1, r_2, r_3) = (10, 60, 160)$ at the beginning of a cycle. Assume that their weights are $(w_1, w_2, w_3) = (1, 2, 3)$. Then, $Y_o = \min\left\{C\big/\sum_{i=1}^{N} w_i r_i, \sum_{i=1}^{N} r_i\big/\sum_{i=1}^{N} w_i\ r_i\right\} = \min\{200/610, 230/610\} = 200/610 < \min_{1\leqslant i\leqslant N}\{1/w_i\} = 1/3$ and thus $\sum_{i=1}^{N} r_i = 230 > C = 200$. Therefore, the optimal allocation is obtained as $X_1 = w_1 Y_o = 200/610, X_2 = w_1 Y_o =$ s(2)(200/610) $= 400/610$, and $X_3 = 600/610$ by Proposition 2.

However, if $Y_o > \min_{1\leqslant i\leqslant N}\{1/w_i\}$, a global search mechanism is required to solve the problem. For example, consider an EPON system with $C = 200$. Suppose that there are three ONUs with demand of $(r_1, r_2, r_3) = (100, 60, 80)$ so that $\sum_{i=1}^{N} r_i = 240 > C$. Assume that their weights are $(w_1, w_2, w_3) = (1, 2, 3)$. Then, $Y_o = \min\left\{C\big/\sum_{i=1}^{N} w_i r_i, \sum_{i=1}^{N} r_i\big/\sum_{i=1}^{N} w_i r_i\right\} = \min\{200/460, 240/460\} = 200/460 > \min_{1\leqslant i\leqslant N}\{1/w_i\} = 1/3$. In this case, the result of Proposition 2 such that $X_3 = 600/460$ does not guarantee feasibility.

3.2. Solution algorithms

Eq. (1) is not concave with respect to decision variable X_i. Therefore, we need an approach which can accommodate this. This section develops two heuristic algorithms. First, we develop a heuristic algorithm based upon the formulation of the optimal solution. Considering each objective individually, there are two feasible solutions with maximum throughput or maximum fairness, denoted as S1 and S2:

$$\text{Solution S1:}\quad X_i = w_i\left(\min\left\{Y_o, \min_{1\leqslant i\leqslant N}\{1/w_i\}\right\}\right) \tag{4}$$

$$\text{Solution S2:}\quad X_i = \min\left\{C\Big/\sum_{i=1}^{N} r_i, 1\right\} \text{ for all } i \tag{5}$$

Since solutions S1 and S2 are feasible as shown by Proposition 3, the better one is a lower bound to our problem. We now devise a heuristic algorithm called H1.

Heuristic algorithm H1: Select the solution with the higher objective value between S1 and S2.

Proposition 3. *Algorithm H1 guarantees an optimal solution for the bandwidth allocation problem satisfying the relationship of* $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$.

Proof. To show the feasibility of S1, we can classify the problem into two classes of $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$ and $Y_o > \min_{1\leqslant i\leqslant N}\{1/w_i\}$. First, if $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$, S1 has the value of $X_i = w_i Y_o$. The solution X_i satisfies $0 \leqslant X_i \leqslant 1$ for all i and $\sum_{i=1}^{N} r_i X_i \leqslant C$ since $Y_o = \min\left\{C\big/\sum_{i=1}^{N} w_i r_i, \sum_{i=1}^{N} r_i\big/\sum_{i=1}^{N} w_i r_i\right\}$ by definition. Second, if $Y_o > \min_{1\leqslant i\leqslant N}\{1/w_i\}$, then the resulting solution of S1 becomes $X_i = w_i(\min_{1\leqslant i\leqslant N}\{1/w_i\})$ which satisfies the relationship of $0 \leqslant X_i \leqslant 1$ for all i, and $\sum_{i=1}^{N} r_i X_i = \min_{1\leqslant i\leqslant N}\{1/w_i\}\sum_{i=1}^{N} w_i r_i < Y_o\left(\sum_{i=1}^{N} w_i r_i\right) = \min\left\{C, \sum_{i=1}^{N} r_i\right\} \leqslant C$. Similarly, we can show the feasibility of S2 by classifying the problem into two cases of $C\big/\sum_{i=1}^{N} r_i \leqslant 1$ and $C\big/\sum_{i=1}^{N} r_i > 1$. If $C\big/\sum_{i=1}^{N} r_i \leqslant 1$, the solution becomes $X_i = C\big/\sum_{i=1}^{N} r_i$ for all i which has the relationship of $0 \leqslant X_i \leqslant 1$ for all i, and $\sum_{i=1}^{N} r_i X_i = C$. Otherwise, $X_i = 1$ for all i, and the solution leads to the relationship of $\sum_{i=1}^{N} r_i X_i = \sum_{i=1}^{N} r_i \leqslant C$ by the precondition of this case.

To consider optimality, note that H1 always selects S1 as its solution since solution S1 has the objective value of $\min\left\{\sum_{i=1}^{N} r_i, C\right\}$ when $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$. If $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$, the solution S1 of $X_i = w_i Y_o$ is the optimal solution by Proposition 2. This completes the proof. □

However, as an example for H1, consider two ONUs with $(r_1, r_2) = (100, 60)$ and $(w_1, w_2) = (1, 2)$ on a link with capacity C=150. Then, $Y_o = \min\{150/220, 160/220\} > \min_{1\leqslant i\leqslant N}\{1/w_i\} = 0.5$, and solution S1 is $X_i = w_i(\min\{Y_o, \min_{1\leqslant i\leqslant N}\{1/w_i\}\})$; $(X_1, X_2) = (0.5, 1)$ with its objective function value 110, and solution S2 is $X_i = \min\left\{C\big/\sum_{i=1}^{N} r_i, 1\right\} = 150/160$ for all i and has objective value of 130.95. Therefore, the solution of H1 is S2, which has the larger objective function value. But we cannot guarantee optimality of solution S2 since the problem does not satisfy the condition $Y_o \leqslant \min_{1\leqslant i\leqslant N}\{1/w_i\}$ of Proposition 3, and there are other better solutions such as $(X_1, X_2) = (0.9, 1)$ which has the objective function value of 138.921.

To address the shortcomings of H1, we develop another heuristic algorithm using solutions S1 and S2. This paper uses Particle Swarm Optimization (PSO) for the second heuristic algorithm since PSO is an effective and computationally speedy global optimization technique for multi-modal optimization problems in the real number domain [10,11]. The PSO is a global optimization heuristic inspired by nature, in this case by bird flocking and fishing schooling. The unique aspect of this heuristic is that it balances individual (particle) knowledge with group (swarm) knowledge. The PSO use a population called a *swarm* which is composed of multiple *particles*. The PSO conducts search by moving the particles according to a velocity vector which establishes the search gradient for each particle. This vector is set by individual knowledge (where the particle's best location has been) and group knowledge (where the swarm's best location has been) using the cognition and social coefficients, respectively. The best solution found by the swarm at the end of the search iterations is returned as the "optimal" solution. This paper uses M particles $\left\{X^1, X^2, \ldots, X^j, \ldots, X^M\right\}$, where each particle X^j is composed of N variables $X^j = \left(X_1^j, X_2^j, \ldots, X_i^j, \ldots, X_N^j\right)$, $j = 1, 2, \ldots, M$. For our problem, X_i^j denotes the allocated proportion of bandwidth for the jth particle of ONU i.

Heuristic algorithm H2

Step 1. Set initial values of the parameters including particle swarm size, inertia weight (w), cognition and social weights ($C1, C2$), and maximum iterations ($I_{\max}$) of the PSO. Read the problem data of the bandwidth requirement r_i and weight w_i for each ONU i, $i = 1, 2, \ldots, N$. Randomly generate M particles $\{X^j\}$ as an initial population, $j = 1, 2, \ldots, M$ where each particle j has N terms. Let $k = 1$.

Step 2. For each particle j, calculate its fitness F_j using Eq. (7) and find the personal best particle for each, $pBest_i^j$, $i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, M$, that is, the best solution that a given particle has identified over the search. For the first iteration, the personal best is the initial solution of particle j.

Step 3. Find global best particle, $gBest_i$, $i = 1, 2, \ldots, N$ over all M personal best particles.

Step 4. Update the allocated proportion X_i^j of each particle such that

$$X_i^j = X_i^j + V_i^j, \quad i = 1, 2, \ldots, N; \; j = 1, 2, \ldots, M$$

where $V_i^j = w * V_i^j + C1(R1)\left(pBest_i^j - X_i^j\right) + C2(R2)\left(gBest_i - X_i^j\right)$ for given parameters w, $C1$ and $C2$, and random numbers $R1$ and $R2$ generated from the uniform distribution on domain [0,1]. Let the velocity V_i^j have $V_{\min} \leqslant V_i^j \leqslant V_{\max}$ by setting $V_i^j = V_{\min}$ if $V_i^j < V_{\min}$, and $V_i^j = V_{\max}$ if $V_i^j > V_{\max}$. Additionally, let $X_i^j = X_{\min}$ if $X_i^j < X_{\min}$, and $X_i^j = X_{\max}$ if $X_i^j > X_{\max}$.

Step 5. Repeat Steps 2–4 by letting $k = k + 1$ until $k > I_{\max}$.

Step 6. Find the best feasible solution (satisfying Eqs. (2) and (3)). If the solution of Step 3 is feasible, then select it as the best solution. Otherwise, take the last $gBest_i$ which is feasible.

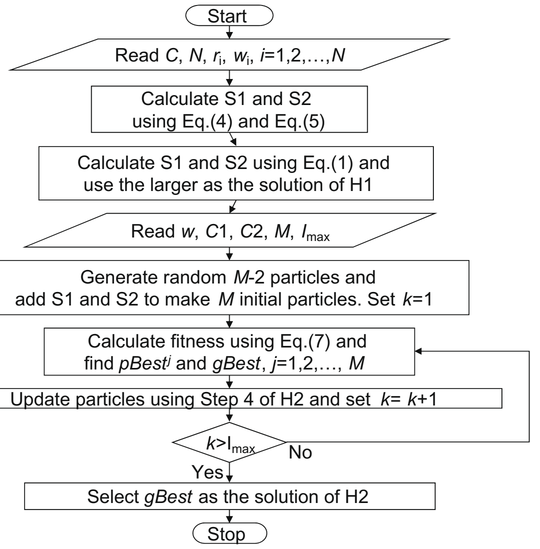

Fig. 2. Flowchart of H2.

For Step 1, this paper uses the values in Table 1 for the parameters. Dong et al. [12] recommended using 20–40 particles for composition a swarm. Our test results show that the value of Eq. (1) is only slightly affected by the swarm size so we use 20 particles to reduce computational time. We select solutions S1 and S2 of Eqs. (4) and (5) as two particles of the initial swarm (termed "seeding" the swarm). Since algorithm H2 uses S1 and S2 as particles, H2 also guarantees an optimal allocation when $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$ by Proposition 3. This paper uses $C1 = C2 = 2$ for Step 4. The inertia weight w is known to have a role in the performance of PSO. Elbeltagi et al. [10] used a linear decreasing function on the domain of [0.4, 1.2] which we use also:

$$w = W_{\max} - (W_{\max} - W_{\min})(ITER)/I_{\max} \tag{6}$$

where $W_{\max} = 1.2$, $W_{\min} = 0.4$, and $ITER$ represents the number of iterations performed. The maximum number of iterations denoted as $I_{\max}$ is 500 according to the result of [12] although our results show that the value of $I_{\max}$ has only a slight effect on Eq. (1).

Some solutions of Step 1 may not be feasible. We remedy this by the following three steps. First, we let $X_i^j = X_{\min}$ if $X_i^j < X_{\min}$ and $X_i^j = X_{\max}$ if $X_i^j > X_{\max}$ at Step 4, where $X_{\min} = 0$ and $X_{\max} = 1$ since the decision variables $\{X_i\}$ must satisfy Eq. (3). Second, we let $V_i^j = V_{\min}$ if $V_i^j < V_{\min}$, and $V_i^j = V_{\max}$ if $V_i^j > V_{\max}$ at Step 4, where we set $V_{\min} = -1$ and $V_{\max} = 1$. Third, we let the fitness function F_j of Step 2 have a penalty term for infeasible solutions as follows:

$$F_j = \left(\sum_{i=1}^{N} r_i X_i^j\right)\left(\sum_{i=1}^{N} X_i^j \Big/ w_i\right)^2 \Big/ \left[N \cdot \sum_{i=1}^{N}\left(X_i^j/w_i\right)^2\right] - \max\left\{0, \sum_{i=1}^{N} r_i X_i^j - C\right\} \tag{7}$$

Yiqing et al. [13] noted that the performance of a penalty function method was satisfactory when an optimization problem is not highly constrained, which is true here. However, even though these steps are used to find a feasible solution, the resultant solutions may not be feasible. So, Step 6 is added to select a best feasible solution as the final PSO solution. The overall procedure of H2 is given in Fig. 2.

Table 1
Parameters for the computational experiments.

Parameter	Value
Number of particles (M)	20 (including S1 and S2)
$C1$, $C2$	2
Inertia weight (w)	$W_{\max} - (W_{\max} - W_{\min})(ITER)/I_{\max}$
$W_{\min}$, $W_{\max}$	$W_{\min} = 0.4$, $W_{\max} = 1.2$
Number of iterations ($I_{\max}$)	500
$V_{\min}$, $V_{\max}$	$V_{\min} = -1$, $V_{\max} = 1$
$X_{\min}$, $X_{\max}$	$X_{\min} = 0$, $X_{\max} = 1$
Number of ONUs (N)	16, 32, 64
Link capacity (C)	1000 Mbps
Bandwidth requirement of each ONU (r_i)	N(mean, 0.0019^2)
Offered load	10–100% (symmetric load) in increments of 10% 60–100% (asymmetric load) in increments of 5%
Weight of each ONU (w_i)	U[1,6]
Number of problems	20

4. Computational experience

For the evaluation of H1 and H2, we consider an EPON system with 32 ONUs using a common fiber of 1000 Mbps speed as described in the lower part of Table 1. This problem is taken from Bai et al. [2]. To consider different numbers of ONUs, we then developed problems with N = 16 and 64.

We assume that the incoming traffic on the EPON system is composed of 20% of Expedited Forwarding (EF), 40% of Assured Forwarding (AF), and 40% of Best Effort (BE) traffic, where the size of the EF traffic is 70 bytes, and each size of the AF and BE traffic is uniform 64–1518 bytes [2]. If we suppose that the size of a packet is distributed according to a uniform distribution and there are many packets, we can use the normal distribution with a standard deviation 0.0019 Mbps as an approximate distribution of input traffic. We consider both symmetric and asymmetric traffic. Symmetric traffic consists of generating balanced traffic in each ONU for an offered load of 10–100%. Asymmetric traffic is balanced except for one ONU, where that ONU makes an offered load of 60–100% and the other N-1 ONUs maintain traffic of 60% load. Each ONU has its class of service according to a SLA. This paper considers six classes of service [14]. The classes are randomly generated from a discrete uniform (DU) distribution on the domain of [1,6]. For each combination of parameters, we devised 20 randomly generated problems having $Y_o > \min_{1 \leqslant i \leqslant N}\{1/w_i\}$ and solved them with H1, H2 and the method of Bai et al. [2] using code in Microsoft visual studio C++ on a PC (Intel Pentium 1.4 GHz). Bai et al. [2] considered only asymmetric traffic with their algorithm W-DBA and use a weight w_i for each ONU i.Since W-DBA is a superior performing method, we use it for computational comparison. We select mean deviation as the performance measure where mean deviation is the proportion of improvement in the objective function value equation (1) over the H1 solution. Since H2 tries to improve the result of H1, a larger mean deviation implies a more improved solution. However, the mean deviation will be zero if the problem instance being considered satisfies the relationship $Y_o \leqslant \min_{1 \leqslant i \leqslant N}\{1/w_i\}$ since both H1 and H2 guarantee an optimal solution in this case.

The test results are given in Figs. 3–6, where Figs. 3 and 4 show the mean deviation of H2 and W-DBA from H1. The figures show the deviation on the y-axis and the percent offered load on the x-axis. Fig. 4 shows that the mean deviation of H2 does not decrease with the offered load and H2 improves the result of H1 up to about 4%. For asymmetric traffic, the improvement becomes larger as the offered load increases. Therefore, the PSO of H2 is effective in allocating bandwidth and improves upon the single objective H1.

When we consider the results from [2], H1 is much better than that of Bai et al.'s W-DBA for all data sets except for symmetric traffic with 100% offered load (31.25 Mbps). This can be verified in Figs. 5 and 6, where we calculated the throughput and fairness value part of Eq. (1) for varying percentages of maximum offered load. Each point in Figs. 5 and 6 display a pair of fairness and throughput values for a given offered load. The point for each optimization method closest to the x-axis (that is, the smallest y value) is for the lowest offered load. Since our objective is to maximize both fairness and throughput, the ideal point of these graphs lies at the upper right corner (maximum fairness and maximum throughput). For asymmetric traffic, H2 dominates W-DBA. For symmetric traffic, this is true except for the maximum offered load. W-DBA tries to improve fairness by assigning the excess bandwidth to overloaded ONUs as fairly as possible, thus the algorithm does well on asymmetric traffic with a small number of extraordinary (overloaded) ONUs. However, our heuristics assign bandwidth fairly for all ONUs including under loaded ones. Furthermore, for symmetric traffic W-DBA leads to more overloaded ONUs because W-DBA can assign bandwidth to an ONU exceeding its requirement while our heuristics do not assign more bandwidth than an ONU requires. Thus, the fairness of W-DBA decreases as the offered load increases since the number of overloaded ONUs increases. However, after a point, the fairness increases again with load because the amount of excess bandwidth is much smaller. The inflexion point at 50% (15.625 Mbps) offered load in Figs. 4 and 6 shows the phenomenon of W-DBA for symmetric traffic (which was not studied in the original paper [2]). As expected in both graphs, H2 has less fairness than H1 (which maximizes fairness) however, H2's throughput is higher.

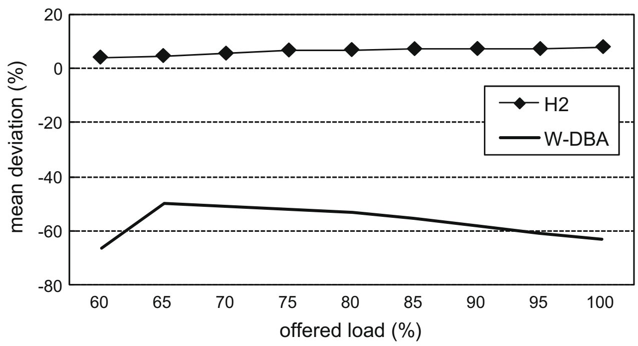

Fig. 3. Mean deviation for asymmetric traffic.

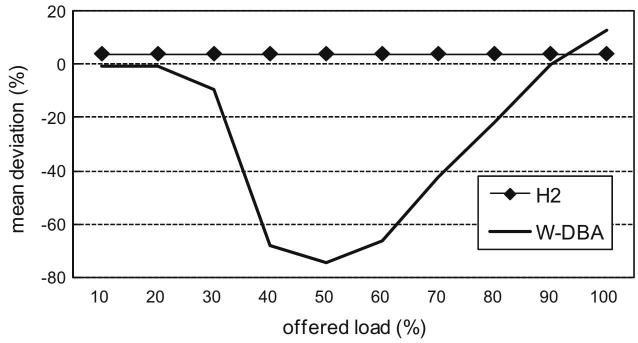

Fig. 4. Mean deviation for symmetric traffic.

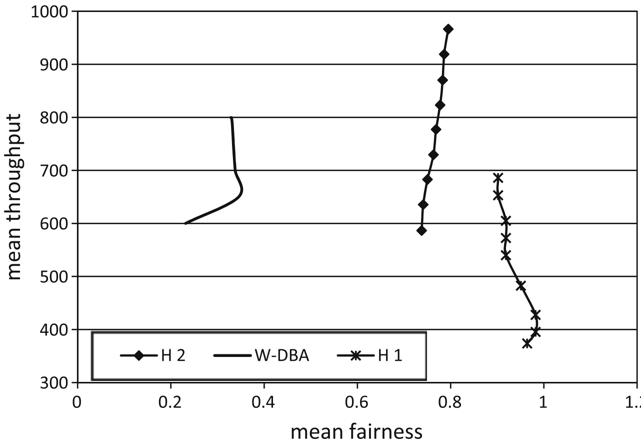

Fig. 5. Mean throughput and fairness for asymmetric traffic at different offered load amounts where the point closest to the x-axis for each series is the lowest offered load.

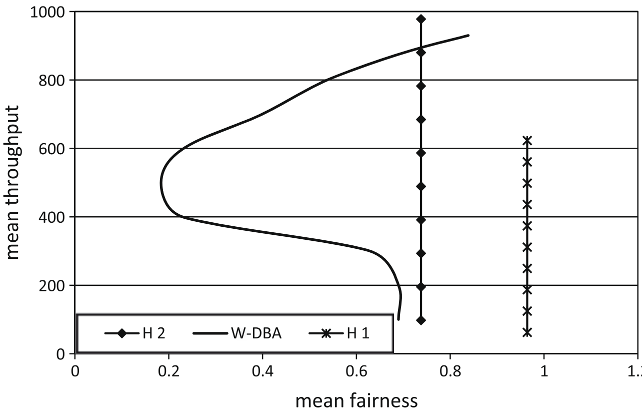

Fig. 6. Mean throughput and fairness for symmetric traffic at different offered load amounts where the point closest to the x-axis for each series is the lowest offered load.

We next considered the effect of number of ONU's. Figs. 7 and 8 show the performance of H2 relative to H1 according to the number of ONUs, $N = 16, 32, 64$. In all cases H2 improved over H1 and the improvements were especially significant as the offered load increased for asymmetric traffic.

For computational effort, it is difficult to make a definitive assessment. Certainly, H2, an iterative meta-heuristic, takes more CPU time than either H1 or the W-DBA algorithm. H2 ranged from 0.119 CPU seconds to 0.126 CPU seconds for each problem with

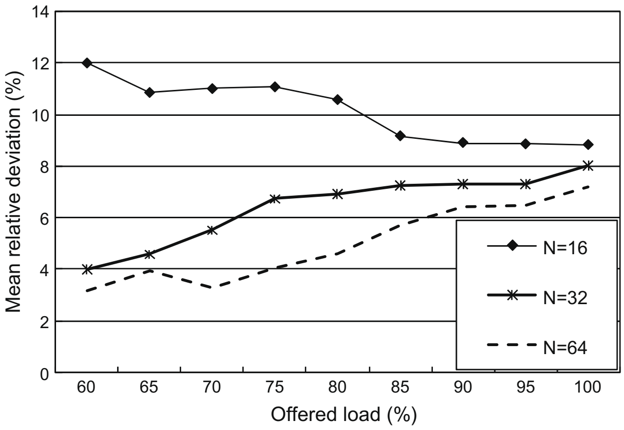

Fig. 7. Mean relative deviation of H2 compared with H1 on asymmetric traffic where the point closest to the *x*-axis for each series is the lowest offered load.

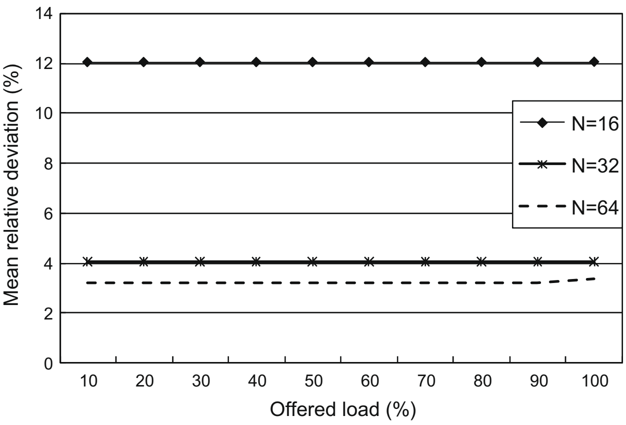

Fig. 8. Mean relative deviation of H2 compared with H1 on symmetric traffic where the point closest to the *x*-axis for each series is the lowest offered load.

$N = 32$ on an Intel Pentium 1.4 GHz PC. So, H2 is very fast and would be viable in the near real time world of EPON bandwidth allocation.

5. Conclusions

This paper considers an EPON system which uses a fiber for each direction of traffic. An efficient bandwidth allocation method is required to use the common fiber. This paper formulates the bandwidth allocation problem to maximize throughput and weighted fairness simultaneously. The optimal allocation is characterized, and two heuristic algorithms are designed based upon the characterization. The first algorithm is a very fast heuristic one, which is optimal under certain restricted conditions, and the second one is a PSO meta-heuristic. This is the first known application of PSO to the EPON bandwidth allocation problem. We test both heuristics and compare to recent results from the literature. The results show that the heuristic algorithms perform well over a range of problem instances and types (symmetric versus asymmetric and different load values and varying numbers of ONU's). The PSO nearly dominates a standard method from the literature (W-DBA from [2]).

Even though this paper considers the multiplicative objective function form of throughput and fairness, other forms such as throughput plus fairness could be also solved with our algorithms by modifying Eq. (7) of the PSO. As another avenue of research, for a given amount of allocated bandwidth to each ONU, the sequencing rule of Weighted Shortest Processing Time (WSPT) could be used to reduce delay time. But, bandwidth allocation and scheduling of packets need to be considered simultaneously, which is complicated.

References

[1] G. Kramer, B.G. Mukherjee, G. Pesavento, IPACT: a dynamic protocol for an Ethernet PON (E-PON), IEEE Communications Magazine (2002) 74–80.
[2] X. Bai, A. Shami, C. Assi, On the fairness of dynamic bandwidth allocation schemes in Ethernet passive optical networks, Computer Communications 29 (2006) 2123–2135.
[3] G. Kramer, B. Mukherjee, S. Dixit, Y. Ye, R. Hirth, Supporting differentiated classes of service in Ethernet passive optical networks, Journal of Optical Networking 1 (2002) 280–298.
[4] C.M. Assi, Y. Ye, S. Dixit, M.A. Ali, Dynamic bandwidth allocation for quality-of-service over Ethernet PONs, IEEE Journal on Selected Areas in Communications 21 (2003) 1467–1477.
[5] F.-T. An, H. Bae, Y.-L. Hsueh, K.S. Kim, M.S. Rogge, L.G. Kazovsky, A new media access control protocol guaranteeing fairness among users in Ethernet-based passive optical networks, IEEE Optical Fiber Communications 11 (2003) 134–136.
[6] Y.-M. Yang, J.-M. Nho, N.P. Mahalik, K. Kim, B.-H. Ahn, A traffic-class burst-polling based delta DBA scheme for QoS in distributed EPONs, Computer Standards & Interfaces 28 (2006) 721–736.
[7] J. Zheng, H.T. Mouftah, Media access control for Ethernet passive optical networks: an overview, IEEE Communications Magazine (2005) 145–150.
[8] G. Kramer, Ethernet Passive Optical Networks, McGraw-Hill, New York, 2005.
[9] R. Jain, A. Durresi, G. Babic, Throughput fairness index: an explanation, in: ATM Forum, 1999. Available from: <http//www.cse.ohio-state.edu/~jain/atmf/a99-0045.html>.
[10] E. Elbeltagi, T. Hegazy, D. Grierson, Comparison among five evolutionary-based optimization algorithms, Advanced Engineering Informatics 19 (2005) 43–53.
[11] F. van den Bergh, A.P. Engelbrecht, A study of particle swarm optimization particle trajectories, Information Science 17 (2006) 937–971.
[12] Y. Dong, J. Tang, B. Xu, D. Wang, An application of swarm optimization to nonlinear programming, Computers and Mathematics with Applications 49 (2005) 1655–1668.
[13] L. Yiqing, Y. Xigang, L. Yongjian, An Improved PSO Algorithm for Solving Non-convex NLP/MINLP Problems with Equality Constraints, Computers and Chemical Engineering, 2006. Available from: <http://www.sciencedirect.com>.
[14] ITU-T Y.1541, Network Performance Objectives for IP-based Services, International Telecommunications Union, 2002.