Optimal Scheduling Algorithms in WDM Optical Passive Star Networks

Hongjin Yeh, Kyubum Wee, and Manpyo Hong

College of Information and Communication, Ajou University, 442-749 Suwon, Korea
{hjyeh, kbwee, mphong}@madang.ajou.ac.kr

Abstract. All-to-all broadcast scheduling problems are considered in WDM optical passive star networks where k wavelengths are available in the network. It is assumed that each node has exactly one tunable transmitter and one fixed tuned receiver. All transmitters can tune to k different wavelengths, and have the same tuning delay δ to tune from one wavelength to another. In this paper, we take δ to be a nonnegative integer which can be expressed in units of packet durations. When all-to-all broadcasts are scheduled periodically in the network, the lower bounds are established on the minimum cycle length depending on whether each node sends packets to itself or not. And then, we present optimal scheduling algorithms in both cases for arbitrary number of wavelengths and for arbitrary value of the tuning delay. [1]

1 Introduction

In optical communications, one of the most popular packet switched network architecture is the passive star network because of its simplicity; packets may be transmitted directly from source to destination without any routing process. Additionally, WDM(Wavelength Division Multiplexing) is used to overcome the bottleneck of communication bandwidth due to optoelectronic components.[1] It is well known that tuning delay of transmitter/receiver can affect performance of the network. One of critical issues is how to minimize the effect of tuning delay for scheduling static or dynamic traffic patterns.

Let's consider WDM optical passive star networks with N nodes and k wavelength channels available in the network. All-to-all broadcast scheduling algorithms on WDM passive star networks may be classified into five categories depending on the number and tunability of transmitters/receivers for each node:

- STT-SFR, SFT-STR, MFT-SFR, SFT-MFR
- MTT-SFR, SFT-MTR
- STT-STR, MFT-MFR, STT-MFR, MFT-STR
- MTT-STR, MFT-MTR, STT-MTR, MTT-MFR
- MTT-MTR

[1] This research was supported by the Institute of Information Technology Assessment under Grant 97-G-0574.

where S: single, M: multiple, T: tunable, F: fixed, T: transmitter, R: receiver.

We assummed that each node has exactly one tunable transmitter (e.g., laser) and one fixed tuned receiver (e.g., filter). The network is asynchronous so that transmitters may send a packet or tune its wavelength at anytime. All transmitters can tune to k different wavelengths, and have the same tuning delay δ to tune from one wavelength to another. In this paper, we take δ to be an arbitrary nonnegative integer which can be expressed in units of packet durations. We also normalize the unit time such that all packets can be transmitted in the unit time so that the packet size is ignored for transmissions.

A special case of the general scheduling traffic patterns is the all-to-all broadcast in which every node has a single packet to be transmitted to the other. The problem is, given values k and δ, how one can schedule all-to-all broadcasts using k wavelengths within a repeating cycle of minimum length. A transmission schedule is a timetable that shows each packet is transmitted at which time using which wavelength. In order to tranfer a packet, the transmitter should tune to the wavelength of the receiver. A transmitter cannot transmit packets while it is tuning to another wavelength. Hence it needs to transmit all the packets that use the same wavelength without interruption if possible, and only then tune to another wavelength. Also a wavelength should be used many times, since the number of wavelengths used in WDM is limited compared to the number of nodes. Thus, obtaining an optimal schedule amounts to finding how to overlap tansmissions of packets and tuning delays as much as possible.

There are two kinds of all-to-all broadcast scheduling problems depending on whether each node sends packets to itself or not. In case each node sends single packet to itself, N^2 packets should be transmitted for each all-to-all broadcast. Pieris and Sasaki [5] proposed a scheduling algorithm under the assumtion that k divides N, and showed that the schedule length of their algorithm falls between $\max\{\delta + \frac{N^2}{k}, N\sqrt{\delta}\}$ and $\max\{\delta + \frac{N^2}{k}, k\delta + N \Leftrightarrow \frac{N}{k} + \frac{N^2}{k^2}\}$. Later it was shown that the schedule of [5] is actually optimal by Choi, Choi, and Azizoglu[2].

For the case where k does not divide N, they also proposed optimal scheduling algorithms for three different cases depending on the range of δ except for a very narrow range where their solution is within the bound of $\frac{13}{12}$ of the optimal schedule length [3]. In the case where each node does not send packets to itself, the number of packets to be sent for each all-to-all broadcast is $N(N \Leftrightarrow 1)$. Lee, Choi, and Choi[4] showed that $\max\{\frac{N(N-1)}{k}, k\delta + N \Leftrightarrow 1\}$ is a lower bound for the cycle lengths of schedules and proposed an algorithm for obtaining a transmission schedule that meets this lower bound under the assumption that k divides N. Chang, Yeh, and Wee[6] proposed an algorithm for obtaining a schedule where each node receives packets with a regular interval, and showed that the optimal schedule length in that case is $\max\{C(N \Leftrightarrow 1), \lceil \frac{\delta}{C-1} + \frac{1}{C-1} \lceil \frac{C-1}{k} \rceil \rceil\}$, where $C = \frac{N}{k}$ is an integer.

In this paper, we present scheduling algorithms that finds optimal schedules for arbitrary values of N, k, and δ.

2 Lower Bounds to Schedule Cycle Lengths

Let us consider WDM optical passive star networks with N nodes n_0, n_1, $\cdots$, n_{N-1}, where are k wavelengths $w_0, w_1, \cdots, w_{k-1}$ available in the network. We take arbitrary number of wavelengths k and arbitrary tuning delay δ for a given network composed of N nodes.

In this paper, it is assumed that all-to-all broadcasts are repeated periodically after initialization of the network. The problem is to schedule packet transmissions for each node pairs so that all nodes transmit once to every node within a repeating cycle as short as possible. The transmission schedules are called *optimal* if their cycle lengths meet a lower bound as shown in the following. All-to-all broadcast scheduling problems can be divided into two categories; one is the case that every node transfers packets to all nodes including itself, another is the case that each node transfers packets to the other nodes except itself.

Firstly, let us talk about the case that every node transmits once to itself for a cycle. For a given $2 \leq k \ll N$, let R_i denote the set of nodes in which their receivers' wavelengths are identically fixed to w_i for $0 \leq i \leq k \Leftrightarrow 1$. To maximize sharing all the wavelengths, let's define $R_i = \{n_j | j \equiv i \pmod{k}, 0 \leq j \leq N \Leftrightarrow 1\}$. Then, $|R_i| = \lceil \frac{N}{k} \rceil$ or $|R_i| = \lfloor \frac{N}{k} \rfloor = \lceil \frac{N}{k} \rceil \Leftrightarrow 1$. Specially, it is noted that $\max_{0 \leq i \leq k-1} |R_i| = |R_0| = \lceil \frac{N}{k} \rceil$ by the definition.

Theorem 1. *For values k and δ, the cycle length of any all-to-all broadcast schedule is at least*

$$\max\{\left\lceil \frac{N}{k} \right\rceil N, k\delta + N\}$$

where each of N^2 node pairs has single packet to be transmitted for a cycle.

Proof. Each node must receive N packets via only one wavelength fixed to its receiver. These packets are transmitted from all the nodes in the network including itself. Since more than two packets can not be transmitted via the same wavelength at a time, the cycle length is at least the number of packets, $|R_i|N$ for $0 \leq i \leq k \Leftrightarrow 1$, to be transmitted using a wavelength w_i for a cycle. Thus, $|R_0| = \lceil \frac{N}{k} \rceil$ means that the cycle length is at least $\lceil \frac{N}{k} \rceil N$.

On the other hand, every node transmits N packets using k different wavelengths for a cycle. If we take into account the fact that all-to-all broadcasts are repeated periodically, each node has to tune its transmitter k times for a cycle. Thus, as for a node, the cycle length is at least the sum of total tuning delay for a cycle and transmission time for N packets i.e., $k\delta + N$.

Another case is that nodes do not transfer any packets to themselves through all-to-all broadcast. That is, the number of packets transmitted to a set of nodes R_i using the wavelength w_i could be different if the transmitting node is an element of R_i for $0 \leq i \leq k \Leftrightarrow 1$. Thus, the lower bound to optimal schedule length is slightly changed.

Theorem 2. *For values k and δ, the cycle length of any all-to-all broadcast schedule is at least*

$$\max\{\left\lceil \frac{N}{k} \right\rceil (N \Leftrightarrow 1), k\delta + N \Leftrightarrow 1\}$$

where nodes do not tranfer any packets to themselves; each of $N(N \Leftrightarrow 1)$ node pairs has single packet to be transmitted for a cycle.

Proof. The proof is in the same manner as theorem 1.

In the next section, we present optimal scheduling algorithms for the latter which may be more complicated than the former.

3 Optimal Scheduling Algorithms

The main ideas to construct an optimal scheduling algorithm can be derived from two important system parameters, *the number of wavelengths* and *the tuning delay*, that affect performance of WDM optical passive star networks. The more wavelength channels are available in the network, the more packets can be transmitted simultaneously. On the other hand, tuning delay is dependent on practical devices and would limit the network throughput; while nodes spend time to tune from one wavelength to another, they have to delay packet transmissions. Therefore, transmission schedules compromise the following two contradictory conditions:

- As for a node, packet transmissions must be scheduled as soon as possible after every tuning delay for a cycle.
- As for a wavelength, packet transmissions must be scheduled as soon as possible from one node to another.

Remark In the previous research, it would be better that the packets transmitted via the same wavelength in a node are scheduled all together before tuning the transmitter to another wavelength. It should be noted, in the case of repeating all-to-all broadcasts periodically, that the cycle length includes the time for tuning to the wavelength used for the beginning of the next broadcast. That is, a wavelength looks like to be used one more time in a cycle with viewpoint of one stage of repeating all-to-all broadcasts.

In our proposed transmission schedules, the order of wavelengths used by each node is $w_j, w_{j-1}, \cdots, w_1, w_o, w_{k-1}, \cdots, w_{j+1}$, where $0 \leq j \leq p \bmod k$. That is, the first group to transmit packets to is R_j with $|R_j| = \lceil \frac{N}{k} \rceil$. The order of nodes using the same wavelength is $n_0, n_1, \cdots, n_{k-1}$, or a circular shift of this order. Without loss of generality, we can assume that the order of wavelengths used by each node is $w_0, w_{k-1}, \cdots, w_1$, since we know that $|R_0| = \lceil \frac{N}{k} \rceil$. Also in the following we notate w_0 and w_k interchangeably. The same comment applies to R_0 and R_k.

Let $N \times k$ matrix M that represents relations between a node n_p for $0 \leq p \leq N \Leftrightarrow 1$ and the set of nodes R_{k-i} containig n_p for $0 \leq i \leq k \Leftrightarrow 1$. Then, the matrix $M = (m_{pi})$ is defined by

$$m_{pi} = \begin{cases} 1 \text{ if } p + i \equiv 0 \pmod{k} \\ 0 \text{ otherwise} \end{cases}$$

for $0 \leq p \leq N \Leftrightarrow 1$ and $0 \leq i \leq k \Leftrightarrow 1$. Note that the indices of m_{pi} starts from zero. For example, in the case of $N = 8$ and $k = 3$, the matrix M is the following:

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus, $m_{12} = 1$ means that the receiver of the node n_1 is fixed tuned to the wavelength $w_{3-2} = w_1$.

Lemma 1. *For $N \times k$ matrix $M = (m_{pi})$,*

$$\sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j} \geq 0$$

for $0 \leq p \leq N \Leftrightarrow 1$ and $0 \leq i \leq k \Leftrightarrow 1$.

Proof. By the definition of the matrix M, the values $\sum_{j=0}^{i} m_{pj}$ and $\sum_{j=0}^{i-1} m_{(p+1)j}$ are always 0 or 1. If $\sum_{j=0}^{i} m_{pj} = 1$, $\sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j} \geq 0$ because the value $\sum_{j=0}^{i-1} m_{(p+1)j}$ is at most 1.

Otherwise, i.e., If $\sum_{j=0}^{i} m_{pj} = 0$, $p + j \not\equiv 0 \pmod{k}$ for $0 \leq j \leq i$. Since $(p+1) + (j \Leftrightarrow 1) \not\equiv 0 \pmod{k}$ for $1 \leq j \leq i$ implies $(p+1) + j \not\equiv 0 \pmod{k}$ for $0 \leq j \leq i \Leftrightarrow 1$, $\sum_{j=0}^{i-1} m_{(p+1)j} = 0$. Thus, $\sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j} = 0$. Therefore, given inequality always holds.

Each node n_p starts transmitting packets to the nodes in group R_{i-1} using wavelength w_{i-1} right after finishing transmitting packets to the nodes in group R_i using wavelength w_i delayed only by tuning delay δ. The time s_p when node n_p transmits the very first packet using wavelength w_k (i.e., w_0) is set as follows:

$$s_p = \left\lceil \frac{N}{k} \right\rceil p \Leftrightarrow \left\lceil \frac{p}{k} \right\rceil + 1$$

Algorithm 1 shows the initial transmission schedules compromising tuning delay and lack of the wavelengths. In the algorithm, the statement $T(n_p, t) = n_q$ means that node n_p transmits a packet to node n_q at time t. The wavelength used in this case is $w_{q \bmod k}$. In the algorithm, the cycle length of optimal schedules L_{opt} is defined by $\max\{\lceil \frac{N}{k} \rceil (N \Leftrightarrow 1), k\delta + N \Leftrightarrow 1\}$.

Algorithm 1 *Initial Schedules of an All-to-all Broadcast*

```
For p = 0 to N ⇔1 do
    t ← s_p
    For r = k to 1 do
        i ← r mod k
        For j = 0 to |R_i| ⇔1 do
            If (p ≠ jk + i)
                T(n_p, t) ← n_{jk+i}
                t ← t + 1
        t ← t + δ
```

Figure 1 shows the transmission schedule constructed by Algorithm 1 when $N = 8, k = 3$, and $\delta = 5$. The numbers on the leftmost column represent the nodes, and time goes on from left to right.

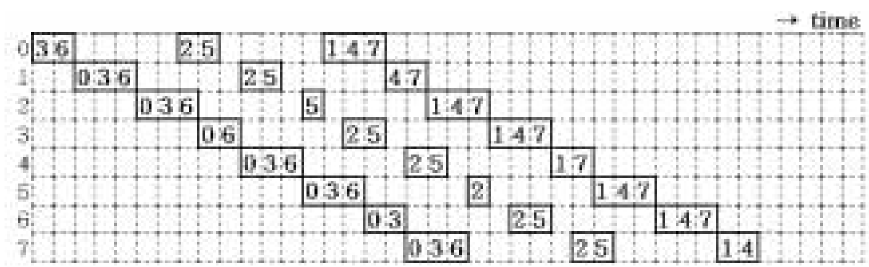

Fig. 1. An example of initial schedules

Lemma 2. *In the schedules generated by Algorithm 1, no two nodes use the same wavelength at the same time.*

Proof. We only have to check that two adjacent nodes n_p and n_{p+1} do not use the same wavelength simultaneously. Let α be the time when node n_p finishes transmitting packets using wavelength w_{k-i}, and β be the time when node n_{p+1} finishes tuning its transmitter to start packet transmissions using wavelength w_{k-i}. Then α and β are expressed as follows:

$$\begin{aligned}\alpha &= s_p + \textit{the number of packets transmitted} + \textit{total tuning delay} \\ &= \left(\left\lceil \frac{N}{k} \right\rceil p \Leftrightarrow \left\lceil \frac{p}{k} \right\rceil + 1\right) + \sum_{j=0}^{i}(|R_{k-j}| \Leftrightarrow m_{pj}) + i\delta\end{aligned}$$

Similarly,

$$\beta = \left(\left\lceil \frac{N}{k} \right\rceil (p+1) \Leftrightarrow \left\lceil \frac{p+1}{k} \right\rceil + 1\right) + \sum_{j=0}^{i-1}(|R_{k-j}| \Leftrightarrow m_{(p+1)j}) + i\delta$$

Hence

$$\beta \Leftrightarrow \alpha = \left\lceil \frac{N}{k} \right\rceil \Leftrightarrow \left\lceil \frac{p+1}{k} \right\rceil + \left\lceil \frac{p}{k} \right\rceil \Leftrightarrow |R_{k-i}| + \sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j}$$

$$\geq \left\lceil \frac{p}{k} \right\rceil \Leftrightarrow \left\lceil \frac{p+1}{k} \right\rceil + \sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j}$$

since $|R_{k-i}| \leq \lceil \frac{N}{k} \rceil$. Recall that $\sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j} \geq 0$ by Lemma 1.

Hence $\beta \Leftrightarrow \alpha \geq 0$ because $\lceil \frac{p}{k} \rceil \Leftrightarrow \lceil \frac{p+1}{k} \rceil = 0$ unless p is a multiple of k. Now, we have to consider the case when $p \equiv 0 \pmod{k}$. In this case, $\lceil \frac{p}{k} \rceil \Leftrightarrow \lceil \frac{p+1}{k} \rceil = \Leftrightarrow 1$. By the definition of matrix M, $p + 0 \equiv 0 \pmod{k}$ implies that $m_{p0} = 1$ is unique nonzero element in p-th row. Similarly, $(p+1) + (k \Leftrightarrow 1) \equiv 0 \pmod{k}$ implies that $m_{(p+1)(k-1)} = 1$ is unique nonzero element in $(p+1)$-th row. Thus, $\sum_{j=0}^{i} m_{pj} \Leftrightarrow \sum_{j=0}^{i-1} m_{(p+1)j} = 1 \Leftrightarrow 0 = 1$ for $0 \leq i \leq k \Leftrightarrow 1$. Therefore, $\beta \Leftrightarrow \alpha \geq 0$ always holds for any value p.

Now let us denote the time for each node n_p to transmit all the packets for an all-to-all broadcast as L_p. Note that the tuning delay for next all-to-all broadcast is not considered.

Lemma 3. *In the schedules generated by Algorithm 1,* $L_p = (k \Leftrightarrow 1)\delta + N \Leftrightarrow 1$ *for any* p *in* $0 \leq p \leq N \Leftrightarrow 1$.

Proof. Lemma 2 guarantees that each node does not have to wait for using a wavelength that is being used by another node. In other words, each node can start transmitting packets as soon as tuning to the wavelength is completed. By the way, every node n_p transmits $N \Leftrightarrow 1$ packets. And every node n_p needs to tune $k \Leftrightarrow 1$ times. Hence, $L_p = (k \Leftrightarrow 1)\delta + N \Leftrightarrow 1$ for any p.

Since all L_p are identical for $0 \leq p \leq N \Leftrightarrow 1$, let us denote L_p simply as L. Now we claim that if we repeat the transmission schedule constructed by Algorithm 1 by many times, any time span of L_{opt} starting from any position of the repeated schedule is an optimal schedule. Algorithm 2 finds an optimal schedule. Algorithm 2 differs from Algorithm 1 in that time t is replaced by $t \Leftrightarrow L_{opt}$ if t is later than L_{opt}.

For example, Figure 2 shows the repetitions of the transmission schedule of Figure 1. Note that in Figure 2, $L_{opt} = 22$ is calculated with the values N, k and δ. If we take out any portion of Figure 2 spanning 22 time units, that constitutes an optimal schedule.

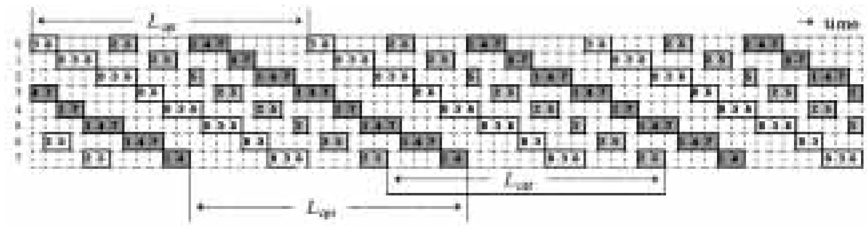

Fig. 2. An example of repeating initial schedules

Algorithm 2 *Optimal Schedules for an All-to-all Broadcast*

For $p = 0$ *to* $N \Leftrightarrow 1$ *do*
    $t \leftarrow s_p$
    For $r = k$ *to 1 do*
        $i \leftarrow r \bmod k$
        For $j = 0$ *to* $|R_r| \Leftrightarrow 1$ *do*
            If $(p \neq jk + i)$
                If $(t > L_{opt})$ $T(n_p, t \Leftrightarrow L_{opt}) \leftarrow n_{jk+i}$
                else $T(n_p, t) \leftarrow n_{jk+i}$
                $t \leftarrow t + 1$
        $t \leftarrow t + \delta$

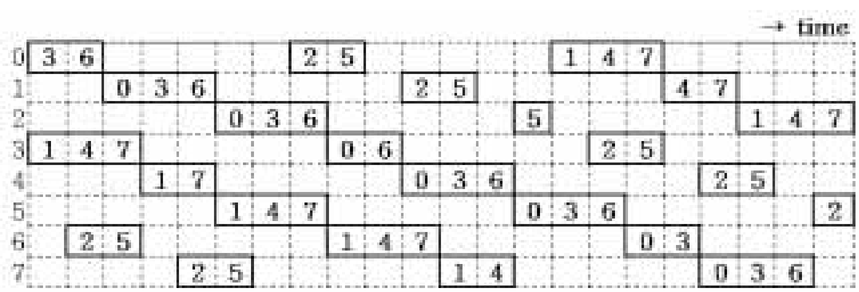

Fig. 3. An example of optimal schedules with $N(N-1)$ packets

Theorem 3. *For arbitrary values k and δ, the cycle length of transmission schedules by Algorithm 2 is optimal.*

Proof. All we have to do is to show that $L_{opt} \Leftrightarrow L$ is no less than tuning delay δ. Recall that $L = (k \Leftrightarrow 1)\delta + (N \Leftrightarrow 1)$ by Lemma 3. If $L_{opt} = \lceil \frac{N}{k} \rceil (N \Leftrightarrow 1)$ then

$$\lceil \frac{N}{k} \rceil (N \Leftrightarrow 1) \geq k\delta + N \Leftrightarrow 1 \Leftrightarrow \lceil \frac{N}{k} \rceil (N \Leftrightarrow 1) \Leftrightarrow \{(k \Leftrightarrow 1)\delta + (N \Leftrightarrow 1)\} \geq \delta$$
$$\Leftrightarrow L_{opt} \Leftrightarrow L \geq \delta$$

Otherwise, $L_{opt} = k\delta + N \Leftrightarrow 1$ implies

$$L_{opt} \Leftrightarrow L = k\delta + N \Leftrightarrow 1 \Leftrightarrow \{(k \Leftrightarrow 1)\delta \Leftrightarrow (N \Leftrightarrow 1)\} = \delta$$

Remark In the above example, $L_{opt} = 3 \times 5 + 8 \Leftrightarrow 1 = 22$ is affected by the value $\delta = 5$. Our algorithm also works in case that the value δ is small enough to determine L_{opt} only with values N and k. Figure 4 shows the optimal schedule generated by Algorithm 2 when $\delta = 4$ with the same values $N = 8$ and $k = 3$. In such case, $L_{opt} = \lceil \frac{8}{3} \rceil (8 \Leftrightarrow 1) = 21$.

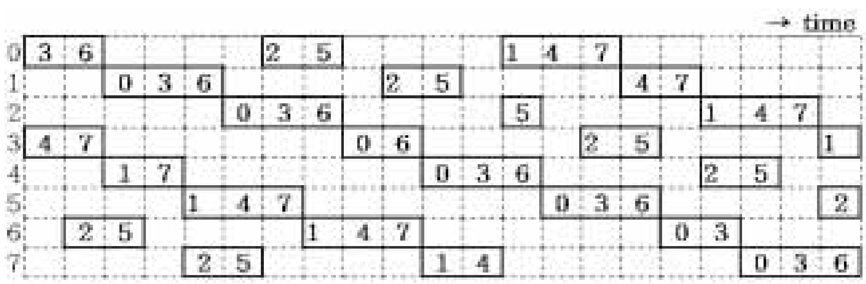

Fig. 4. An example of optimal schedules for $\delta = 4$

Remark Another scheduling problem for all-to-all broadcasts takes place when every node transmits packets to itself. Our algorithms can be applicable to this problem if the line 6 in Algorithm 2 is deleted. In fact, each node has the same number of packets that are transmitted using the same wavelength. This property makes the proposed algorithms more simple. The proof of optimality in transmission schedules is omitted. The following example is depicted to compare two versions of optimal schedules when $N = 8, k = 3$ and $\delta = 5$.

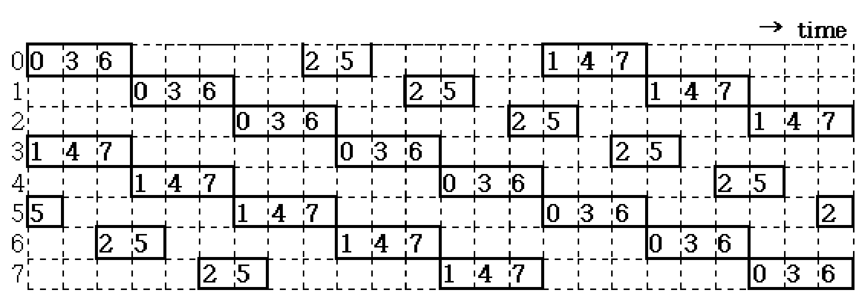

Fig. 5. An example of optimal schedules with N^2 packets

4 Conclusions

We presented optimal scheduling algorithms for all-to-all broadcasts repeated periodically in WDM optical passive star networks where each node has one tunable transmitter and one fixed receiver. Our algorithms work for arbitrary values k and δ. There are two versions: one for the case where nodes send packets to themselves and the other case where they do not. Both versions find optimal schedules.

As long as we use electrical buffers, it is important to guarantee enough time to electrically process received packets which are usually of very big size in optical networks. Hence, we are currently working on optimal scheduling algorithms with additional condition that each node receives packets at a regular interval. On the other hand, finding optimal schedules for the other classes as mentioned in the Introduction, particulary the case where both transmitters and receivers are tunable, would be an interesting future research topic.

References

1. C. A. Brackett, "Dense Wavelength Division Multiplexing Networks: Principles and Applications," IEEE J. on Selected Areas in Communications. Vol. 8, No. 6, pp. 948-964, Aug. 1990.
2. H. Choi, H-A. Choi, and M. Azizoglu, "Efficient Scheduling of Transmissions in Optical Broadcast Networks," IEEE/ACM Trans. on Networking, Vol. 4, No. 6, pp. 913-920, Dec. 1996.
3. H. Choi, H-A. Choi, and M. Azizoglu, "On the All-to-All Broadcast Problem in Optical Network," Proceedings of Infocom '97, 1997.
4. S-K. Lee, A. D. Oh, H. Choi, H-A. Choi, "Optimal Transmission Schedules in TWDM Optical Passive Star Networks," Discrete Applied Mathematics, Vol. 75, No. 1, pp. 81-91, Jan. 1997.
5. G. R. Pieris and G. H. Sasaki, "Scheduling Transmissions in WDM Broadcast-and-select Networks," IEEE/ACM Trans. on Networking, Vol. 2, No. 2, pp. 105-110, April, 1994.
6. S. Chang, H-J. Yeh, K. Wee, "Regular Transmission Schedules for All-to-All Broadcast in Optical Passive Star Interconnections," Ajou Univ. J. of Research Institut for Information and Communication, Vol. 1, No. 1, pp. 165-173, Feb. 1998.