유전자 알고리즘을 이용한 WDM 네트워크 최적화 방법

Genetic Algorithm based Methodology for Network Performance Optimization

양 효 식*
Hyo-Sik Yang*

요 약

WDM 네트워크는 높은 전송속도와 낮은 지연시간으로 메트로폴리탄 네트워크뿐만 아니라 최근 기가비트 이더넷 등을 이용하여 근거리 망에서도 많은 연구가 진행되어 왔다. 네트워크의 성능은 네트워크 구조의 파라미터 값들과 사용되는 Medium Access Control 프로토콜의 파라미터 값들에 많이 의존한다. 또한 네트워크 효율성과 지연시간은 주로 상반된 관계를 보여 한쪽의 희생이 불가피 하였다. 네트워크를 효율적으로 운용하기위해서는 효율성과 지연시간이라는 성능의 최적값을 찾아야 상황에 맞게 운용할 수 있다. 본 논문에서는 Arrayed Waveguide Grating (AWG) 기반의 성형 WDM 네트워크상에서 효율성의 최대화와 지연시간의 최소화라는 두 개의 서로 상반된 목적 함수를 유전자 알고리즘 기반의 방법론을 이용하여 파레토 최적화 곡선이라는 최적의 값들을 찾아내었다. 이를 이용하여 구한 최적의 네트워크 구성을 위한 파라미터 값들과 MAC 프로토콜의 파라미터 값들을 이용하여 상황에 따른 최적의 네트워크 성능을 유추할 수 있게 되었다. 본 논문에 사용된 유전자 알고리즘을 이용한 최적화 방법은 이와 유사한 상반된 목적 함수를 갖는 네트워크의 성능을 최적화하는데 사용될 수 있을 것이다.

Abstract

This paper considers the multi-objective optimization of a multi-service arrayed waveguide grating-based single-hop WDM network with the two conflicting objectives of maximizing throughput while minimizing delay. This paper presents a genetic algorithm based methodology for finding the optimal throughput-delay tradeoff curve, the so-called Pareto-optimal frontier. Genetic algorithm based methodology provides the network architecture parameters and the Medium Access Control protocol parameters that achieve the Pareto-optima in a computationally efficient manner. The numerical results obtained with this methodology provide the Pareto-optimal network planning and operation solution for a wide range of traffic scenarios. The presented methodology is applicable to other networks with a similar throughput-delay tradeoff.

Keywords : Genetic algorithm, Multi-objective optimization, Pareto-optima, WDM network

I. Introduction

Optical single-hop wavelength division multiplexing (WDM) networks have the potential to provide high throughput and low delay connectivity in metropolitan and local area settings [1, 2].The throughput-delay performance of these single-hop WDM networks is typically very sensitive to the setting of the architecture parameters and the medium access control (MAC) protocol parameters [3].For good network performance, these parameters must be set properly, which is a challenge due to the large search space of possible parameter combinations and the typically computationally demanding evaluation of a particular parameter combination. The objectives to maximize the throughput while minimizing the delay are typically conflicting. With certain combinations of parameter settings, the networks achieve a small delay and moderate throughput, which is perfectly suited for delay sensitive traffic. On the other hand, certain combinations of parameter settings achieve a large throughput but introduce some moderate delays, which is perfectly suited for throughput sensitive

* Dept. of Computer Eng. Sejong University
논문 번호 : 2007-4-39 접수 일자 : 2007. 11. 20
심사 완료 : 2008. 1. 3

traffic that can tolerate some delays.

The existing studies employing genetic algorithms for WDM network optimization typically optimize a single objective. For example, wavelength assignment in the design of wavelength routed mesh WDM networks have been optimized [4]. The optimal placement of wavelength converters in WDM mesh networks is studied in [5]. Optimality issues in planning of survivable WDM networks have been investigated in [6]. In contrast, multi-objective optimization problem is considered in this paper.

This paper presents a genetic algorithm based methodology for solving the multi-objective optimization problem of maximizing throughput and minimizing delay in single-hop WDM networks. As an example of application of developed methodology, Arrayed-Waveguide Grating (AWG) based network is considered throughout this paper [7]. This methodology finds the optimal tradeoff curve and the parameter combinations attaining the curve in a computationally efficient manner. This methodology applies analogously to networks with a similar throughput delay tradeoff.

II. AWG based Single Hop WDM Network

The basic architecture of the single-hop WDM network [7] is based on a D×D AWG, as shown in Fig. 1. At each AWG input port, a wavelength insensitive S×1 combiner collects data from S attached nodes. Similarly, at each AWG output port, signals are distributed to S nodes by a wavelength insensitive 1×S splitter. Each node is composed of a transmitting part and a receiving part. The transmitting part of a node is attached to one of the combiner ports. The receiving part of the same node is located at the opposite splitter port. The network connects N=D · S nodes. At each AWG input port, R adjacent Free Spectral Ranges (FSRs) of the AWG are exploited. Each FSR consists of D contiguous wavelengths. The total number of wavelengths at each AWG input port is Λ=D · R. architecture. For interesting readers, please refer [7].

In the MAC protocol, time is divided into cycles. Each cycle consists of D frames. Each frame contains F slots. Each frame is partitioned into the first M, 1≤M<F, slots and the remaining (F-M) slots. In the first M slots, control signals are transmitted based on a modified slotted ALOHA protocol and all nodes must be tuned (locked) to one of the Light Emitting Diode (LED) slices carrying the control information. In every frame within the cycle, the nodes attached to a different AWG input port send their control packets. In the considered traffic scenario, a node that is not backlogged generates a new packet with probability σ at the beginning of its transmission cycle. The generated packet is long (has size F slots) with probability q, and is short (has size K=F-M slots) with probability (1-q).

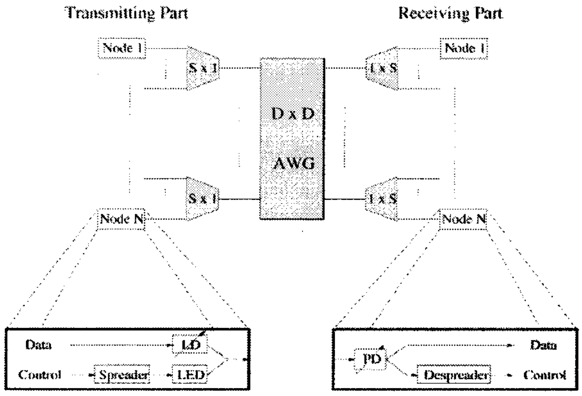

Fig. 1. Architecture of AWG based WDM network.

The two key performance metrics of AWG based single-hop WDM networks are the mean throughput and the mean delay. The typical goal of the optimization of single-hop WDM networks is to maximize the throughput while minimizing the delay. For the AWG based network, the mean throughput and the mean delay have been derived as follows [7]:

$$\text{Throughput} = D^2 \cdot \frac{F \cdot E[L] + K \cdot E[S]}{F \cdot D}, \quad (1)$$

and

$$\text{Delay} = \left\{ \frac{S}{D \cdot (E[L] + E[S])} - \frac{1-\sigma}{\sigma} \right\} \cdot D \cdot F \quad (2)$$

where, *E[L]* is the expected number of successfully scheduled long packets (of size F slots) from a given (fixed) AWG input port to a given (fixed) AWG output port per cycle (of length F · D slots), and *E[S]* is the expected number of successfully scheduled short packets (of length K=F-M slots) from a given (fixed) AWG input port to a given (fixed) output port per cycle.

III. Genetic Algorithm based Methodology

3.1 Multi-Objective Problem

The familiar notion of an optimal solution becomes somewhat vague when a problem has more than one objective function, as is the case in this metro WDM network optimization. A solution that gives very large throughput may also give large delay and thus rate poorly on the minimize delay objective. The best strategy is to find a set of optimal tradeoff solutions, i.e., solutions that give the largest achievable throughput for a given tolerable delay, or equivalently the smallest achievable delay for a required throughput level. After a set of such optimal tradeoff solutions is found, a user can use higher level considerations to make a choice. A feasible solution to a multi-objective optimization problem is referred to as *efficient frontier* or *Pareto optimal solution* [8]. As illustrated in Fig. 2, two objectives, maximizing throughput and minimizing delay are presented. The goal of multi-objective optimization is to find such a feasible efficient frontier. Classical methods for generating the Pareto-optimal solution set aggregate the objectives into a single, parameterized objective function [9].

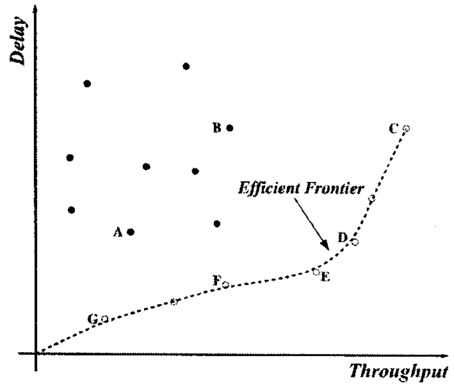

Fig. 2. Illustration of efficient frontier for maximizing throughput-minimizing delay

3.2 Decision Variables and Constraints

This section identifies the decision variables in this optimization problem and identifies the constraints on the decision variables. The AWG degree D is selected as the (independent) decision variable for the network (hardware) architecture; the other architecture parameters R and S as functions of D (and the given N and Λ) are determined, as discussed shortly. Generally, the decision variable D can take any integer satisfying

$$D \geq 2 \quad \text{and} \quad D \leq \Lambda \tag{3}$$

where Λ is the maximum number of wavelength channels accommodated by the fast tunable transceivers employed in the considered network. In other words, Λ is the maximum tuning range of the employed transceivers divided by the channel spacing and is thus very technology dependent. To use transceivers with a negligible tuning time (and a small tuning range) Λ is set to 8 in this paper. Note that the number of ports of commercially available photonic devices is typically a power of two. This constraint can be easily incorporated by restricting D to the set 2, 4, 8,....

The number of used FSRs R depends on the (independent) decision variable D and the given tuning range Λ of the transceivers. Generally, R must be an integer satisfying $R \cdot D \leq \Lambda$, i.e., $R \leq \Lambda/D$. The larger R, the more parallel channels are available between each input-output port pair of the AWG, and hence the larger the throughput. Therefore, R is set to the largest integer less than or equal to Λ/D, i.e., $R=\lfloor \Lambda/D \rfloor$. Note that the tuning range Λ and degree D are typically powers of two for commercial components. Hence, Λ/D is a power of two for practical networks, thus $R=\Lambda/D$. The combiner/splitter degree S depends on the decision variable D and the given number of nodes in the network N. In determining the combiner/splitter degree S, it is natural to assume that the nodes are equally distributed among the D AWG input/output ports; i.e., each input/output port serves at least $\lfloor N/D \rfloor$ nodes. This arrangement minimizes the required combiner /splitter degree S, which in turn minimizes the splitting loss in the combiners/splitters, $S=\lceil N/D \rceil$. As for the protocol (software) parameters, three decision variables, F, M, and p, are identified. Generally, the number of slots per frame F can take any positive integer, i.e., $F \geq 1$, while the number of control slots per frame can take any positive integer less than or equal to F, i.e., $1 \leq M \leq F$. Note that the size of the packets to be transported may impose additional constraints on F and M. With a given maximum packet size, F must be large enough to accommodate the maximum size packet in a frame. If short packets have a specific size requirement, F-M should be large enough to accommodate that packet size. In this methodology, however, genetic algorithm determines the F and M values that give the optimal throughput-delay performance, subject only to $F \geq 1$ and $1 \leq M \leq F$. The packet re-transmission probability p may take any real number in the interval [0, 1]. To reasonably limit the search space restriction on p to [0, 0.05, 0.10, 0.15,..., 1.0] in this numerical work.

3.3 Operation of Genetic Algorithm

In the genetic algorithm, a population of individuals are considered. Each individual is represented by a string of the decision variables. In the terminology of genetic algorithms the string of decision variables is referred to as *chromosome,* while each individual decision variable is referred to as *gene.* The quality of an individual in the population with respect to the two objective functions is represented by a scalar value called *fitness.* After generating the initial population (by randomly), each individual is assigned a fitness value. The population is evolved repeatedly, generation by generation, using the crossover operation and the mutation operation. The crossover and mutation operations produce offspring by manipulating the individuals in the current population that have good fitness values. The crossover operation swaps portions of the chromosomes as shown in Fig. 3. The mutation operation changes the value of a gene. Individuals with a better fitness value are more likely to survive and to participate in the crossover (mating) operation. After a number of generations, the population contains members with better fitness values. The Pareto- optimal individuals in the final population are the outcome of the genetic algorithm.

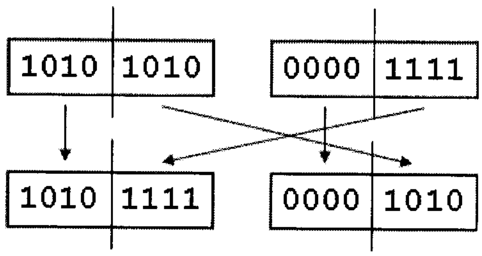

Fig. 3. Crossover operation

1) Fitness Function

The fitness function is typically a combination of objective functions. Three commonly used types of fitness function are considered. This methodology generates 20 generations, each with a population size of 200 to compare the quality of the fitness functions. The probability of crossover is set to 0.9 and the probability of mutation is set to 0.05. The genetic algorithm outputs are compared with the true Pareto-optimal solutions which were found by conducting an exhaustive search over all possible combinations of the decision variables in limited condition. Mean arrival rate is set to 0.6 for this evaluation.

First, the Vector Evaluated Genetic Algorithm (VEGA) is evaluated, which is easy to implement. The VEGA algorithm divides the population into two subpopulations according to two objective functions. The individuals in each subpopulation are assigned a fitness value based on the corresponding objective function. The main disadvantage of VEGA is that typically after several generations, the algorithm fails to sustain diversity among the Pareto-optimal solutions and converges near one of the individual solutions.

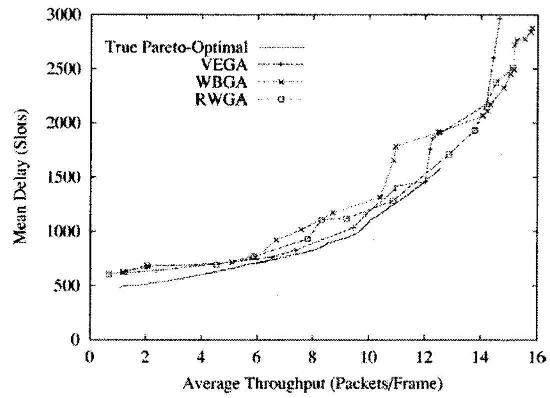

Fig. 4. Efficient frontier obtained with different fitness functions without elitism.

Next, the Weight Based Genetic Algorithm (WBGA) which uses the weighted sum of the objective functions is used as fitness function. The main difficulty in WBGA is that it is hard to choose the weight factors. Same weight factor of 1/2 for each objective function are used. The fitness function used is

$$Fitness = 0.5 \times Th - 0.5 \times Delay \qquad (3)$$

Finally, the Random Weight Genetic Algorithm (RWGA) which weighs the objective functions randomly is evaluated. A new independent random set of weights is drawn each time an individual's fitness is calculated. The fitness function used is

$$Fitness = \varepsilon \times Th - (1-\varepsilon) \times Delay \qquad (4)$$

where ε is uniformly distributed in the interval (0, 1).

As observed from Fig. 4 that the RWGA efficient frontier is relatively far from the true efficient frontier in the throughput range from 8~10 pkts/frame.

Elitism is one of the schemes used to improve the search; with elitism the good solutions in a given generation are kept for the next generation. This prevents losing the already found good solutions in the subsequent crossover operation(s), which may turn good solutions into bad solutions. The results obtained with elitism are given in Fig. 5. The efficient frontiers are closer to the true efficient frontier of the problem. According to the observations made in this section, RWGA with elitism is used throughout the remainder of this paper.

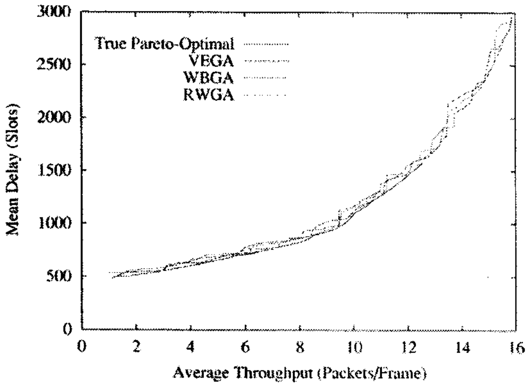

Fig. 5. Efficient frontier obtained with different fitness functions with elitism.

2) Population size and the number of generation

The population size trades off the time complexity (computational effort) and the number of optimal solutions. In order to accommodate all Pareto-optimal solutions, the population should be large enough. However, as the population size grows, the time complexity for processing a generation increases. On the other hand, for a smaller population, the time complexity for the population decreases while the population may lose some Pareto-optimal solutions. As a result, the smallest population size which can accommodate all Pareto-optimal solutions is preferable. Efficient frontiers with different population size, *P*, and the number generations, *G*, are shown in Fig. 6.

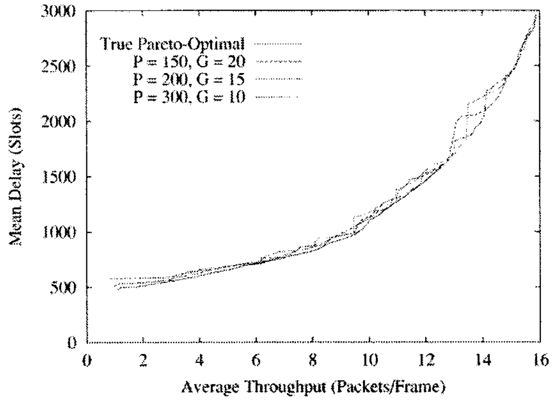

Fig. 6. Efficient frontier with different population size and the number of generations.

Ⅳ. Numerical Results

In this section, the developed genetic algorithm based methodology is employed in the preceding section to optimize the AWG-based single-hop WDM network. The methodology determines the settings of the network architecture parameter that give Pareto-optimal throughput-delay performance. The random weight genetic algorithm (RWGA) with elitism is used with the parameter settings found in the preceding section, a population size of P=200, G=40 generations, crossover probability 0.9, and mutation probability 0.05. Data packets can have one of two lengths. A data packet is long packet with probability *q*, and short packet with probability (*1-q*). To reasonably limit the search space the number of slot in a frame, F, is set to be no larger than 400 slots. The number of nodes in the network is set to 200 and the transceiver tuning range is fixed at 8 wavelengths.

Efficient frontier with mean arrival rate of 0.6 is shown in Fig. 7. We also conduct an optimization where the traffic load and the fraction *q* of long packet traffic are free decision variables, which gives the vest achievable network performance, refer to as *network frontier*. The number of Pareto-optimal solutions with D=2, 4, and 8 are given in Table 1. The complete parameter vectors corresponding to the Pareto-optimal solutions are given in [10].

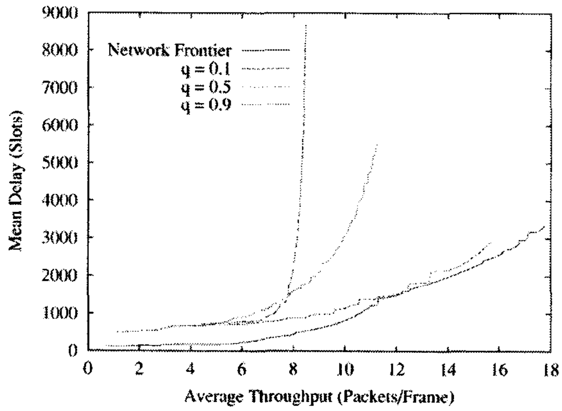

Fig. 7. Efficient frontier with mean arrival rate of 0.6 with different fraction of *q* of long packet traffic and network frontier.

The proper setting of the AWG degree D is considered in detail. The setting of this network architecture (hardware) parameter has a profound impact on the network performance, once the network hardware for a particular D value has been installed, it is very difficult and costly to change D; whereas the protocol parameters F, M, and p can easily be changed by modifying the network protocol (software). For this reason, the proper setting of D warrants special attention. According to Table 1, for predominantly long packet traffic (i.e., large q), D=2 is the best choice for all levels of traffic load σ. For predominantly short packet traffic (i.e., small q), on the other hand, the choice is not so clear. For light traffic loads, D=2 is the best choice, whereas for heavy traffic loads, D=4 turns out to be the best choice.

Table 1. Number of Pareto-optimal Solutions with D=2, 4, and 8

	σ=0.1			σ=0.3			σ=0.6			σ=0.8		
q	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
D=2	148	132	133	108	84	158	31	102	121	23	105	135
D=4	0	1	8	2	65	4	86	46	5	102	46	3
D=8	0	0	0	0	2	2	1	4	1	0	4	1
Total	148	133	141	110	151	164	118	152	127	125	155	139

V. Conclusion

In this paper a genetic algorithm based methodology for the multi-objective optimization problem of maximizing throughput while minimizing delay in an AWG-based metro WDM network is developed. This methodology finds the Pareto-optimal throughput delay trade-off curve in a computationally efficient manner. The optimal tradeoff curve can be used to optimally provide varying degrees of small delay (and moderate throughput) or large throughput (and moderate delay) packet transport services. Developed methodology thus facilitates efficient multi service convergence for increased cost effectiveness in metropolitan and local area networks.

The developed genetic algorithm methodology can be applied in analogous fashion to networks with a similar throughput delay tradeoff. The methodology is especially useful for the multi-objective optimization of networks with complex, highly non-linear characterizations of the network throughput and delay.

References

[1]. B. Mukherjee, "WDM optical communication: Progress and Challenge," *IEEE Jourl. Sel. Areas on Comm.* vol. 18, No. 18, pp.1810-1824, 2000

[2]. K. Benji, "Performance of single-hop WDM LAN's supporting real-time services," *Photonic Network Communications*, vol. 1, no. 4, pp. 287-301, 1999

[3]. J. Spaeth, "Dynamic wavelengths under decentral ized management." in *TransiNet Workshop*, 2001

[4]. K. Singh and R. Bhattacharjee, "A novel GA based wavelength allocation scheme in WDM systems", *Proc. of ICOCN*, pp. 299-302, 2005

[5]. S. Runggeratigul, P. Thongsri, P. "Optimal number of wavelength converters for WDM network design", *Proc. of TENCON*, vol. 3, pp. 93-96, 2004

[6]. Y. S. Kavian and M. Naderi, "An evolutionary approach to design WDM telecommunication survivable networks", *Proc. of ICIS*, pp. 1-6, 2006

[7]. M. Maier, M. Sheutzow, M. Reisslein, A. Wolisz, "Wavelength reuse for efficient transport of

variable size packets in a metro WDM network," Proc. of *INFOCOM*, pp. 1432-1441, 2002

[8]. K. Deb, "Multi-objective optimization using evolutionary algorithm," New York: Wiley, 2001

[9]. V. T. Le, X. Jiang, S. Horiguchi, and Y. Inoguchi, "A new fitness function for GA-based dynamic RWA algorithms in optical WDM networks", *Proc. of ICC,*, vol 2, pp. 667-671, 2005

[10]. H.-S. Yang, M. Maier, and M. Reisslein, "A genetic algorithm based methodology for optimizing multi-service convergence in a metro WDM network (Extended version)", *Tech. Report,* Arizona State Univ., http://fulton.asu.edu/~mre

Hyo-Sik Yang is an assistant professor at Sejong University, Dept. of Computer Eng. Korea. He received the B. E. degree from Myongji University, Korea, in 1998 and the M. S. and Ph. D. degree in Electrical Eng. from Arizona State Univ. U.S.A, in 2000 and 2005, respectively. His research interests are WDM all optical networks, WDM packet switching, and WDM MAN including node architecture, optimization, MAC, and routing in mobile ad-hoc networks. He is also working on Substation Automation System (SAS).