Scientific
Research
Publishing

Design of a Heuristic Topology Generation Algorithm in Multi-Domain Optical Networks

Lei Wang[1], Huayang Feng[1], Li Lin[2], Li Du[1*]

[1]College of Computer Science and Engineering, Northeastern University, Shenyang, China
[2]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, USA
Email: *duli@mail.neu.edu.cn

Abstract

Designing an excellent original topology not only improves the accuracy of routing, but also improves the restoring rate of failure. In this paper, we propose a new heuristic topology generation algorithm—GA-PODCC (Genetic Algorithm based on the Pareoto Optimality of Delay, Configuration and Consumption), which utilizes a genetic algorithm to optimize the link delay and resource configuration/consumption. The novelty lies in designing the two stages of genetic operation: The first stage is to pick the best population by means of the crossover, mutation, and selection operation; The second stage is to select an excellent individual from the best population. The simulation results show that, using the same number of nodes, GA-PODCC algorithm improves the balance of all the three optimization objectives, maintaining a low level of distortion in topology aggregation.

Keywords

Topology Generation, Genetic Algorithm, Crossover Operation, Topology Aggregation

1. Introduction

The scale of optical networks is rapidly expanding with the development of optical communication technology [1], the structure of optical networks has also changed greatly. To adapt to the new structure, optical networks are already moving towards multi domain in recent years. The excessive network information, meanwhile, causes bottlenecks in the scalability of optical networks and increases the necessity to provide a better scalability. The generation of original topology is one of the impetuses to solve the problems of scalability in large-scale multi-domain optical networks [2]. Meanwhile, original topology also influences

the performance of all upper-layer topology-related strategies [3].

The generation of original topology is a NP (Non-deterministic Polynomial) complete problem. Since the network topology is needed in urgent, a heuristic algorithm is proposed quickly. The available resources are taken as the constraint to determine the network topology. The network topology design problem is formulated as a K-Maximum Spanning Tree Problem with degree bound and has been proven NP-Hard [4]. Owing to the 3-Dimensional integration has the higher packing density and the shorter wire length, they design a novel 3D torus topology along with simplified inter-layer and vertical optical routers, in order to achieve a better performance [5]. The other paper also addresses an NP-hard problem, referred to as network topology design with minimum cost subject to a reliability constraint (NTD-CR), to design a minimal-cost communication network topology that satisfies a pre-defined reliability constraint [6]. For the complex topology network, someone introduces an algorithm to construct a plant topology from analyzing correlations in operations data and iteratively combining pieces of information to the final result [7]. Considering the design of survivable virtual network mapping in multi-domain optical networks, the researchers propose a heuristic algorithm for survivable virtual network mapping using a partition and contraction mechanism (PCM) and based on cut set graph theory with the objective of minimizing the total network link cost [8]. Most topology generation algorithms aim at optimizing the overall topological performance at the expense of other topological performance when concurrently optimizing some objectives. Therefore, there is a demand to improve the balance among all optimized objectives.

2. GA-PODCC Algorithm

2.1. Topology Generation Model

Genetic algorithm utilizes group search to obtain optimal solution, especially for complicated nonlinear problems [9]. The model of multi-objective optimization is given before designing the topology generation algorithm. The objective functions are as follows:

$$\min f_1(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \sum_{s} t_{ij}^{s} \tag{1}$$

$$\max f_2(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} w_{ij} \tag{2}$$

$$\min f_3(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \sum_{s} u_{ij}^{s} \tag{3}$$

x_{ij} represents the situation of link configuration between nodes i and j; $x_{ij} = 1$ if there is an available link between nodes i and j; Otherwise, $x_{ij} = 0$. t_{ij}^{s} represents the delay required for the business request s passing the link (i, j). Therefore, Equation (1) represents the total link delay of the network; w_{ij} represents the bandwidth configured for the link (i, j), thus Equation (2)

represents the situation of static resource configuration. u_{ij}^s represents the bandwidth required for the business request s passing the link (i, j), hence Equation (3) represents the situation of resource consumption. Equation (4) is constraint conditions.

$$\begin{cases} \sum_s u_{ij}^s \le \lambda_{max} \\ \sum_{j=1}^{n} x_{ij} < b_i \\ \sum_{j=1}^{n} x_{ij} < a_j \\ x_{ij} \in \{0,1\} \end{cases} \tag{4}$$

where λ_{max} represents the maximum amount of business requests that a link can carry, so the first constraint represents the sum of business requests of link (i, j) is less than λ_{max}. b_i is the number of optical transmitters owned by node i; The second constraint limits the number of links from node i to other nodes less than b_i, representing an out-degree constraint; Similarly, the third constraint is an in-degree constraint; The last constraint limits the value of x_{ij} to 0 or 1.

2.2. GA-PODCC Algorithm

Based on the above generation model, we propose GA-PODCC algorithm, which optimizes the link delay, resource allocation and resource consumption simultaneously. In order to continuously approach the optimal population and obtain a solution that optimizes the three objectives, the politropism and full search of solutions are achieved through selection and competition of population between parents and offspring.

2.2.1 Gene, Individual and Population

GA-PODCC algorithm needs to initialize a population which consists of many individuals (also known as chromosome). **Figure 1** shows the chromosome's structure, in which each element on the chromosome is a gene of the chromosome and each gene represents the situation of link configuration between two nodes.

Gathering a number of individuals forms a population, the structure of a population is shown in **Figure 2**.

2.2.2. Crossover Operation

The crossover times of the proposed operation is half of the individual number in parent population. **Figure 3** shows a schematic diagram of the crossover operation.
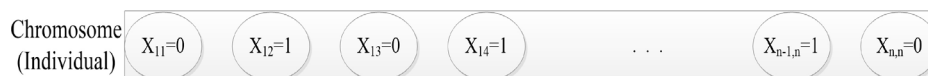

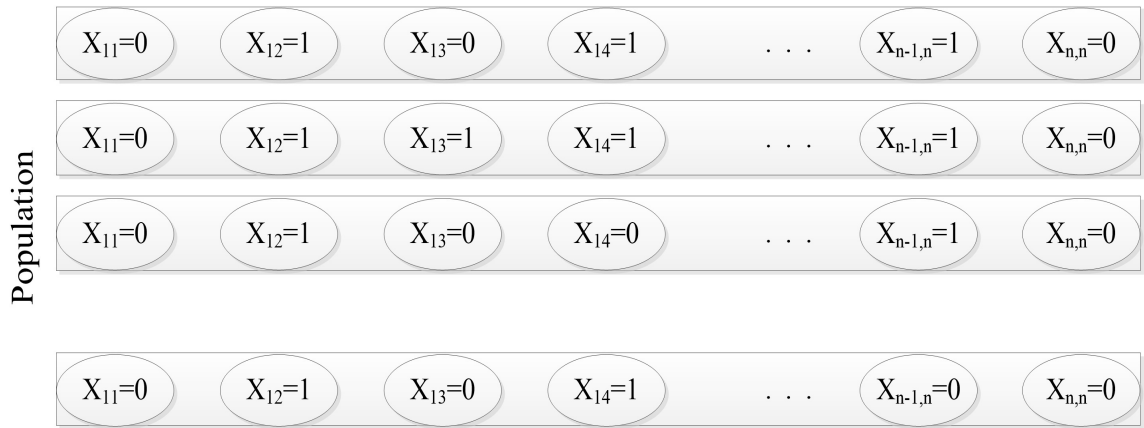
Figure 1. Chromosome's structure.
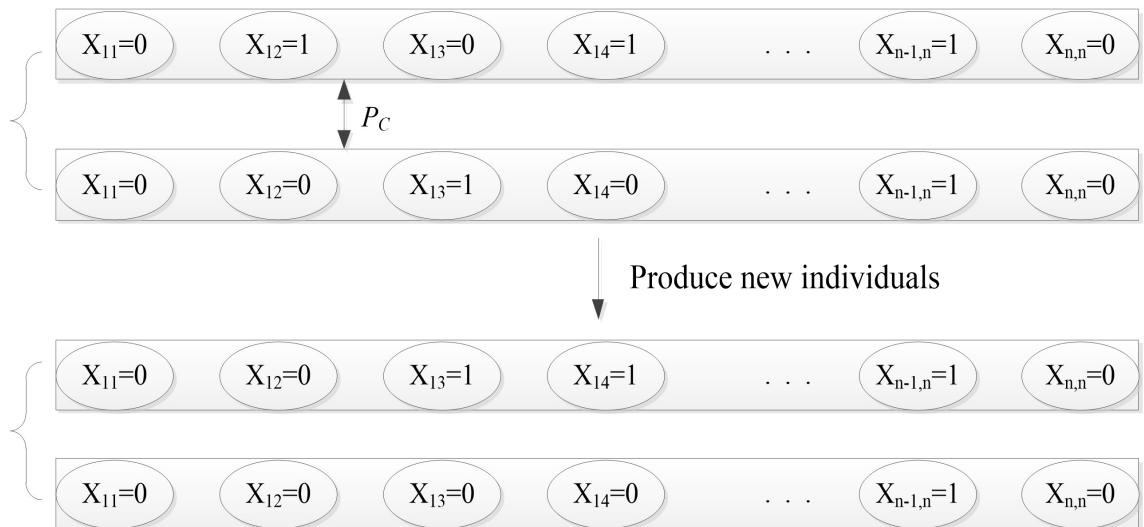
Figure 2. Population's structure.

Figure 3. Schematic diagram of the crossover operation.

The crossover type we adopted is uniform crossover, in which two parent individuals are randomly selected from the parent population, and all the genes on the chromosome are exchanged with a crossover probability p_c. Two new individuals are produced after the exchange. Then repeat the above crossover operation until n new individuals are produced. Equation (5) provides the crossover probability of the ith gene on the chromosome.

$$p_c(i) = \min(p_c) + \{\max(p_c) - \min(p_c)\} \times i / Max\_Gen \quad (5)$$

$\min(p_c)$ is the minimum crossover probability; $\max(p_c)$ is the maximum crossover probability; *Max_Gen* is maximum genetic algebra. After a large number of experimental simulation tests, the three values are set as $\min(p_c) = 0.3$, $\max(p_c) = 0.7$, and *Max_Gen* = 5000, respectively.

The purpose of crossover operation is to produce new individuals, enabling the genetic algorithm to enhance global searching ability. To eliminate the bias about locations and orders of genes on chromosomes, uniform crossover is chosen. Therefore, the search of non-inferior solutions can be improved.

2.2.3. Mutation Operation

Mutation operation is applied after the crossover operation. Combined with the chromosome's structure, the so-called genetic mutation is that gene on chromosome with a certain probability changes from the original 0 to 1, or from 1 to 0. Figure 4 shows the mutation operation.

The main purpose of introducing mutation operation is to provide the genetic algorithm has a characteristic of being easy to fall into the local optimum. Genetic mutation with a certain probability can make the algorithm jump out of the local optimum. When selecting mutation probability p_m, adjustability should be considered. Equation (6) gives the mutation probability of the *i*th gene on the chromosome.

$$p_m(i) = \min(p_m) + \{\max(p_m) - \min(p_m)\} \times i / Max\_Gen \quad (6)$$

We set $\min(p_m) = 0.01$, $\max(p_m) = 0.1$, and $Max\_Gen = 5000$, respectively, providing a low mutation probability and an excellent global searching ability in the initial stage of the algorithm execution. In the middle and later stages of the algorithm execution, the mutation probability becomes higher in order to maintain the diversity of the population and prevent it from falling into the local optimum. Similarly, the mutation type adopted in this paper is uniform mutation.

2.2.4. Selection Operation

Through crossover and mutation operations, *n* parent individuals in the initial population can produce *n* new offspring individuals. Here, the selection operation selects *n* individuals from parent and offspring individuals as the parent individuals of the next generation population P_1, and the remaining *n* individuals are eliminated. In this way, while producing more outstanding individuals, the outstanding individuals of the original parent population are preserved. Figure 5 shows the selection operation.

The key of the selection operation is to determine the selection mechanism, which means that to figure out the design of the fitness function. The design principle of the fitness function follows that the value of individual fitness function is directly proportional to the adaptability. The design of the fitness function needs to consider three objective functions. The first objective function represents the total link delay of the network, which is a minimum objective function. Therefore, the design of the fitness function is to invert the minimum objective function to the maximum objective function, and use the inverted function as the first fitness function Equation (7). Similarly, the other two fitness functions follow the same design.
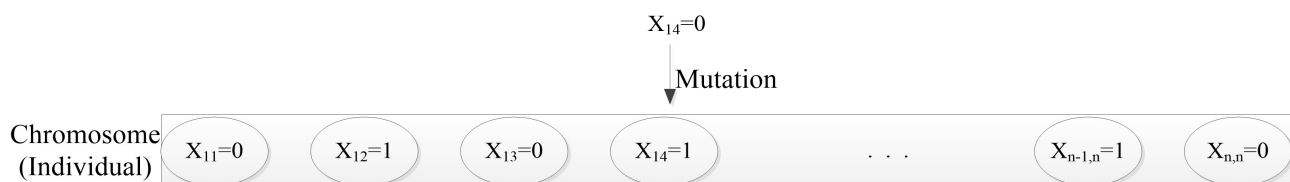

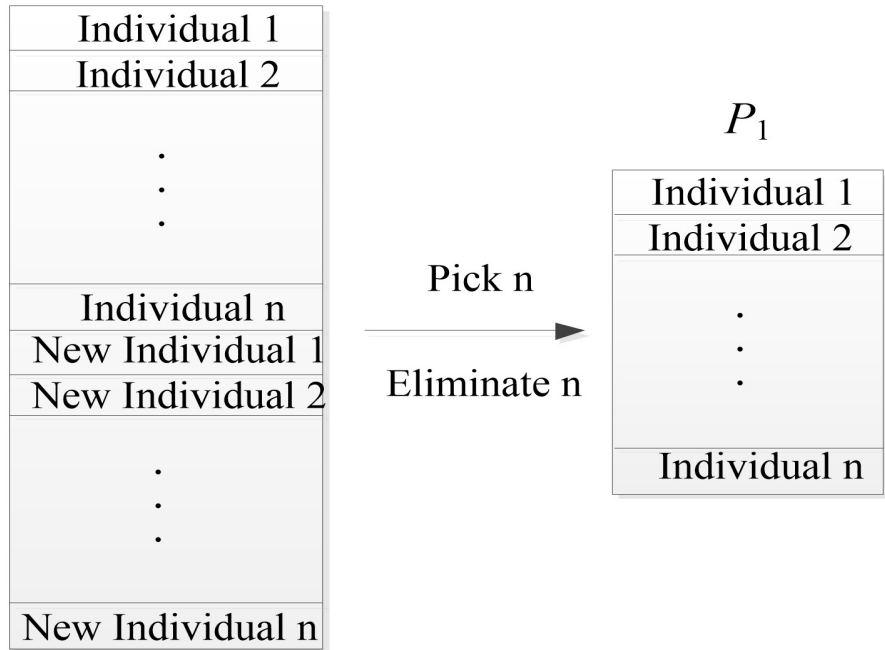**Figure 4.** Schematic diagram of the mutation operation.

Figure 5. Schematic diagram of the selection operation.

$$h_1(x) = -\sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}\sum_s t_{ij}^s \tag{7}$$

$$h_2(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}w_{ij} \tag{8}$$

$$h_3(x) = -\sum_{i=1}^{n}\sum_{j=1}^{n}x_{ij}\sum_s u_{ij}^s \tag{9}$$

2.2.5. Genetic Operation

Genetic operation is the last step of GA-PODCC algorithm. The genetic operation is divided into two stages. In the first stage, the crossover, mutation, and selection operation are performed for internal individuals of the population, producing the next generation population. The first stage of genetic operation is repeatedly executed until the genetic algebra equals *Max_Gen*. **Figure 6** depicts the first stage of the genetic operation.

Subsequently, three sets of the optimal solutions corresponding to different optimization objectives are obtained. There are n optimal solutions for each optimization objective and the amount of total optimal solutions equals to $3n$. Each set of optimal solution is sorted from high to low according to the value of fitness function.

The second stage of the genetic operation is to select one optimal solution that optimizes the three objectives from the $3n$ solutions. The introduction of a new elimination mechanism is required, the average computing time complexity of algorithm is approximately to n^3.

Figure 7 shows the second stage of the genetic operation. First, the optimal solutions of the link delay are tested in fitness function Equation (8). If the test
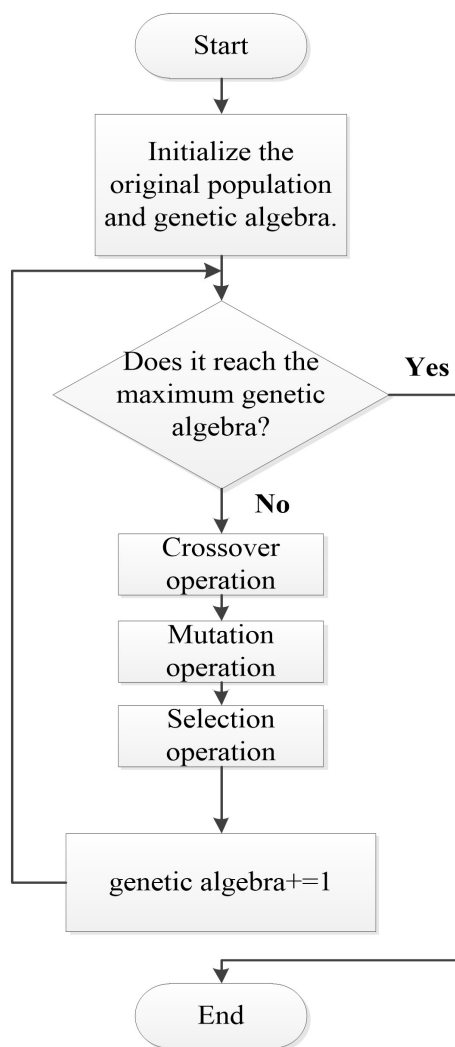
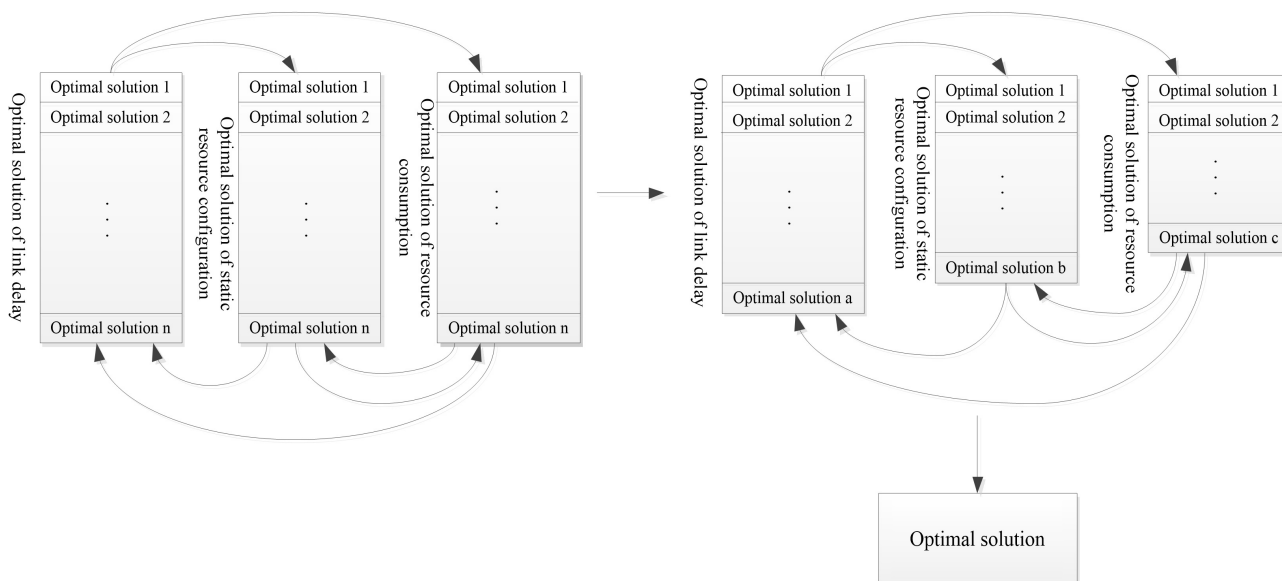**Figure 6.** Flow chart of the first stage in the genetic operation.

Figure 7. Second stage of the genetic operation.

results are better than optimal solution *n* of the static resource configuration, which means the test results pass the test of fitness function Equation (8), corresponding optimal solutions of the link delay are remained, or eliminated. In the same way, the optimal solutions of the link delay are tested in fitness function Equation (9). After the two tests, there are *a* (*a* < *n*) optimal solutions of the link delay. Similarly, some optimal solutions of the static resource configuration and consumption are eliminated. The number of the remaining optimal solutions is *b* (*b* < *n*) and *c* (*c* < *n*), respectively.

The second stage of the genetic operation is repeated until the single optimal solution is obtained. The second stage of the genetic operation takes the performance of three optimization objectives into account. All optimal solutions need to be tested by the other two fitness functions, so the last remaining optimal solution must be the best. After executing two stages of the genetic operation, an excellent network topology can be obtained.

3. Simulation

In addition to the above three performance indexes, a new performance index called generational distance (*GD*) is added to the simulation experiment. Here, *GD* is a convergence evaluation index [10] and is used to evaluate the approximation quality to global non-inferior optimal region. It can be acquired by the formula:

$$GD = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} \tag{10}$$

n is the number of individuals in the non-inferior solution set. d_i is the Euclidean distance between the individual and the global non-inferior optimal region. A smaller *GD* value can lead to a close distance between the population individuals and the global non-inferior optimal region, thus achieve a better performance.

3.1. Experimental Scheme 1

Under the same experimental conditions, the algorithm performance indexes of the GA-PODCC, GA-POCDC, and DC-GALD are compared in the experimental Scheme 1.

The simulation parameters of the network model are set as follows: The number of individuals is 50 and the value of *Max_Gen* is 5000. When genetic algebra equals to 10, 100, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500 or 5000, the corresponding performance indexes are calculated respectively for the experimental analysis. Finally, it shows all the topologies generated in the last generation and chooses the best network topology.

3.1.1. Comparison of the Link Delay
The following conclusions can be drawn from the data shown in Figure 8: In

terms of the link delay, the solutions found by the GA-PODCC algorithm are significantly better than that of the other two algorithms at the beginning. That is to say, GA-PODCC algorithm has a superior searching ability in the link delay. When genetic algebra larger than 1000, the optimal solutions in the link delay of the GA-PODCC algorithm basically remains stable. While the link delay of the GA-POCDC and the DC-GALD algorithms is basically stable only when the genetic algebra is larger than 3000 generations. This is because the GA-PODCC algorithm puts the parent individuals and the offspring individuals together to produce individuals of the next population. Since the space of the selection operation is enlarged, it also guarantees that the elite parent individuals can enter into the next population when offspring individuals are produced.

3.1.2. Comparison of the Static Resource Consumption

The following conclusions can be drawn from the data shown in **Figure 9**: The static resource consumptions of GA-POCDC and GA-PODCC algorithm are almost identical at the beginning, but the GA-PODCC algorithm decreases faster with the increase of genetic algebra and thus converges to the optimal solution earlier. After the genetic algebra is larger than to 200, the static resource consumption of the GA-PODCC algorithm basically remains stable. However, the GA-POCDC algorithm does not stabilize the static resource consumption until
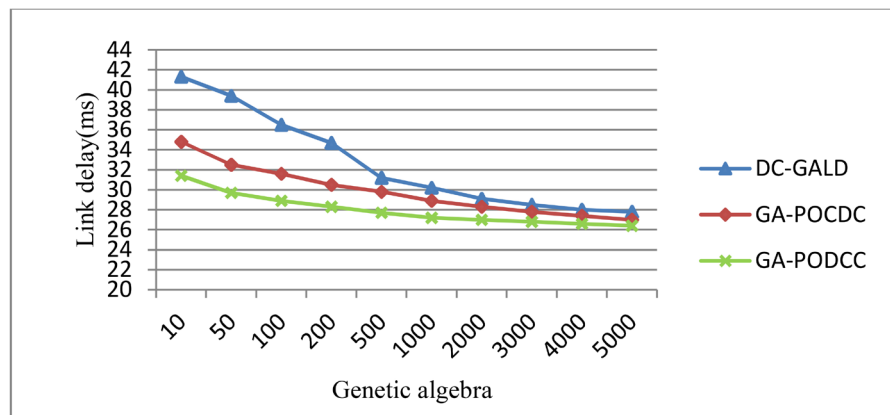

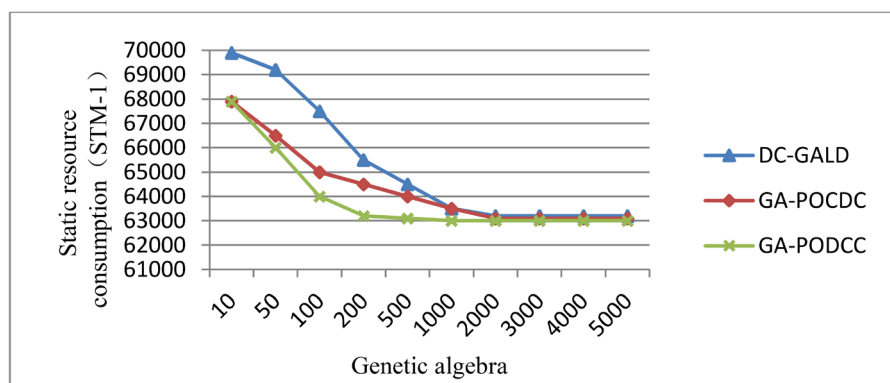**Figure 8.** Comparison of the link delay of three algorithms.

Figure 9. Comparison of the static resource consumption of three algorithms.

the genetic algebra larger than 1000. This is because GA-PODCC algorithm has two stages. Introducing a new elimination mechanism based on the first stage of the genetic operation can greatly speed up the convergence.

3.1.3. Comparison of the GD Value

The following conclusions can be drawn from the data shown in Figure 10: The GA-PODCC algorithm has a smaller GD value than the other two algorithms. Because the GA-PODCC algorithm has a fair fitness allocation mechanism, while the other two algorithms adopt an equal probability to the randomly assign fitness value, the GA-PODCC algorithm has better performance in the convergence evaluation.

From the analysis of the above three performance indexes, the GA-PODCC algorithm has less link delay and static resource consumption, obtaining a more satisfactory convergence evaluation index and a better foundation for routing and survivability strategy of optical network.

3.2. Experimental Scheme 2

Experimental scheme 2 is designed to verify the impact of the topology generation mechanism on subsequent topology aggregation algorithms. We used ML-S algorithm [11] in this experiment. It has two performance indexes: Error Rejected Ratio (ERR) and Crank Back Ratio (CBR). ERR often results in a waste of the available resources of the network and CBR misleads the service requests into an area without meeting the physical path to its QoS requirements.

Because the two important indexes of the aggregation algorithm are used to describe the accuracy of the original network topology information, ERR and CBR can better reflect the influence of underestimation and overestimation distortions on network performance. The formulae of ERR and CBR are as follows:

$$\text{ERR} = \frac{\text{Number of feasible requests rejected}}{\text{Number of overall feasible requests}} \tag{11}$$

$$\text{CBR} = \frac{\text{Number of unfeasible requests accepted}}{\text{Number of requests accepted}} \tag{12}$$

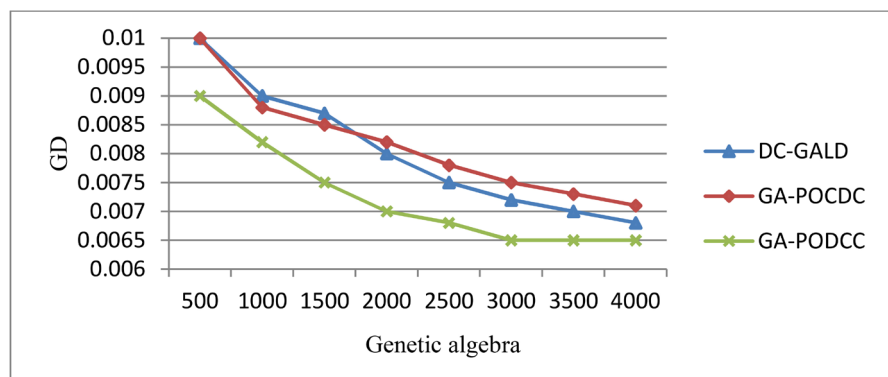The two indicators, ERR and CBR, constitute the description accuracy of the

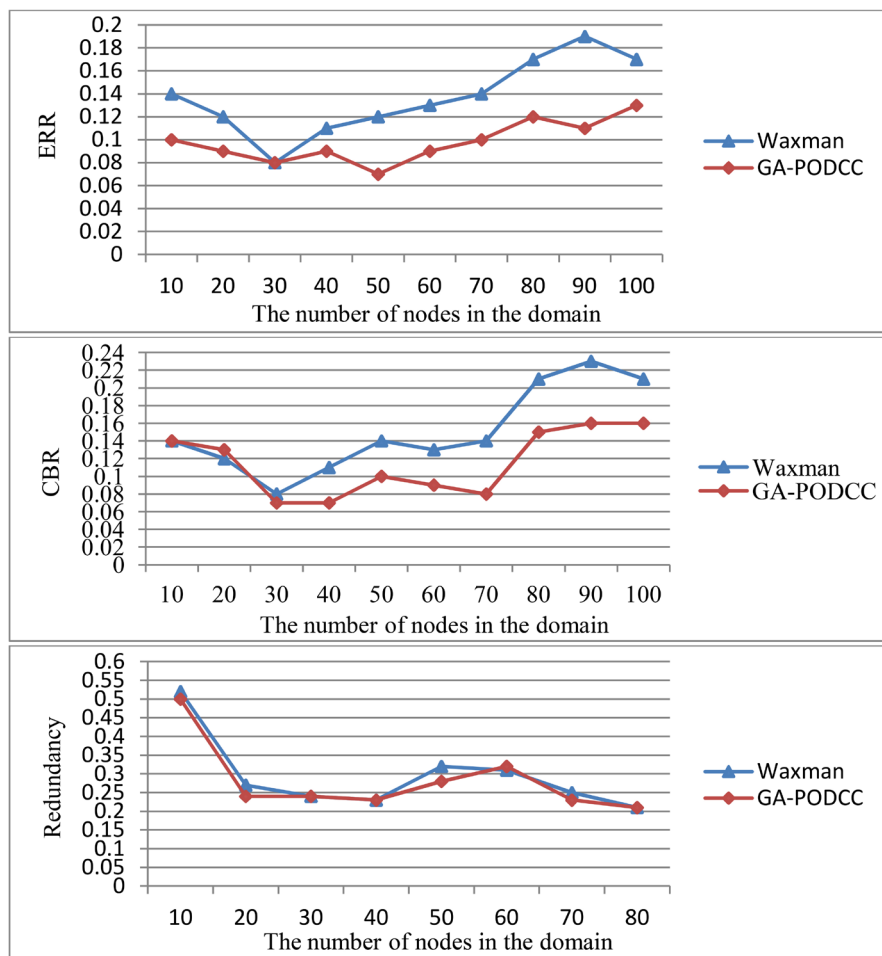**Figure 10.** Comparison of the GD value of three algorithms.

Figure 11. Comparison of the aggregation performance indexes under two different topologies.

network, combine the aggregation degree, and maintain the balance between the aggregation degree and the description accuracy. They are the focus of this study.

The simulation of the experimental scheme 2 is carried out on the random network topology. By comparing the aggregation performance indexes of the network topology generated by the GA-PODCC algorithm and the Waxman model, the Waxman topology is closer to the real network and is more widely used in various simulations of routing algorithm. In Waxman topology, all nodes are randomly distributed in a designated area. Each node is determined according to randomly generated coordinates, and a link is generated between two points with a certain probability.

50 Waxman and GA-PODCC topologies were randomly generated each time, The number of nodes in the domain were selected as {10, 20, 30, 40, 50, 60, 70, 80, 90, 100} in turn. We calculate the average value of the 50 ERR and CBR values as the final experimental data. Between the designated boundary nodes of the network topology, 500 random requests were randomly generated. Under the same network resource allocation conditions, the aggregation performance indexes (ERR, CBR, Redundancy) of the GA-PODCC and the Waxman topology

are compared.

The following conclusions can be drawn from the data shown in Figure 11: The ML-S algorithm based on GA-PODCC algorithm outperforms the ML-S aggregation based on the Waxman topology in ERR and CBR. The index tends to be stable with the increase of the node number in the domain. In terms of the redundancy, their performance is almost identical. That is to say, GA-PODCC algorithm is beneficial to the improvements of the topology aggregation distortion and the overall performance of multi-domain optical network.

4. Conclusion

Aiming at the topology generation problem of multi-domain optical network, we propose a two-stage genetic algorithm GA-PODCC based on a multi-objective model. GA-PODCC provides an optimized link delay, resource configuration, and static resource consumption. Simulation results show that GA-PODCC algorithm achieves a better balance among the three optimization objectives. In addition, compared to the topology generated by the Waxman model in ERR and CBR, GA-PODCC algorithm has a better performance and is more preferable to improve the topology aggregation distortion.

References

[1] Chen, L.K., Li, M. and Liew, S.C. (2015) Breakthroughs in Photonics 2014: Optical Physical-Layer Network Coding, Recent Developments, and Challenges. *IEEE Photonics Journal*, **7**, 1-6. https://doi.org/10.1109/JPHOT.2015.2418264

[2] Sheikh, A.A., Felemban, E., Alhindi, A., Naseer, A., Ghaleb, M. and Lbath, A. (2017) OpToGen-A Genetic Algorithm Based Framework for Optimal Topology Generation for Linear Networks. 2017 13*th IEEE International Conference on Intelligent Computer Communication and Processing* (*ICCP*), Cluj-Napoca, Romania, 7-9 September 2017, 255-262. https://doi.org/10.1109/ICCP.2017.8117012

[3] Gao, C.Y., Cankaya, H.C. and Jue, J.P. (2014) Survivable Inter-Domain Routing Based on Topology Aggregation with Intra-Domain Disjointness Information in Multi-Domain Optical Networks. *IEEE/OSA Journal of Optical Communications and Networking*, **6**, 619-628. https://doi.org/10.1364/JOCN.6.000619

[4] Huang, J.S. and Lien, Y.N. (2017) Topology Design for Multihop Cellular Network. 2017 19*th Asia-Pacific Network Operations and Management Symposium* (*APNOMS*), Seoul, 27-29 September 2017, 284-287. https://doi.org/10.1109/APNOMS.2017.8094129

[5] Guo, L., Hou, W.G. and Guo, P.X. (2017) Designs of 3D Mesh and Torus Optical Network-on-Chips: Topology, Optical Router and Routing Module. *China Communications*, **14**, 17-29. https://doi.org/10.1109/CC.2017.7942191

[6] Elshqeirat, B., Soh, S., Rai, S. and Lazarescu, M. (2015) Topology Design with Minimal Cost Subject to Network Reliability Constraint. *IEEE Transactions on Reliability*, **64**, 118-131. https://doi.org/10.1109/TR.2014.2338253

[7] Gutermuth, G. and Hoernicke, M. (2017) Automatic Generation of Plant Topologies by Analysing Operations Data. 2017 22*nd IEEE International Conference on Emerging Technologies and Factory Automation* (*ETFA*), Limassol, 12-15 September 2017, 1-8.

https://doi.org/10.1109/ETFA.2017.8247670

[8] Hong, S., Jue, J.P., Park, P., Yoon, H., Ryu, H. and Hong, S. (2016) Survivable Virtual Topology Design in Multi-Domain Optical Networks. *IEEE/OSA Journal of Optical Communications and Networking*, **8**, 408-416.
https://doi.org/10.1364/JOCN.8.000408

[9] Wang, Y., Chen, X., Li, J. and Huang, W. (2016) A Genetic-Algorithm-Based Information Evolution Model for Social Networks. *China Communications*, **13**, 234-249. https://doi.org/10.1109/CC.2016.7897547

[10] Liu, L.M., Wang, N.P. and Li, F.C. (2009) Study on Convergence of Self-Adaptive and Multi-Population Composite Genetic Algorithm. 2009 *International Conference on Machine Learning and Cybernetics*, Hebei, 12-15 July 2009, Vol. 5, 2680-2685. https://doi.org/10.1109/ICMLC.2009.5212122

[11] Wang, L., Lin, L. and Du, L. (2017) Design and Simulation of a Topology Aggregation Algorithm in Multi-Domain Optical Networks. *Communications and Network*, **9**, 235-248. https://doi.org/10.4236/cn.2017.94017