Network Planning Using GA For Regular Topologies

Gutierrez Lopez, Jose Manuel; Imine, Mohamed; Madsen, Ole Brun

Published in:
IEEE International Conference on Communications, 2008. ICC '08

Publication date:
2008

Document Version
Early version, also known as pre-print

Network Planning Using GA For Regular Topologies

José Manuel Gutiérrez López, Mohamed Imine and Ole Brun Madsen
Department of Electronic Systems, Network and Security Section, Aalborg University, Denmark
Emails: jgl@es.aau.dk, mi@es.aau.dk, obm@es.aau.dk

***Abstract*—Evolutionary Algorithms (EAs) and Genetic Algorithms (GAs) have successfully been applied to solve constrained problems in network design. Network systems are being designed with fiber optic, as the users' requirements are growing all the time. Selecting a suitable depiction of candidate solutions to the problem at hand is the crucial issue for applying GAs. Well defined network topologies, such as Single Ring, Double Ring and Grid, have been intensively scrutinized. As the number of nodes increases, optimization of the network for a given topology can be challenging for any human or specialized algorithm. For EAs/GAs where the objective is optimization of some characteristics even thought that there is no guaranty that the best optimized solution, better solutions after many runs can be found. These solutions are getting near the optimal one at each time until reaching some satisfaction.**

***Keywords*—Regular Network Topologies, Single Ring, Double Ring, Torus-Grid, Availability, Genetic Algorithm**

I. Introduction

Genetic Algorithms (GAs) have been used in multiple studies to optimize network structures. The purpose of these studies were based on optimizing budget, reliability and diameter of the network [1], [2] and [3].

On the other hand, the goal of this study state the bases to design a tool that not only treats the theoretical properties of the network but also its practical performance.

In this paper, a GA is proposed for solving bicriteria network design problems to optimality, or near-optimality, which are to minimize connection cost and maximize network reliability. The objectives are to minimize the cost needed to design the network systems and also to minimize the average communication delay, considering a robust network.

Constraint topology design problems for telecommunication networks, since the networks size get bigger, have received interest by many researchers, for instance network designers, network analysts and network administrators [4]. The problem of how to efficiently design a network where limitations do exist and objectives can be met, is extremely essential in many real world applications, for instance in the telecommunications, electrical grid network, computer networking, oil and gas lines, water duct system, sewage systems, etc. Usually, large networks consist of an assortment of small networks joined together by means of a backbone. The highest investment on Fibre Optic networks is the civilian construction to install the fibre. Hence, it is preferable that the network topological architecture is composed of a well-known and defined topology where characteristics (e.g. cost, delay, traffic, reliability, availability and security) can be optimized in the same time reducing its total budget cost. The criteria performances of these structures are essential and mostly determined by the network topology used. This paper will only treat three different types of regular topologies as a starting point: Single Ring, Double Ring and Torus-Grid.

Some of the problems found at previous studies were that the GA had to calculate the reliability and the diameter in every solution obtained to be able to find the optimal option. The GA had to calculate all the paths for every pair of nodes to calculate the diameter, which makes the GA´s convergence to an optimized solution more difficult and slower [2].

The use of regular topologies allows, *a priori*, to identify parameters as diameter, number of independent or disjoint paths and average path distances just by knowing the number of nodes involved in the network. Furthermore, future networks will demand higher level of SQoS, that it can be defined as a number of metrics and properties related with the logical structure of the network [4]. The use of the properties of regular topologies allows to determine the level of SQoS with no path calculation at all [4]. These SQoS levels can also be improved using table-free routing methods which are defined for regular topologies and have been proposed as a feasible alternative to the traditional table routing methods [5].

The selection of an optimal network topology is an extremely complex combinatorial optimization problem that can be categorized as an NP-hard optimization problem [6]. Consequently heuristics algorithms that are based on EAs/GAs are recommended to solve network topology design.

The remaining of the paper is organized as follows: the review of the networks design using GAs and the topologies used is introduced in Section II; Section III presents the GA being used where the used **crossover** and **mutation** operators, in addition to the **evaluation function** are described; Section IV deals with network planning and the limitation parameters; Section V describes the simulations; and finally, Section VI concludes the paper. explains the future improvements.

II. GAs and Networks Design, a Background

In this Section background concerning GA and network topologies are breefly explained and some references are given for further information.

Fig. 1. Crossover and Mutation Operations

A. Genetic Algorithm Introduction

Evolutionary Algorithms (EAs) and in particular Genetic Algorithm (GA) have successfully been applied to solve constrained problems with multi-objectives, such as transportation problems [7], production process planning problems [8] and network topology design problems [1]- [3], [9] and [10].

EAs/GAs were investigated for several kinds of encoding methods [8] where most of them can not effectively encode/decode without getting some infeasible solutions that require some repair before being considered. This repair can have the reverse effect on the performance of the EAs/GAs.

Genetic operators have very large influence on the GA performances because of their ability to mimic the process of heredity in the creation of new offsprings at each generation.

Crossover is the main genetic operator. It operates on two parents (chromosomes) at a time and generates offspring by combining both chromosomes' features, see Fig. 1 (a).

In network design, crossover operator plays the role of exchanging each partial route of two chosen parents in such a manner that the offspring produced by the crossover represents a feasibly solution that contains routes from both parents.

Mutation is an operator which produces spontaneous random changes in various genes. A simple way to achieve mutation would be to alter one or more genes. Several mutation operators have been proposed for permutation representation, such as swap mutation, inversion mutation, and insertion mutation, and so on [8] and [9].

The selection is intended to improve the quality of the population by giving the high-quality chromosomes, i.e., a better chance to get copied into the next generation. The selection focuses on the exploration on promising regions in the solution space. A type of fitness-proportional selection adopted is used in the GA implemented.

Picking up a proper description of candidate solutions to the problem at hand is the crucial issue for applying GAs. Once the number of nodes increases ($n > 20$), optimization of the network characteristics (total length, maximum link length, availability, budget, etc) for, even, a known topology can be challenging for any human or dedicated algorithms except perhaps for generalized heuristics algorithms such as EAs/GAs where the objective is optimization of the characteristics of these topologies even thought that there is no guarantee that the best solution can be found, nevertheless the guarantee is that after many runs better solutions can be found [1] and [2].

Fig. 2. Single Ring, Double Ring and Torus-Grid structures

B. Regular Network Topologies

The two main reasons for using regular topologies are:

a) It is possible to define and document well-known parameters and metrics (i.e. number of independent paths) which ease network characterization. Besides, based on well-known metrics it is easy to compare different designs in a proper way.

b) Based on regular topologies it is possible to define topological routing techniques which allow faster transmissions and the reduction of routing traffic within the network [5].

These paragraphs discuss the three topologies implemented. The explanation only treats the structure, for further information about the properties and advantages it is recommended to read the given references for each of the topologies.

Single Ring (SR): The number of nodes, **N**, is any positive integer larger than 2. All nodes in a SR network are connected to two other nodes; thus the nodes in the structure are of second degree. See Fig. 2(a).

Double Ring (DR):It consists of two rings denoted inner and outer rings. These rings each contain the same number of nodes (p); hence the number of nodes, $N = 2{\cdot}p$, is any positive even integer larger or equal to 6. The rings are interconnected by links between each corresponding pair of nodes in the inner and outer ring. The DR network is a degree three network structure. [11], see Fig. 2(b).

Torus-Grid (TG): It is obtained from a rectangular grid network by adding links between opposite nodes at the border grid [12]. The result is a fourth degree structure, see Fig. 2 (c). Depending on the number of nodes, there can be more than one possible grid structure. The rectangular structure, $N = a \cdot b$, that has the minimal average distance is considered first in addition to the maximum link and total length.

III. Genetic Algorithm on Network Planning

A key issue in GAs is how to effectively encode a solution of the network topology problem into a GA chromosome? Knowing the topology used (Single Ring, Double Ring, Torus-Grid), each solution is evaluated in view of the links that the considered topology provides by having the nodes of the solution as inputs to the logical link table. The goal is to find a solution that orders the nodes in such a way that when provided as input to the logical link table, the solution provides minimum total length an the same time minimizes any maximum link length. Each solution is represented by

Fig. 3. Logical Connection Table

a string of N numbers from 1 to N. Random solutions are generated to initialize the GA population.

After the evaluation of all the individual solutions the population is sorted based on the evaluation function from minimum to maximum value. Based on this classification a new population is created using the following operations: selection, crossover, mutation and evaluation function. The selection operation consists on picking the parents of the children solutions that will be part of the next generation. The more an individual has a better solution the more it will be selected for participating in the generation of the next population. The selected parents are then crossover to provide the new individuals with their genetic characteristics based on a cross-over probability, Pr.

Each chromosome contains the list of nodes that constitutes the network. The order/rank of each node is linked to the logical link table to decide its connections to the nodes that are its neighbors in the topology in question, see Fig. 3.

Many critical issues are to be carefully considered when designing an appropriate method to build a GA that can, easily, solve the problem. Several kinds of classification of encoding methods can be considered, such as a) Characteristic Vectors-based Encoding; b) Edge-based Encoding; and c) Node-based Encoding. Applied to such chromosomes, positional crossover and mutation operators will generate infeasible solutions, requiring again penalization or repair. The GA implemented considers only node based encoding.

The selection used herein is a combination of the roulette wheel and elitist approaches, in order to enforce the GA to freely search solution space. The roulette wheel selection, considered as a fitness-proportional technique, is applied to arbitrarily replicate new generations and the elitist method is utilized to conserve the fittest chromosome for the next generation. The selection process, help maintain the best chromosomes from the current generation to the next one.

Multi-point crossover (or uniform crossover) is used in the implemented GA, this type of crossover is achieved by picking two parent solutions and randomly taking a component from one parent to form the corresponding component of the offspring. The remaining of the chromosome is taken from the other parent taking in consideration that new alleles are added only if they were not already the randomly chosen from the first parent (see Fig. 1 (a)). This way of doing the crossover will not require any additional repair strategy, to avoid any illegal chromosome or to modify the newly constructed chromosome.

Swap mutation was use in the implemented GA; this operation simply selects two positions in the chromosome at random and swaps their contents (see Fig. 1 (b)).

Evaluation function considered in this GA implementation consists of a multi-value optimization: minimize the total length of the network (budget), minimize the maximum link and optimize the number of 9s in the availability value.

IV. Network Planning

In this Section the given network parameters to the GA to optimize the solution are explained. The first two parameters chosen (among many others that can be added in future work) are the *Availability* and *Budget*.

A. Limitation Parameters

Each one of these limitation parameters involves different essential properties when planning a network:

Availability: Is a value related to the tolerance to failures of the network. At other GA studies [3], the value analyzed, was the reliability. The difference is that the reliability does not consider the maintenance of the network. Hence, by testing the availability, the optimization can be more realistic and the maintenance budget can be estimated. This value is related with the longest logical path, the number of independent paths of the network and longest physical link. The way this value is calculated gives the worst case possible. Therefore, it can be assumed as the limit to guarantee in all possible transmissions. There are ways to improve the accuracy of this value. In future improvements instead of using just the longest link the idea is to use a combination of the longest links in the network.

There are three steps to calculate the availability:

1) *Availability of the links*, A_l: The value is calculated using Formulas (1) and (2).

$$A_l = \frac{MTTF}{MTTF + MFT} \tag{1}$$

$$MTTF = L_{Max\_physical\_link} * FITS \tag{2}$$

By calculating MTTF (Mean Time To Fail) using the value of the maximum physical link, it can be assumed that any link in the network will have an availability that is greater or equal, but never less.

2) *Availability of the path*, A_p: This value is related with the topology used. The use of regular topologies gives by direct operation the maximum logical length of the paths (MaxHopsForThePath) [11] and [12], it is unnecessary to calculate the paths from all the nodes to all the nodes, which is unavoidable for irregular topologies to find the diameter. Formula (3) is used to calculate the path availability.

$$A_p = A_l^{MaxHopsForThePath} \tag{3}$$

3) *Total Availability*, A_t: This value is calculated by Formula (4), being *"n"* the number of independent paths, given by the degree of the nodes at each structure. The standard availability required by most of the telecommunications providers of optical fibre communications is *five* 9s (0,99999) [13].

$$A_t = 1 - \prod_{i=1}^{i=n} (1 - A_{pi}) \tag{4}$$

Budget: This value, in this study, only covers the expenses for the civilian construction (ditches, ducts, etc). This parameter is related with the total length of the digging necessary to install the fibre for the network. It is assumed that all the links are physically independent to each other to be able to guarantee at any situation the total independency of the different paths. Hence, the total length of the ditches is calculated by the sum of all the links "N_L", see Formula (5). The budget can be easily calculated by Formula (6) after obtaining the total length of the ditches. The value of "X" is the price of civilian construction.

$$L_{TotalLinksLenght} = \sum_{i=1}^{i=N_L} L_{linki} \tag{5}$$

$$Budget = L_{TotalLinksLenght} * X \tag{6}$$

A deeper explanation of the formulas can be found in [13] and their practical use at [14]. The given default parameters are used for the simulation of the scenarios. At the implementation of the tool these parameters can be introduced by the users depending on the real values at each case and deployment are conditions. For example the value of the civilian construction is not the same at every country, it must be a variable value.

The values of the length of the links are calculated as straight line. The real value of those link can be estimated by the statistic or mathematical methods with the ratio straight link-road link, a complete explanation of this topic can be found at [15]. So far the tool is implemented using the straight links, the road factor will be added in future work.

This whole idea of combing GA optimization methods with regular network topologies has big potential on wireless networks as well. Considering that the link implementation is not as critical as in wired networks since there is no need of investing on civilian constructions (ditches), the wireless networks can be organized and optimized to improve their performance such as power, link synchronization, etc . The limitation parameters and constraints are different on wired networks, but with proper modifications, the decision criteria can be easily implemented. In this case only the wired networks are studied and simulated but the wireless networks could be an interesting scope for this tool as well.

V. SIMULATION

This Section illustrates the previous ideas at real geographical locations. The scenario chosen is to interconnect 28 of the most important cities in Europe. The scenario could seem not realistic due to the dimension of the project, many factors can influence an international network such as political agreements, and also in reality this network probably would be implemented in different hierarchal levels. The scope of this simulation is to illustrate the mechanism of the tool at its first step to obtain results for continuing improvement. The procedure of the simulation is very simple. The coordinates of the cities and the logical connection tables are given as inputs. Then, the GA, depending on the goal function desired, returns the potential solution. Fig. 4 illustrates the three results obtained depending on the constraints when planning a network. Fig. 4(a) shows the result when the lowest possible budget is required. The resulting topology is a SR. This SR and the links forming it are optimized in terms of physical distance. The use of a SR topology for this network does not meet the optical communications availability levels. The maximum availability using a SR structure is 0.999, or *three* 9s ($1 - 10^{-3}$), which is far away from the required value. Fig. 4(b) shows the result when the lowest possible budget is required but the availability levels must be acceptable. The result is a DR. In this case the budget has increased from the one obtained at the SR case due to the third link implementation at every node. The result is an availability of 0.99999, *five* 9s ($1 - 10^{-5}$), just enough to be accepted as an optical communications network. Probably in the future the required performance of the network will be much more restricted, therefore, the last test is run the GA with no budget constraint trying to optimize the availability of the system. The result as expected is the TG, Fig. 4(c), which obtains a extremely high availability, 0.999999 or *six* 9s ($1 - 10^{-6}$). The improvement is not worthy due to the long links needed to complete the TG and the relatively small number of nodes at this huge area, but a result is always good to be able to discard possibilities.

Table I presents the results of the three scenarios in terms of availability, budget, maximum link length (Link represented with dotted line at Fig. 4) and total length.

Taking a closer look to Fig. 4 we can see that there are some implemented links at the SR solution which would not be implemented at the DR . The proposed solutions, depending on the scenario, are the same topologies as the ones obtained by the simulation, but with a small difference. Even though the requirements are met, there is a possibility that in the future the network will need to be upgraded due to the increment of the requirements of the services and applications. Therefore, the network can be planned starting from a better configuration solution, e.g. if only a SR is required, the DR is considered. When implementing the network, the SR will be installed using some of the links that form the DR solution. In case of an upgrading demand, the modification will be cheaper and with better characteristics than if the DR was never considered. Of course we cannot forget that the SR installed would be more expensive than the one obtained with the simulation. The decision will be a matter of future expectations or believes of the engineers. The tools can help and facilitate the design tasks but we can never forget human factors.

(a) Single Ring Structure

(b) Double Ring Structure

(c) Torus-Grid Structure

Fig. 4. European Simulation

VI. Conclusion

This GA method used for the planning of wired networks using regular topologies has achieved the following results:

a) Determination of the minimum cost for the civilian construction for the obtained topology with given performance.

b) SQoS level of the communications guaranteed at all possible situations, between any pair of nodes in the network.

c) The minimum characteristics and performances (budget, availability, etc..) can help the financing parties of network construction (such State, City, etc..) to make judgmental decision concerning any offer from network construction parties (such as infrastructure companies, etc) and vise versa to make better negotiation before doing any complex designs (such as those that include roads infrastructures for example).

	Single Ring	Double Ring	Torus-Grid
Total length (km)	15723	27040	47375
Max Link (km)	1414	1824	3764
Budget (Millions of €)	314.5	540	947
Availability	0.999	0.99999	0.999999

TABLE I
RESULTS FOR SINGLE RING, DOUBLE RING AND TORUS-GRID

d) The proposed topologies (SR, DR and TG) allow us to precisely get the well known parameters and metrics (diameter, number of independent possible path, average distance, etc...) to be able to compare them and characterize the network depending on the constraints.

e) Regular topologies allow more precise topological routings without the need of routing tables to be defined in advance or additional header in the transmitted data packet that can result in faster communications.

The method has been tested as a possible future tool for the wired network planning. The bases have been established but it is still a lot of work to be done. For this idea to succeed many future ideas must be develop be able to design a real and effective networking tool.

Acknowledgment

The authors would like to thank Rasmus Hjorth Nielsen and Muhammed Tahir Riaz for their help on improving this work.

References

[1] Bassam Al-Bassam, Abdulmohsen Alheraish, Saad Haj Bakry, “A tutorial on using genetic algorithms for the design of network topology”.
[2] Anup Kumar, Rakesh M. Pathak, M. C. Gupta, “Genetic algorithm based approach for designing computer network topology”.
[3] Abdullah Konak and Alice E. Smith “Designing Resilient Networks Usign a Hybrid Genetic Algorithm Approach” GECCO´05, Washington, USA.
[4] J.M. Pedersen “Structural Quality of Service in Large-Scale Network”. PhD Thesis. Control Departament, Aalborg University. April 2005.
[5] J.M. Pedersen, T.P. Knudsen, O.B. Madsen “Topological Routing in Large-Scale Networks”.
[6] Jong Ryul Kim and Mitsuo Gen, “Genetic Algorithm for Solving Bicriteria Network Topology Design Proble”,
[7] M. Gen, A. Kumar, and R. Kim, “Recent network design techniques using evolutionary algorithms”
[8] L. Lin and M. Gen,“Bicriteria network design problem using interactive adaptive-weight GA and priority-based encoding method”
[9] Kumar, A., R., M. Pathak, and Y. P. Gupta, “Genetic-algorithm based reliability optimization for computer network expansion”
[10] Dengiz, B., F. Altiparmak, and A. E. Smith, “Efficient optimization of all-terminal reliable networks using evolutionary approach”,
[11] T. Jørgensen, L.Pedersen and J.M. Pedersen“Reliability in single, double and N2R ring network structures” CIC’05 Las Vegas, USA, Pag. 2-4.
[12] G. Barrenechea, B. Beferull-Lozano, and M. Vetterli, “Lattice sensor networks: Capacity limits, optimal routing and robustness to failures”.
[13] Wayne D. Grover, “Mesh-Based Survivable Networks, Options and Strategies for Optical, MPLS, SONET and ATM Network” Prentice Hall PTR, 1st edition (August 14, 2003).
[14] Jose Gutierrez, Ruben Cuevas, Jens Pedersen, Ole Madsen “Backbone structure for a future multipurpose network”. ICACT 08, South Korea.
[15] J. Fernandez, P. Fernandez, and B. Pelegrin. “Estimating actual distances by norm functions: a comparison between the $l_{k,p,\theta}$-norm and the $l_{b1,b2,\theta}$-norm and a study about the selection of the data set”. Comput. Oper. Res., 29(6):609−623, 2002.