NEW GENETIC MULTICAST ROUTING ALGORITHM FOR AD-HOC NETWORKS

PIOTR SZYMIAK, MACIEJ PIECHOWIAK, PIOTR ZWIERZYKOWSKI

Chair of Communication and Computer Networks
Poznan University of Technology
Polanka 3, 60-965 Poznan, POLAND
mpietch@ukw.edu.pl, pzwierz@et.put.poznan.pl

Abstract. Real-time multimedia applications in computer networks require techniques of providing data to a group of users at the same time. Building efficient multicast trees is a challenge for multicast routing protocols and algorithms. The article focuses on the implementation of genetic multicast algorithms in ad-hoc networks. It also proposes a new approach in ad-hoc networks modelling. Tests were carried out for a wide range of parameters defining properties and characteristics of test networks topologies.

Key words. Multicast, genetic algorithms, ad-hoc networks.

1 Introduction

The multicast routing problem is known in literature as the Minimum Steiner Tree Problem, which has been shown to be $NP$-complete for one and more link parameters (metrics) [6]. Literature presents many efficient heuristic solutions solving this problem in polynomial time [7, 8, 13, 18]. In recent years, the genetic algorithms (GA) have become the field of interest for many researchers, because of their robust and efficient search in complex space. Researchers particularly focus on the adaptation of genetic operators in multicast routing algorithms [1–4]. The advantages of genetic algorithms are commonly known. GAs code solutions as bit strings, thus difficult problems can be easily solved using long strings. Genetic operations, such as crossover and mutation, are easy to implement. The main advantage of GA is the parallelization of code - with a pool of chromosomes, GAs search the solution space at different corners in parallel. Randomized genetic operations, such as mutation, can keep the search from being trapped by local optima [1]. The common goal of the above-mentioned multicast algorithms (both heuristics and GAs) is their implementations in networks that represent real Internet topologies (at domain level and AS-level). This area is well-studied and the influence of network topology on the efficiency of QoS multicast heuristic algorithms has been widely examined [19, 20]. Following the increased interest in the transmission in wireless networks, there is also a need to analyse the efficiency of routing algorithms dedicated mainly for ad-hoc networks. These are wireless networks of decentralized structure, where mo-
bile nodes can function both as a client (end terminal) and a router. The models of network structures reflecting the real topologies of ad-hoc networks have not been studied in great detail. However, Li and Yu carried out some research on the properties of such networks [9]. Therefore, there is a need to create a tool which will allow to generate ad-hoc networks for algorithm analysis, which allow to compute them in similar topologies.

The article discusses the effectiveness of GA algorithms in ad-hoc networks. Chapter 2 presents a network model. Chapter 3 is an overview of implemented algorithm. Chapter 4 includes the results of the simulation of the implemented algorithms along with their interpretation. The final chapter sums up the discussion.

2 Network model

2.1 Problem formulation

It is assumed in the article that the network on which the algorithms will be tested is represented by a coherent undirected graph $G = (V,E)$, where $V$ is a set of nodes and $E$ a set of links between the nodes of the network. The existence of the link $e = (u,v)$ between the nodes $u$ and $v$ effects in the existence of the link $e' = (v,u)$ for any number of $u,v \subset V$ (the equivalent of the bi-directional links in communications networks). Each link $e \subset E$ is associated with two parameters: cost $C(e)$ and delay $D(e)$. The cost metric describes the general cost of setting up a connection with the execution of a given link. The cost function can be composed of: the flow capacity of the link, capacity of the buffers, error rate (link quality), etc. The delay metric represents the delay in the propagation of the signal carried by a given link.

In research studies the Constrained Shortest Path Tree (CSPT) is commonly used. It contains constrained shortest paths between the source and each destination node. The CSP problem (Constrained Shortest Path), also known as DCLC (Delay Constrained Least Cost), can be stated as the problem of minimizing: $z^* = \min \sum_{i,j} c_{ij} e_{ij}$, subject to: $\sum_{i,j} d_{ij} e_{ij} \leq \Delta$, where: $e_{ij} \in \{0,1\}$ [5].

The problem of finding a Steiner tree is a $NP$-complete problem [6], regardless the number of constraints. This means that finding a Steiner tree in large graphs is possible only through the application of suboptimal algorithms.

2.2 Propagation model

Ad-hoc network topology is analyzed in many works, including [12] and [11]. These publications provide detailed analyses on modeling topologies for ad-hoc networks as well as sensor networks, methods for controlling topologies, models of mobility of nodes in networks and routing protocols in wireless ad-hoc networks. Wireless ad-hoc networks are formed by devices that have mobile energy source with limited capacity. It is essential then for the energy consumption to be maintained at a possibly low level in order to prolong the time duration of autonomous operation of the device.

The adopted model of the costs of links between the devices takes into consideration energy used by the antenna system of a device. The proposed implementation assumes that network devices have isotropic radiators. The power of electromagnetic wave $P_r$ received by the antenna can be expressed by the following dependency:

$$P_r \sim \frac{P_s}{d^\alpha},$$

where $d$ expresses the distance between the transmitter and the receiver, and $P_s$ denotes transmitting power. If radiation propagates in vacuum, then $\alpha = 2$. However, in real environment $\alpha \in \langle 2, 6 \rangle$ [12]. In the present investigation, the value $\alpha = 3.5$ was adopted. This value was calculated as an arithmetic mean from the middle ranges of the variability of the parameter $\alpha$, published in [12] and [11]. The required power of the received electromagnetic wave $P_r$ was adopted as
constant. These assumptions determine the following dependency between transmitting power and distance:

\[ P_s \sim d^{3.5}. \]  

(2)

For the purposes of the study, it is assumed that future, far more advanced, devices will have the capability of precise fine tuning of the transmitting power level to that required by the receiver.

2.3 The scenario generator for ad-hoc network simulations

In order to compare effectiveness of the algorithms for constructing multicast trees in graphs with topologies similar to those used in wireless ad-hoc networks, the scenario generator for appropriate simulation was worked out. The process of topology generation can be divided into three stages.

The first stage involves the construction of a square grid \( N = \lceil \sqrt{n} \rceil^2 \) with evenly distributed nodes, where \( n \) is the target number of nodes. In the next step, the superfluous nodes are removed in a pseudo-random way. Then, the nodes are joined by edges depending on the distances between them and the ranges of the nodes \( r \in \mathbb{N} \land r \in (1, 4) \).

The adopted value of maximum range of links results from the necessity of providing nodes obstacle-free mobility one the one hand, and securing high resistance to damages and providing coherency to the whole of the network, on the other.

The establishment of links between nodes depends on the decision making process (maximum values of ranges of both nodes). This random element is used to simulate obstacles in the environment that block transmission in the assumed directions.

The last step is to assign appropriate costs and delays to the edges. Figure 1 shows the ways of constructing links with different ranges and an exemplary network generated by the discussed generator.

The arrangement of evenly distributed nodes in a network on the surface usually does not correspond much to the structure of a real network. To simulate a random arrangement (placement) of devices combined with unforeseeable anisotropy of environment, the value of transmitting power, calculated in Equation 2, is modified by a pseudorandom factor that scales the result within the range \((50\%, 150\%)\), which diversifies values of the required transmitting power. The ultimate form of the link cost is expressed by the following dependency:

\[ C = \text{rand}(0.5; 1.5) \cdot d^{3.5}. \]  

(3)

Delay in the wireless link is dependable on the delay imposed by the queueing process, switching and transmission in the source node, the transmission channel and by the receiver. Delay imposed by the transmitter and the receiver has a constant value, whereas delay in transmission medium is variable and depends on the distance between devices. In the projects concerning future standards in wireless transmission it is postulated that SC-FDMA modulation should be used for transmission in the direction from mobile devices [17, 21]. The SC-FDMA modulation is characterized by a long processing route, so a delay imposed by a transmitting medium is relatively not important. The above allows us to derive the following formula for the link delay:

\[ D = 8 + \text{round}(d). \]  

(4)

The dependency determined in Formula 4 implies that delay will adopt values within the range \((10, 12)\).

3 Proposed algorithm

The proposed algorithm (Algorithm 1) is characterized by a far bigger exploration capability that allows to generate trees with significantly lower costs than those resulting from the application of the algorithms known from the literature, such as GADVM [3] or WWH [15, 16] or WWH [15, 16].
The number of the simulations carried out in the study made it possible to select the following set of parameters [14]:

- random reproduction with minimal diversity of alleles for at least 40% of genes,
- uniform crossover with the probability of 100%,
- mutation on one randomly-selected gene at the maximum, with the probability of 65%,
- each offspring that was a copy of any of the chromosomes from the base population, was replaced by a random combination of alleles,
- elitist succession, the size of the elite was 50% of the base population,
- number of paths to each of the destinations was 34,
- evolutions with the participation of 100 organisms through 500 generations.

The operation of the algorithm is illustrated by its pseudocode (Algorithm 1).

4 Simulation results

The proposed genetic algorithm was put to a comparative performance study with a number of reference tests with genetic algorithms presented in the relevant literature.

The tests were carried out with the application of the proposed ad-hoc generator.

Each of the genetic algorithms was applied within the solution space that assumed:

Figure 1: Graphs representing networks of 20 devices, in which all feasible links with the range \( r = 1 \) (a), \( r = 2 \) (b), \( r = 3 \) (c) and \( r = 4 \) (d) are denoted. Graph e shows an exemplary network of 20 devices constructed with the application of the ad-hoc network generator [14].
**Algorithm 1** Proposed GA for Ad-hoc networks

1: generate 34 path delays that satisfy criteria for each of the destinations...
2: ...with the application of \( k \) Dijkstra shortest paths algorithm
3: number the generated paths
4: \( p \leftarrow \) number of organisms
5: for \( i = 1 \) to \( p \) do
6: generate genetic code in a pseudo-random way
7: determine the value of its fitness function
8: end for
9: for number of generations do
10: for \( i = 1 \) to \( p/2 \) do
11: single out two organisms
12: \( j \leftarrow 0 \)
13: while organisms differ in at least 40\% genes \( \land j < 4 \) do
14: single out the other organism
15: \( j \leftarrow j + 1 \)
16: end while
17: make a copy of the singled out organisms
18: perform uniform crossover on copies
19: if mutation probability condition satisfied then
20: substitute random gene of the first copied organism with a random allele
21: end if
22: if mutation probability condition satisfied then
23: substitute random gene of the second copied organism with a random allele
24: end if
25: end for
26: for \( a = 1 \) then \( p \) do
27: for \( b = 1 \) then \( p \) do
28: if derivative organism No. \( a = \) base organism No. \( b \) then
29: generate chromosome in a pseudo-random way and substitute the derivative organism No. \( a \) with it
30: \( b \leftarrow p \)
31: end if
32: end for
33: end for
34: determine value of the fitness function of organisms of the derivative generation
35: determine elite with the population of 50\% of the base generation
36: save \( p \) best fitted organisms from the elite and the derivative population
37: remove the remaining chromosomes
38: end for
39: return best-fitted individual

- 2 paths for each of the receivers (for the WWH algorithm),
- 8 paths for each of the receivers (for the GADVM algorithm),
- 34 paths for each of the receivers (for the authors’ algorithm).

The investigation was carried out using the following set of fixed parameters:
- randomly-selected destination nodes and the source node,
- maximum delay \( \Delta = 150 \),
- range of links between the nodes in the networks within \( r \in (1, 4) \),
- evolution using 100 organisms through 500 generations.

The first stage of the investigation involved the determination of the dependency between the costs of tree and the number of nodes in a network (Fig. 2).

![Figure 2: The dependency between the cost of tree and the number of nodes in the network with the fixed number of destination nodes (\( m = 20 \)).](image-url)

The imposed constraint of maximum path delay considerably influences the quality of results obtained with the application of the algorithms in question in networks with more than 100 nodes, which introduces a necessity of choosing an increasing number of paths with minimal delay in place of paths with minimal costs. Regardless of
the number of nodes, the proposed genetic algorithm allows us to obtain the lowest cost of tree. The next investigation study involved a determination of the influence of the number of receiving nodes on the cost of tree (Fig. 3).

Figure 3: The dependency between the cost of tree and the number of destination nodes with the fixed number of nodes in a network.

An increase in the number of destinations is followed by the increase in the solution space that genetic algorithms have to investigate in order to find optimal combination of numbers of paths. The absence of compensation, through an increase in the number of organisms or duration prolongation in of evolution processes, results ultimately in obtaining solutions that are less mature. The cheapest trees in this study were available after the application of the proposed genetic algorithm.

The next step in the comparative tests was the analysis of the influence of the maximum admissible delay on the cost of tree (Fig. 4). It is easy noticeable that a determination of the maximum delay for less than 100 results in a substantial increase in the costs of the generated trees.

The next stage of the study was to examine the influence of the average node degree [19] on the total cost of the multicast tree (Fig. 5). A change in the node degree was effected by a modification in the range of admissible ranges of links between nodes. When connections with range equal to 1 were only permitted, then the generated networks had average node degree equal 3.6. However, when the range was changed to \( r \in (1, 2) \) – the average node degree was 6.03. A situation in which links within the ranges \( r \in (1, 3) \) were permitted resulted in generating a network with the average node degree equal to 8.23. The adoption of a default range \( r \in (1, 4) \) resulted in obtaining the average node degree equal to 11.45.

Additionally, the influence of the arrangement of the destination nodes on the cost of tree was investigated (Fig. 6). The cost was expressed as the function of the cost of tree from the spreading...
factor of the nodes [10]. The source node was determined from among those nodes for which the Euclidean distance between them and the focal point of the area of the distribution of the destination nodes was maximum. The area of the distribution of nodes was circle-shaped with the middle point in the central part of the network. The side length of the square surface area of the network was 1000.

Observing the results presented in Fig. 6, it is easy to notice that the bigger the area from which destination nodes of transmission are determined the bigger the cost of multicast trees. The adoption of the radius larger than 520 causes the algorithm to construct paths with minimum delay instead those of minimum cost, which is followed by a disproportional increase in costs of the obtained trees.

The next stage of the simulation tests involved the analysis of the influence of the genetic parameters, i.e. the number of organisms (Fig. 7) and the number of generations (Fig. 8).

The presented results show the accuracy in the selection of the solution spaces in the proposed genetic algorithm. Even a three-fold increase in the number of organisms is not followed by any significant improvement in the obtained results. The results, in the form of mature individuals, are achieved as early as the stage of the nominal number of organisms. A decrease in the number of organisms results, in turn, in a deterioration in the obtained results. However, in the case of the WWH algorithm, incomparably small solution space can be effectively searched with the participation of as few organisms as 20. The dependency of the cost of tree on the duration length of the evolution process is shown in Fig. 8.

In this case, the incomparably small solution space has major influence upon the results provided by the WWH algorithm. The worst results in the first generation are yielded by the proposed...
algorithm with its searches of optimal solution in the biggest solution space. However, between the twentieth and the thirtieth evolution cycles, the genetic algorithm provides trees with the lowest costs.

In the presented graphs each point represents the average results from 500 results obtained with the application of the genetic algorithm, with 10 runs in each of the 50 generated networks. The results of the simulations are shown in the charts (Figs. 2–8) in the form of marks with 95% confidence intervals that were calculated after the \( t \)-Student distribution. 95% confidence intervals of the simulation are almost included within the marks plotted in the figures.

5 Conclusions

The article presents the new multicast routing algorithm based on genetic operations. The efficiency of the proposed genetic algorithm is then compared with the efficiency of classical genetic algorithms emphasizing the quality of a network model (precision and accuracy in illustrating real ad-hoc topology).

To analyze the effectiveness of the proposed genetic algorithms, their dependencies on the parameters determining the network topology are explored and discussed.

The results, obtained during the simulation experiments, prove the accuracy of the proposed solution, which allows to reach better results than those obtained with the application of genetic algorithms known form literature.

References


