Abstract—The article investigates representative heuristic algorithms finding the cheapest spanning trees between a source node and a group of destination nodes (multicast connections). The key part of the article includes the efficiency analysis and the influence of the parameters of a given structure (network model generated by BRITE tool) upon the efficiency of the algorithms under scrutiny.

I. INTRODUCTION

MULTICASTING is a way of transmission between a source node and the specified destination nodes - members of a multicast group. In practice, however, there is a given group of nodes receiving the same data that run parallel in the same time. Multicasting requires efficient routing algorithms defining a tree with minimal cost between the source node and particular nodes representing the users. Such a solution prevents duplication of the same data (packets) in the links of the network. Routing of the sent data occurs only in those nodes of the network which lead directly to destination nodes.

To ensure accurate transmission of real time data, a particularly defined bandwidth is needed and, primarily, a situation in which the delay between the source and destination nodes can be maintained at the steady, unchanged level. The occurring jitter phenomenon is an undesirable phenomenon here.

Popular multicast IP routing protocols solve the least-delay routing problem [1]. These protocols are characterized by acceptable time complexity (O(N log N) for MOSPF protocol [2]). They do not take into account quality parameters (i.e., the maximum delay) minimizing path’s cost (and delay) between the sender and each destination node.

To ensure reliable transmission, multimedia applications set up high standards for quality parameters (Quality-of-Service parameters). Quality requirements concerning, inter alia, steady and guaranteed delay values in relation to a transfer of a packet between the source and destination nodes along a defined path in the network still pose a great challenge for those designing the application in real time. Hence, the process of optimization includes the second metrics of the network - the delay (d). With constructions of multicast trees, the maximum delay between two endpoints in the network (Δ) is then the applicable appropriate criterion. In [3], [4], it has been proved that finding the tree poses an NP-complete problem for one or more QoS parameters. Due to the complexity of the problem, the presented algorithms use the techniques approaching the solution - heuristics.

The analysis of the effectiveness of the algorithms known to the authors and the design of the new solutions utilize the numerical simulation based on the abstract model of the existing network. These, in turn, need structures (network models) reflecting in the best adequate way the Internet network.

In modelling the topology of the Internet network, it is not necessary, or even advisable, to describe the whole of the network. The dynamics of the changes of the topology depends on random connections and disconnections of the hosts and does not allow for building a model reflecting a given current structure. From the point of view of the effectiveness of the algorithms under scrutiny, the use of such a approach in the simulation process is not economical and it induces a great complexity of the calculations. The research on the traffic in the particular domains (or autonomous system) as well as the inter-domain traffic is sufficient enough because this research takes into consideration the majority of events taking place in the whole of the network.

If the communication network is presented as a graph, the result of the implementation of the routing algorithm will be a spanning tree rooted in the source node and including all destination nodes in the multicast group. Two kinds of trees can be distinguished in the process of optimization: MST - Minimum Steiner Tree, and the tree with the shortest paths between the source node and each of the destination nodes - SPT (Shortest Path Tree). Finding the MST, which is a NP-complete problem, effects in a structure with a minimal total cost. The relevant literature provides a wide range of heuristics solving the above problem in polynomial time [4], [5], [6]. From the point of view of the application in data transmission, the most commonly used is KMB algorithm [4]. The other method minimalizes the cost of each of the paths between the sender and each of the members of the multicast group by forming a tree from the paths having the least costs. In general, it is first either Dijkstra algorithm [7] or Bellman-Ford algorithm [8] that is used, and the branches of the tree that do not have destination nodes are then cut off.

The article discusses the effectiveness of most commonly used constrained heuristic algorithms as well as their comparative usefulness. Chapter 2 presents Steiner tree problem and its heuristics - constrained algorithms: KPP (Kompella, Pasqualle, Polyzos) [9], CSPT (Con-
strained Shortest Path Tree) [5] and least-delay (LD) algorithm. Chapter 3 presents the methodology of generating structures representing the topology under scrutiny and important network parameters. Chapter 4 includes the results of the simulation of the implemented algorithms along with their interpretation.

II. PROBLEM FORMULATION

A. Network Model and Minimum Steiner Tree

Let us assume that a network is represented by a directed, connected graph \( N = (V, E) \), where \( V \) is a set of nodes, and \( E \) is a set of links. The existence of the link \( e = (u, v) \) between the node \( u \) and \( v \) entails the existence of the link \( e' = (v, u) \) for any \( u, v \in V \) (corresponding to two-way links in communication networks). With each link \( e \in E \), two parameters are coupled: cost \( C(e) \) and delay \( D(e) \). The cost of a connection represents the usage of the link resources. \( C(e) \) is then a function of the traffic volume in a given link and the capacity of the buffer needed for the traffic. A delay in the link is, in turn, the sum of the delays introduced by the propagation in a link, queuing and switching in the nodes of the network. The multicast group is a set of nodes that are receivers of the group traffic (identification is carried out according to a unique address), \( G = \{g_1, ... , g_u\} \subseteq V \), where \( u = |G| \leq |V| \). The node \( s \in V \) is the source for the multicast group \( G \). Multicast tree \( T(s, G) \subseteq E \) is a tree rooted in the source node \( s \) that includes all members of the group \( G \) and it is called Steiner tree. The total cost of the tree \( T(s, G) \) can be defined as \( \sum_{e \in T(s, G)} C(e) \). The path \( P(s, g_i) \subseteq T(s, G) \) is a set of links between \( s \) and \( g_i \in G \). The cost of path \( P(s, g_i) \) can be expressed as: \( \sum_{p \in P(s, g_i)} C(p) \), while the delay measured between the beginning and the end of the path as: \( \sum_{p \in P(s, g_i)} D(p) \). Thus the maximum delay in the tree can be determined as: \( \max_{g_i \in G} [\sum_{p \in P(s, g_i)} D(p)] \).

Steiner tree is a good representation for solving the routing multicast problem. This approach becomes particularly important when we have to deal with only one active multicast group and the cost of the whole group has to be minimal. However, due to the computational complexity of this algorithm (NP-complete problem) [3], heuristic algorithms are most preferable. If a set of the nodes of the minimal Steiner tree includes all nodes of a given network, then the problem comes down to finding the minimal spanning tree (this solution can be obtained in polynomial time).

B. Representative heuristic algorithms

Constrained algorithms are heuristics that find minimal trees using an additional parameter - delay (\( \Delta \)) for each path in the network.

KPP heuristics [9] finds a closure graph of constrained shortest paths between all multicast nodes (including the source node). Each edge in this complete graph represents the cost of the shortest path between these nodes in original graph \( N \). Then, KPP finds a constrained spanning tree of the closure graph. Finally, it replaces the edges of the tree by paths from the original graph \( N \) and removes the loops using Prim’s algorithm. Time complexity of this algorithm is \( O(|V|^2) \).

The Constrained Shortest Path Tree algorithm is an example of the minimum cost path tree heuristics [5]. If the delay on path to any destination exceeds the delay bound for this path, then the path will be replaced by a minimum delay path. Thus, if the least-cost tree (LC) exceeds the delay bound (\( \Delta \)), it constructs a minimum delay tree (LD) and combines both trees. This algorithm finds a constrained multicast tree if one does exist.

Least-delay algorithm (LD) was implemented as a Dijkstra’s shortest path algorithm in which \( C(u, v) = D(u, v) \). It guarantees minimum end-to-end delay from the source node to each destination nodes. Time complexity of Dijkstra’s algorithm is \( O(|V|^2) \).

III. NETWORK TOPOLOGY

The main purpose of this article is to analyze the efficiency of the presented algorithms in identical network conditions - size of the network, adopted topology model, values of metrics for each of the edges, etc.

A. Generative methods

In the research a flat random graph constructed according to the Waxman method was used [10]. This method defines the probability of an edge between node \( u \) and \( v \) as:

\[
P(u, v) = a e^{-\frac{d}{L}}
\]

where \( 0 < \alpha, \beta \leq 1 \), \( d \) is Euclidean distance between the node \( u \) and \( v \), and \( L = \sqrt{2} \) is the maximum distance between two freely selected nodes. An increase in the parameter \( \alpha \) affects in the increase in the number of edges in the graph, while a decrease of the parameter \( \beta \) increases the ratio of the long edges against the short ones.

Another method was proposed by Barabasi in [11]. This model suggests two possible causes for the emergence of a power law in the frequency of outdegrees in network topologies: incremental growth and preferential connectivity. The network growth process consists of incremental addition of new nodes. Preferential connectivity refers to the tendency of a new node to connect to existing nodes that are highly connected or popular. When a node \( u \) connects to the network, the probability that it connects to a node \( v \) (already belonging to the network) is yielded by:

\[
P(u, v) = \frac{d_v}{\sum_{k \in V} d_k}
\]

where \( d_v \) is the degree of a node belonging to the network, \( V \) is the set of nodes connected to the network and \( \sum_{k \in V} d_k \) is the sum of the outdegrees of nodes previously connected.

With the construction of the network models based on Waxman and Barabasi-Albert method, BRITE (Boston university Representative Internet Topology gEnerator) [12] was used as a tool for generation of realistic topologies. The application provides a range of network topology models and appropriate generative methods.

Fig. 1 shows typical topologies generated with the application of the Waxman and Barabasi-Albert method.
A network model was adopted in which the nodes were arranged on a square grid with the size of 1000 × 1000 (Waxman parameters: α = 0.15, β = 0.2). Onto the existing network of connections the cost metric \( c(u, v) \) was applied on each link in network (as a matrix of Euclidean distances between the nodes) and the delay metric \( d(u, v) \) resulting from Euclidean distance between the nodes. It means that each link has two parameters: cost and delay.

It was an important element during the simulation process to maintain a steady average node degree of the graph (for each of the generated networks) defined as: \( D_{av} = \frac{2k}{n} \) (where \( n \) is the number of the nodes of the network, \( k \) is the number of edges) which, in practice, meant the necessity of maintaining a steady number of edges. It is generally accepted that for \( D_{av} \geq 2 \), the so-called two-connected network, that is a connected network, can be constructed. Toronha and Tobagi [13] proved that the efficiency of the routing algorithm implemented in a real network was identical to the efficiency of the same algorithm in a random two-connected network. In the implementations, the adopted degree of the graph was within the 3 to 5 bracket.

**B. Network topology parameters**

The efficiency of multicast algorithms depends on implemented network structure. Thus it is important to define basic parameters describing network topology:

- **average node degree**: 
  \[
  D_{av} = \frac{2k}{n} \tag{3}
  \]
  where \( n \) - number of nodes, \( k \) - number of links,

- **diameter** - is the length of the longest shortest-path between any two nodes (a low diameter corresponds to shorter paths),

- **hop diameter** - is the length of the longest shortest-path between any two nodes; shortest paths are computed using hop count metric,

- **length diameter** - is the length of the longest shortest-path between any two nodes; shortest paths are computed using Euclidean distance metric,

- **clustering coefficient** (\( \gamma_v \)) of node \( v \) is the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them [14].

Let \( \Gamma(v) \) be a neighborhood of a vertex \( v \) consisting of the vertices adjacent to \( v \) (not including \( v \) itself). More precisely:

\[
\gamma_v = \frac{|E(\Gamma(v))|}{\binom{k_v}{2}} = \frac{|E(\Gamma(v))|}{k_v(k_v - 1)} \tag{4}
\]

where \( |E(\Gamma(v))| \) is the number of edges in the neighborhood of \( v \) and \( \binom{k_v}{2} \) is the total number of possible edges between neighborhood nodes.

Let \( V^{(1)} \subset V \) denote the set of vertices of degree 1. Therefore [15], [16]:

\[
\hat{\gamma} = \frac{1}{|V| - |V^{(1)}|} \sum_{v \in V} \gamma_v \tag{5}
\]

**IV. Simulation results**

The research work done consists of network topology parameters influence on the efficiency of multicast heuristic algorithms.

The results presented in Fig. 2 show a comparison of KPP, CSPT and LD algorithms in relation to a number of nodes \( n \) in the network and the number of the members of the group \( m \) (destination nodes) with cost as a parameter. Two cases were considered: i. cost of a link is an Euclidean distance between nodes, ii. cost of a link is randomly generated within the range of 500 to 1000. These results show that LD algorithm constructs multicast tree with highest cost. CSPT and LD returns the same trees when cost of the link is an Euclidean distance.

Literature [17], [18], [19] proved that the total cost of the trees generated with the help of KPP (when \( \Delta \to \infty \)) algorithm was, on average, only 5% worse than that obtained by the optimal method (minimum Steiner tree).

The results of the presented algorithms were made dependent also on other parameters of the network such as diameter and the clustering coefficient. The research was carried out for a constant number of the nodes and a constant node degree of the graph. These parameters are not included directly in the BRITE configuration file but are determined for each of the structure which has been generated by the BRITE.

The results presented in Figs 3 and 4 lead to a conclusion that the cost generated by these algorithms solutions...
does not depend on the hop-diameter and clustering coefficient of the network (for the same size of the network). However, the total cost of multicast tree increases with the increase of length-diameter. The present research shows, however, that this algorithm generates solutions with lesser cost with the applications of the structures generated by the Waxman method, while the Barabasi-Albert method generated structures with the constant hop-diameter values.

<table>
<thead>
<tr>
<th>hop diameter</th>
<th>Waxman method</th>
<th>Barabasi method</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>41.2%</td>
<td>1.6%</td>
</tr>
<tr>
<td>5</td>
<td>58.4%</td>
<td>91.1%</td>
</tr>
<tr>
<td>6</td>
<td>0.4%</td>
<td>7.2%</td>
</tr>
<tr>
<td>7</td>
<td>0%</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

**TABLE I**

**PERCENTAGE DISTRIBUTION OF THE DIAMETER OF THE GENERATED GRAPHS FOR DIFFERENT GENERATIVE METHODS** \((n=40, D_{av}=4)\)

![Graph showing total cost of multicast tree versus the number of network nodes and multicast nodes](image)

**Fig. 2.** Total cost of multicast tree versus the number of network nodes \(n\) (a) and the number of multicast nodes (b) for constrained algorithms \((n=100, m_0=10, D_{av}=4, \Delta=10)\).

V. CONCLUSIONS

The article presents and compares representative routing algorithms for multicast connections emphasizing the quality of the network model (accuracy of the illustration of a real internet topology). To this effect, topology generator BRITE, which is considered to be state-of-the-art and very good tool preferred in any research on optimization techniques of networks, was used. Implementation of algorithms constructing multicast trees close to the MST method (like KPP) seems to be advisable with networks of a small number of nodes (for instance LAN networks). On the other hand, algorithms determining the shortest paths (like CSPT) are more appropriate for large-scale networks in which bit flow of multimedia streams can be low.

While analyzing a variety of the diameters of the networks obtained by the application of different generative method it is noticeable that any obtaining of the tree with the lesser cost can be the result of the application of the generative methods and not only the result of the application of a more efficient routing algorithm. However, while comparing the values of the clustering coefficient obtained for different methods of the network topology generation, it can be noticed that the Barabasi-Albert model, despite being more adequate model of the topology of the Internet, generates structures with greater values of clustering coefficient.

Further researches should focus on implementations of fast and simple routing algorithms supporting scalability and work in distributed mode.

**REFERENCES**

Fig. 3. Total cost of multicast tree versus the clustering coefficient obtained for the Waxman (a) and the Barabasi-Albert (b) method for constrained algorithms ($n=40$, $m=10$, $D_{av}=4$)

Fig. 4. Total cost of multicast tree versus the hop-diameter (a) and the length-diameter (b) obtained for the Waxman and the Barabasi-Albert method for constrained algorithms ($n=40$, $m=10$, $D_{av}=4$)