The Influence of Group Members Arrangement on the Multicast Tree Cost

Maciej Piechowiak*, Maciej Stasiak†, Piotr Zwierzykowski†
*Kazimierz Wielki University, Bydgoszcz, Poland, e-mail: mpiech@ukw.edu.pl
†Poznań University of Technology, Poznań, Poland, e-mail: stasiak,pzwierz@et.put.poznan.pl

***Abstract*—The article proposes a novel multicast routing algorithm without constraints and introduces the group members arrangement as a new parameter for analyzing multicast routing algorithms finding multicast trees. The objective of STA (Switched Trees Algorithm) is to minimize the total cost of the multicast tree using a modification of the classical Prim's algorithm (Pruned Prim's Heuristic) and the SPT (Shortest Path Tree) algorithm that constructs a shortest path tree between a source and each multicast node. In the article, the results of the proposed STA algorithm are compared with the representative algorithms without constrains. The results part of the article also contains some selected statistical properties of the multicast routing algorithms finding multicast trees as part of a wider research methodology.**

***Keywords*-multicasting, multicast tree, network topology, routing algorithm**

I. Introduction

The multicast technology is based on the simultaneous transmission of data concurrently to multiple destinations, from the source node to a group of destinations. Over the last few years, multicast algorithms have become more important due to the speciffc natureof data transmitted in transport networks. Current research work on telecommunication networks considers especially real time data transmission. The increase in traffic capacity of present-day networks has offered great advantages in distributed applications such as multimedia data transmission in real time, video-on-demand, teleconferencing etc.

The implementation of multicasting requires solutions to many combinatorial problems accompanying the construction of optimal transmission trees. In the optimization process one can distinguish: 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 $\mathcal{NP}$–complete problem, results in a structure with a minimum total cost [1]. Relevant literature provides a wide range of heuristics solving this problem in polynomial time [2], [6], [7].

From the point of view of the application in data transmission, the most commonly used is the KMB algorithm [2]. Other methods minimize 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. Doar and Leslie [8] show that the total cost of MST constructed by the KMB heuristic is, on average, 5% worse as compared to the exact cost incurred by the MST algorithm [9].

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) that reflect most accurately 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 introduces a great complexity of the calculations. An investigation into traffic in particular domains (or autonomous system) as well as into inter-domain traffic is usually sufficient enough because it 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 $\mathcal{NP}$-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 [2], [3], [4]. From the point of view of the application in data transmission, the most commonly used is the KMB algorithm [2]. The other method minimizes 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. Conventionally, it is first either the Dijkstra algorithm [12] or the Bellman-Ford algorithm [5] that is used, and then the branches of the tree that do not have destination nodes are cut off.

The remarks presented above indicate the direction of the research work carried out by the authors. The main goal of these investigations is to elaborate a methodology for a reliable comparison of existing solutions and a proposition of a new algorithm [28]. This methodology should define a wide range of network topologies as a base for the simulation process of multicast algorithms. A set of important parameters that describes networks should be applied as well. Some statistical properties of the results of multicast algorithms are examined in the article.

The article is divided into seven sections. Section 2 describes the implemented network model. Section 3 presents an overview of the STA algorithm. In Section 4, multicast group members distribution methods are laid down. In Section 5, the simulation methodology is described. Section 6 includes the results of the simulation of the implemented algorithms (STA and others), while Section 7 sums up the presented study.

II. Network model

Let us assume that a network is represented by an undirected, connected graph $G = (V, E)$, where V is a set of nodes, and E is a set of links. The existence of the link $e = (u, v)$ between the nodes 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 $e \in E$, the cost $c(e)$ parameter is coupled. The cost of a connection represents the usage of the link resources. The multicast group is a set of nodes that are receivers of the group traffic (identification is carried out according to a unique i address), $M = \{m_1, ..., m_m\} \subseteq V$, where $m = |M| \leq |V|$. The node $s \in V$ is the source for the multicast group M. Multicast tree $T(s, M) \subseteq E$ is a tree rooted in the source node s that includes all members of the group M and is called a *Steiner tree*. The total cost of the tree $T(s, M)$ can be defined as $\sum_{t \in T(s,M)} C(t)$. The path $P(s, m_i) \subseteq T(s, M)$ is a set of links between s and $m_i \in M$. The cost of path $P(s, m_i)$ can be expressed as: $\sum_{p \in P(s,m_i)} C(p)$.

A 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 minimum. However, due to the computational complexity of this algorithm ($\mathcal{NP}$-complete problem) [11], heuristic algorithms are most preferable. If the set of the nodes of the minimum Steiner tree includes all nodes of a given network, then the problem comes down to finding the minimum spanning tree (this solution can be obtained in polynomial time).

III. Overview of the Algorithms

The simplest way of running the routing algorithm for multicast connections is the implementation of one of the classic algorithms constructing a minimum spanning tree, i.e. the Kruskal algorithm [22], or Prim's algorithm [23]. The designated spanning tree is also constructed for $n \neq M$, which, in practice, effects in aggravation of unnecessary traffic in the network since routers must determine paths for each of the nodes. The cost of tree is disproportionately high in relation to results returned by the exact algorithm (MST).

These inconveniences can be solved by the pruning mechanism that can be introduced to the resulting spanning tree. The PPH technique (*Pruned Prim Heuristic*) [19] is a modification of the classical Prim's algorithm which is a good and efficient solution for solving the Steiner problem when $m \approx n$. PPH builds a minimum spanning tree in the network represented by an undirected graph and removes unwanted arcs – branches that do not contain multicast nodes. Our analysis of algorithms results shows that PPH can construct multicast trees with lower costs as compared to the results of the popular SPT algorithm when the *group density* parameter [13] is greater than 0.5. The mode of operation of the PPH algorithm is presented in Algorithm 1.

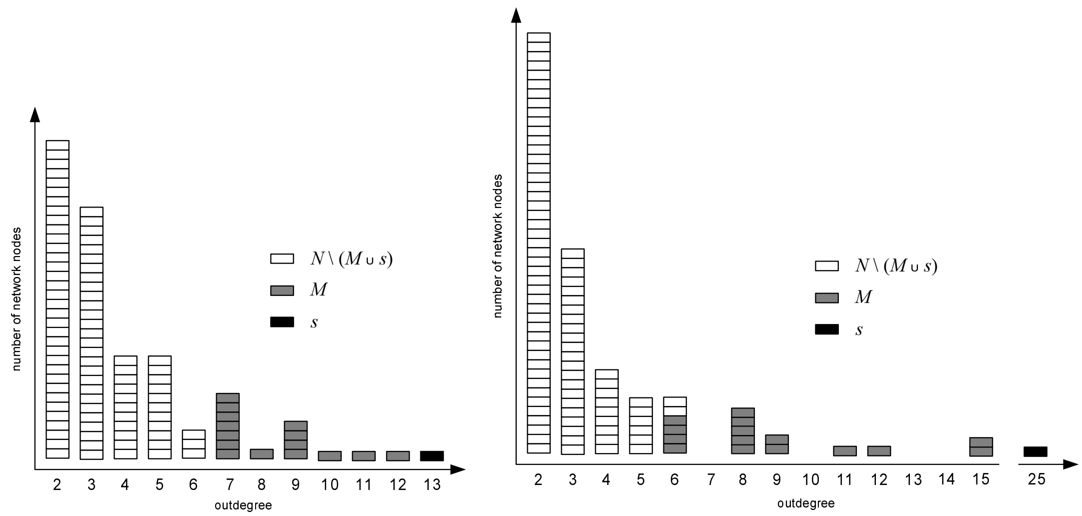
Algorithm 1 Pruned Prim Heuristic

```
1: PPH(C, s, M)
     C – adjacency matrix with costs of links in graph,
     s – source node,
     M – set of multicast nodes m_i ∈ M.
2: T_full ← Prim(C, s)
3: T ← DeleteLeaves(T_full, M)
4: return T
```

Kou, Markovsky and Bermann have proposed the following heuristic algorithm (KMB) determining (constructing) a minimum multicast tree [2]:

- for any cohesive, undirected graph $N = (V, E)$ that includes a set of receiving nodes G, construct a cohesive, undirected graph $N_1 = (V_1, E_1)$ that consists of a sending node s only and of a set of receiving nodes G (the paths between the nodes of graph N_1 are the least cost paths in the original graph N),
- determine a minimum spanning tree T_1 for graph G_1 (if there are more than one solution, choose just one),
- construct a subgraph G_S of graph G by replacing each edge of the tree T_1 with a corresponding path from graph G,
- determine a minimum spanning tree T_S for graph G_S (if there are more then one, choose one),
- construct a Steiner tree T_{KMB} form the tree T_S by removing leaves that do not include receiving nodes.

A good representative for the class of algorithms that construct a multicast tree with the shortest paths is the SPT algorithm (*Shortest Path Tree*). The mode of operation is based on constructing the shortest paths tree only for those nodes that are members of the multicast group (Algorithm 2).



(a) (b)

Figure 1. Histogram of the node outdegree distributions and the illustration of the operation of the algorithm *GroupHighDegree* for an exemplary Waxman (a) and Barabási-Albert network (b) ($n = 100$, $m = 15$, $D_{av} = 4$)

Algorithm 2 Shortest Path Tree

SPT$(C,\ s,\ M)$
C – adjacency matrix with costs of links in graph,
s – source node,
M – set of multicast nodes $m_i \in M$.
for each vertex $m_i \in M$ **do**
  $p_i \leftarrow$ **Dijkstra**(C, s, m_i)
  AddPath(p_i, T)
end for
DeleteLeaves(T, M)
return T

Algorithm 3 Switched Tree Algorithm

STA$(C,\ s,\ M)$
C – adjacency matrix with costs of links in graph,
s – source node,
M – set of multicast nodes $m_i \in M$.
$T_{SPT} \leftarrow$ **SPT**$(C,\ s,\ M)$
$T_{PPH} \leftarrow$ **PPH**$(C,\ s,\ M)$
if $c_{SPT} > c_{PPH}$ **then**
  $T := T_{PPH}$
else
  $T := T_{STA}$
end if
return T

The above observation makes its possible to propose a decision mechanism (STA) which chooses a multicast tree with minimum cost obtained from SPT or PPH that can work concurrently (Algorithm 3). The STA (*Switched Tree Algorithm*) mechanism is based on two well-known optimization algorithms: SPT (*Shortest Path Tree*) and PPH (*Pruned Prim Heuristic*). The combination of these two solutions allows us to achieve a better performance than with the case of each of them working separately. The main concept of the SPT algorithm is to build a shortest paths tree between the source node (s) and each of multicast nodes (m_i) using the Dijkstra algorithm [12]. In the last step of SPT, loops in graphs are removed using the Prim's algorithm and nodes with outdegree 1 which are not multicast members are pruned as well. The STA technique is easy to implement and very fast.

IV. Multicast group members distribution

An essential element of the conducted research study process is a determination of methods for the distribution of the multicast group members in a network [28]. Having in mind various implementations of methods for generating many network topologies, there is also a need for a wider mechanism determining receiving nodes in a network by geographical positioning or one that would be related to the link and node parameters (for example, the outdegree of the network). This will allow us to answer the question whether the way receiving nodes are distributed has any influence on the quality of multicast trees constructed by algorithms. The following methods were used in the study:

- *GroupRandom* method,
  A group of receiving nodes M is constructed by a random choice of m network nodes from among all its nodes (N). The source node s is also randomly chosen from among the number n of nodes in the network. Apparently, this is the only method that has been used so far in any research studies [14], [15], [16], [17], [18].

- *GroupRadius* method,
  Nodes that form a multicast group M and the source node s are chosen from N_r nodes within a circle with radius $r = \frac{d_m}{2}$. This method for creating a group was presented by the authors in [19]. The method reflects the geographical distribution of nodes in a real network.
- *GroupHighDegree* method.
  The algorithm used for the purpose determines the *outdegree* of the network – the number of outcoming links from each node of the network $i \in N$ – and then sorts out the nodes in the diminishing order of this outdegree value. The group $\{M \cup s\}$ is constructed from $m+1$ of the most-preferred nodes (with the highest number of links). Figure 1 shows a histogram for the node outdegree distributions in exemplary (sample) networks generated by the Waxman and the Barabási-Albert methods and explains the operation of the algorithm *GroupHighDegree*.

The method for the receiving nodes M distribution in the network has not been addressed and analyzed earlier in literature.

V. Network topology

A. Generative methods

The Internet is a set of nodes interconnected with links. This simple definition makes it possible to represent this real structure as a graph. In fact, the Internet is a set of domains – a number of grouped nodes (routers) which are under joined administration and share routing information. The Internet consists of thousands of domains and autonomous systems (AS). It is possible to generate those kinds of synthetic structures reflecting real topologies [20].

In the study, a flat random graph constructed according to the Waxman method was used [1]. This method defines the probability of an edge between node u and v as:

$$P(u,v) = \alpha e^{\frac{-d}{\beta L}} \tag{1}$$

where $0 < \alpha, \beta \leq 1$, d is an 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 α effects in the increase in the number of edges in the graph, while a decrease of the parameter β increases the ratio of the long edges against the short ones.

Another method was proposed by Barabasi in [21]. 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 consist 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:

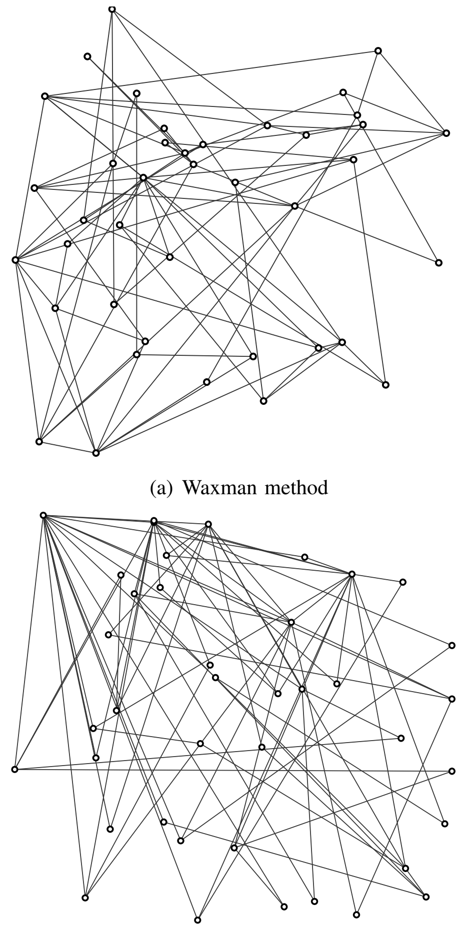
$$P(u,v) = \frac{d_v}{\sum_{k \in V} d_k} \tag{2}$$

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 the nodes previously connected.

With the construction of the network models based on Waxman and Barabasi-Albert method, BRITE (*Boston university Representative Internet Topology gEnerator*) [27] was used as a tool for generation of realistic topologies. The application provides a range of network topology models and appropriate generative methods.

Fig. 2 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: $\alpha = 0.15$, $\beta = 0.2$). Onto the existing network of connections, the cost matrix $C(u,v)$ was applied



(a) Waxman method

(b) Barabasi method

Figure 2. Visualization of network topologies (D_{av}=4, n=40)

(as a adjacency matrix of Euclidean distances between the nodes).

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 the edges) which, in practice, meant the necessity of maintaining a steady number of edges. In the implementations, the adopted degree of the graph was 4.

B. Parameters

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

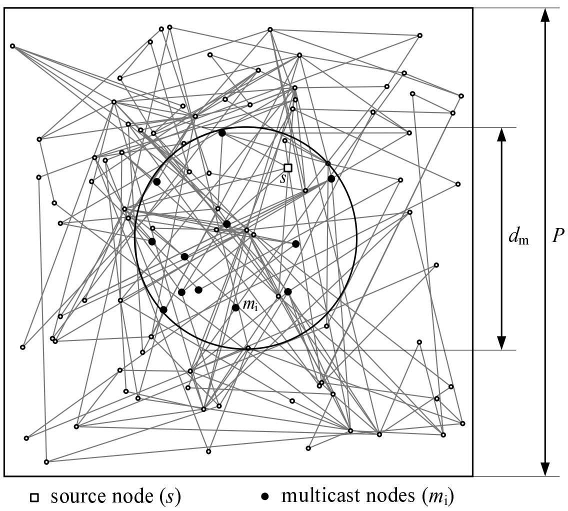


Figure 3. The explanation of the idea of the group spreading factor ε_m ($n = 200,\ D_{av} = 4$)

- *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 the Euclidean distance metric,
- *clustering coefficient* (γ_v) of node v is the proportion of links between the vertices within its neighbourhood divided by the number of links that could possibly exist between them [24].

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

$$\gamma_v = \frac{|E(\Gamma(v))|}{\binom{k_v}{2}} = \frac{|E(\Gamma(v))|}{k_v(k_v - 1)} \quad (3)$$

  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 [25], [26]:

$$\hat{\gamma} = \frac{1}{|V| - |V^{(1)}|} \sum_{v \in V} \gamma_v \quad (4)$$

- *group spreading factor* (ε_m) – describes the arrangement of the multicast group on the plane (Fig. 3). It is defined as the diameter of the area (d_m) containing all multicast nodes divided by the size of plane (P) where all the nodes are situated (1).

$$\varepsilon_m = \frac{d_m}{P}, \quad (5)$$

VI. Numerical results

In the study, a flat random graph constructed according to the Waxman [1] and Barabási [21] methods was used to generate networks topologies to validate the accuracy of our algorithm. With the construction of the network models based on the Waxman and the Barabási-Albert methods, BRITE was used as a tool for generating realistic topologies. The application provides a range of network topology models and appropriate generative methods. The research work was conducted with the application of the networks generated by the above-mentioned methods that were appropriately adopted and unified [29], [30], [31], [32], [33], [35].

The numerical results are divided into three stages. The first stage of the experiment investigates the efficiency of STA and other algorithms. The evaluation of the new constrained algorithm STA and the existing solutions (KMB, DDMC and SPT) bases on the average cost of multicast trees.

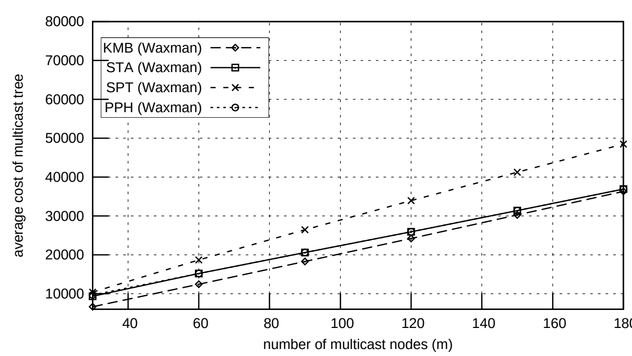


Figure 4. Average cost of multicast tree versus the number of multicast nodes m for Waxman model ($n = 200,\ D_{av} = 4$)

The quality of STA can be observed in Figures 4 and 5. Regarding fixed network parameters, STA bases mainly on PPH heuristic that constructs trees with lower costs as compared to SPT. The convergence of the STA and the KMB results for large groups occurs in the two examined network topology models. The multicast group number growth causes

Table I
VALUE OF THE AVERAGE COST OF MULTICAST TREE CONSTRUCTED BY THE ALGORITHMS WITHOUT CONSTRAINTS FOR DIFFERENT METHODS FOR DISTRIBUTION OF RECEIVING NODES AND METHODS FOR MODELING NETWORK TOPOLOGY ($n = 100$, $m = 20$, $D_{av} = 4$)

model	multicast group	KMB	DDMC	STA	SPT	PPH
Waxman	*GroupRandom*	4732	4739	6442	7252	6651
	GroupRadius$(0, 2)$	5594	5742	6634	6962	7156
	GroupRadius$(0, 4)$	4836	4867	6200	6655	6565
	GroupHighDegree	4797	4819	6294	7230	6427
Barabási	*GroupRandom*	7241	7243	8959	9618	9405
	GroupRadius$(0, 2)$	9058	9463	9902	10273	10409
	GroupRadius$(0, 4)$	7590	7721	8887	9316	9435
	GroupHighDegree	7275	7282	8977	10030	9240

Table II
VALUE OF THE AVERAGE COST OF MULTICAST TREE CONSTRUCTED BY THE ALGORITHMS WITHOUT CONSTRAINTS FOR DIFFERENT METHODS OF DISTRIBUTION OF RECEIVING NODES AND METHODS FOR MODELING NETWORK TOPOLOGY ($n = 500$, $m = 100$, $D_{av} = 4$)

model	multicast group	KMB	DDMC	STA	SPT	PPH
Waxman	*GroupRandom*	19998	20022	25814	30286	25854
	GroupRadius$(0, 2)$	28631	29386	34212	36189	34493
	GroupRadius$(0, 4)$	21909	22102	27554	30523	27692
	GroupHighDegree	20697	20786	25819	30710	25827
Barabási	*GroupRandom*	35230	35243	42489	45979	42777
	GroupRadius$(0, 2)$	43884	45891	47781	50420	48006
	GroupRadius$(0, 4)$	36344	36928	42161	44987	42434
	GroupHighDegree	35468	35505	42470	48175	42533

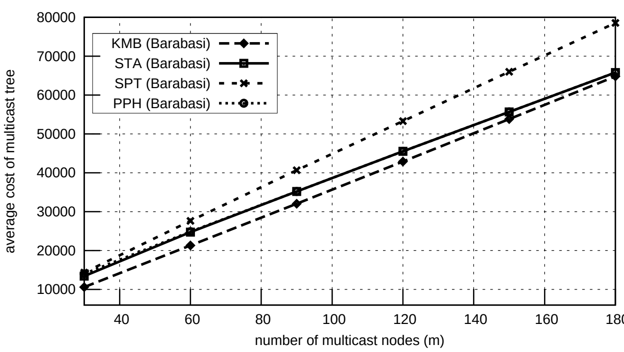


Figure 5. Average cost of multicast tree versus the number of multicast nodes m for Barabási-Albert model ($n = 200$, $D_{av} = 4$)

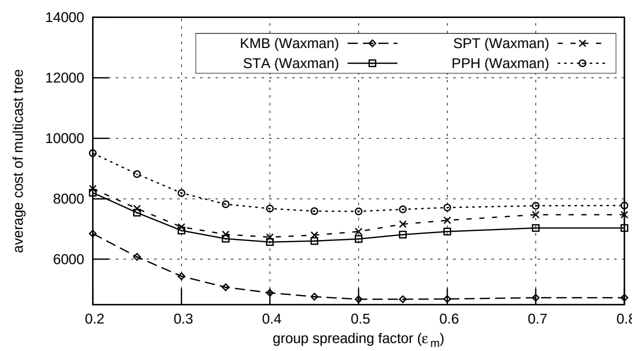


Figure 6. Average cost of multicast tree versus the group spreading factor ϵ_m for Waxman model ($n = 200$, $m = 20$, $D_{av} = 4$)

a growth in the average cost of multicast trees. The costs increase linearly.

Beside conventional parameters of simulation, such as the number of multicast nodes in the network m, we also took into consideration the group spreading factor (ε_m) [19]. A change in this parameter influences the way receiving nodes are distributed in the implemented network. The influence of the group spreading factor ε_m on the average cost of the multicast tree was examined for fixed values of the network parameters,(Figures 6 and 7).

In order to choose and mark nodes in the networks as group members in the area bounded by a circle, the *GroupRadius* method was implemented. The number of network nodes $m = 15$ is required to constitute the multicast group in a narrow area ($\varepsilon_m = 0.2$).

The dependency of group spreading factor is observable for all the examined algorithms. An increase in the area bounded by the circle is followed by a decrease in the average costs of trees constructed by the KMB algorithm. To be more precise, decreasing the factor ε_m to the value 0.2 effects in an increase in the costs of trees 60% on average for the Waxman model (Fig. 6) and 48% for the Barabási-Albert model (Fig. 7). The proposed STA algorithm is less sensitive – the difference is 27%. Above the value $\varepsilon_m = 0.7$, multicast nodes spread randomly throughout the whole network and the costs of trees are similar to those obtained by the

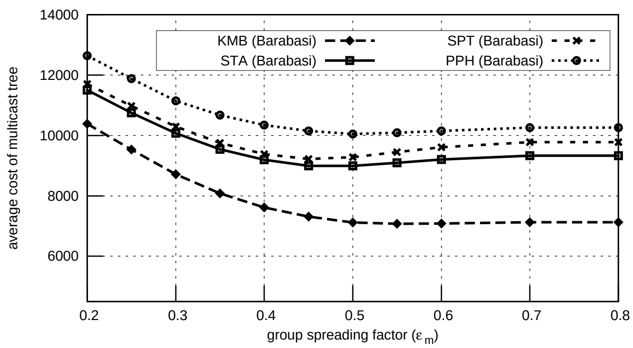


Figure 7. Average cost of multicast tree versus the group spreading factor ϵ_m for Barabási-Albert model ($n = 200$, $m = 20$, $D_{av} = 4$)

GroupRandom method.

Using the methods proposed in Section IV, a research study on the influence of these methods on the average costs of multicast trees was carried out in relation to the following parameters:

- method of generating network topology (Waxman or Barabási-Albert methods),
- method for estimation of the costs of links (as the Euclidean distance and randomly chosen from the interval $10 \ldots 1,000$),
- change in the scale (linear increase of the number of network nodes and receiving nodes) – for ($n = 100$, $m = 20$) and ($n = 500$, $m = 100$).

The indicated studies make sense mainly for algorithms without constraints, where the cost metric is the Euclidean distance, thus being heavily dependant on the distribution of nodes in the plane. The results of the studies carried out for link costs represented as Euclidean metrics are presented in Tables I and II.

Basing on observations related to the influence of the parameter ε_m on the costs of obtained trees, the characteristic values $\varepsilon_m = 0.2$ and $\varepsilon_m = 0.4$ were chosen, for which the above-mentioned costs were respectively the highest and the lowest [19].

The analysis of the average costs of trees indicates the existence of the influence of one particular method applied in the distribution of receiving nodes in the network. Depending on the algorithm under scrutiny, the lowest costs of trees are obtained for the *GroupRandom* method and the *GroupHighDegree* method. The differences in the results for the same algorithm in relation to the applied method fluctuate within the interval 7–21% (Table I, Waxman model), whereas the DDMC algorithm is the most sensitive to the way multicast nodes are distributed (21%), and the least sensitive is the STA algorithm (7%). It should be also noted that these differences are maintained at a similar level for the Barabási-Albert model (10–30%).

Reliability of conducted studies requires an experimental phase that would take into consideration other network parameters. For this purpose, the number of nodes in the network was increased fivefold ($n = 500$) and the number of multicast nodes to ($m = 100$). Similarly to the previous case, the lowest costs of trees were obtained for the *GroupRandom* method and for the *GroupHighDegree* method. The differences in the results for the same algorithm in relation to the applied method fluctuate between 19–46% for the Waxman model (Table II). The most sensitive for the method of multicast nodes distribution is the DDMC algorithm (46%), while the least sensitive is the SPT algorithm (19%). For the Barabási-Albert model, the results are the lowest (10–30%).

The conclusions of thus conducted study, however, are not so conclusive and obvious. There is, indeed, a dependency between the algorithms and the methods for distribution of receiving nodes, but every analyzed algorithm requires, however, an individual approach, with such criteria taken into consideration as: the result of the algorithm (cost of tree or path), the model used and the network parameters, the number of receiving nodes, and so on. For example, the lowest costs of tree are yielded by the DDMC algorithm in networks with 500 nodes and 100 multicast nodes generated by the Waxman model with the application of the *GroupRandom* method as compared with the *GroupRadius*(0.2) method (Table II).

The authors made numerous experiments for different topologies to develop appropriate research methodology. The research investigation was conducted for 1,000 networks.

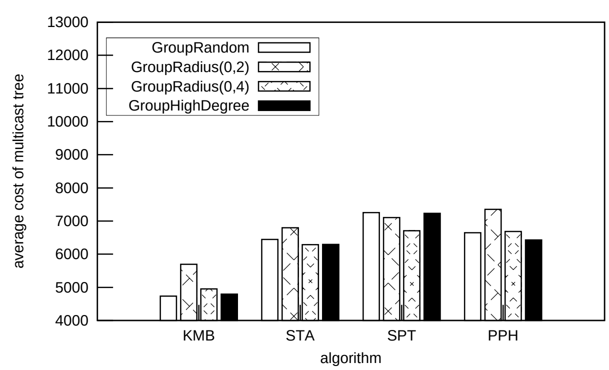


Figure 8. Value of the average cost of multicast tree for different methods for distribution of receiving nodes and Waxman model ($n = 100$, $m = 20$, $D_{av} = 4$)

A similar analysis was made for a cost metric that was a random value taken from within the interval $10 \ldots 1000$ (Figs. 8 and 9). The findings show that the algorithms optimizing cost of trees (KMB and PPH) are more effective with the application of the *GroupRandom* and the *GroupHighDegree* methods. In turn, the SPT algorithm yields the lowest costs of trees for the *GroupRandom* method and the *GroupRadius* method with the parameter $\varepsilon_m = 0.2$. The above observations are correct and accurate both for the

Table III
DESCRIPTIVE STATISTICS OF UNCONSTRAINED ALGORITHMS RESULTS FOR 1,000 NETWORKS ($n = 200$, $m = 20$, $D_{av} = 4$)

model	parameter	KMB	DDMC	STA	SPT	PPH
Waxman	mean value	4731,2	4741,0	7037,8	7471,7	7774,0
	minimum value	3275,0	3275,0	4268,0	4473,0	4268,0
	maximum value	6436,0	6569,0	10385,0	11817,0	12752,0
	standard deviation	527,1	531,9	1012,2	1189,3	1309,0
	variation coefficient	0,111	0,112	0,143	0,159	0,168
	skewness	0,251	0,241	0,227	0,368	0,262
	kurtosis	0,056	0,033	-0,037	-0,038	0,001
Barabási	mean value	7130,1	7134,1	9328,8	9779,9	10260,0
	minimum value	4607,0	4607,0	5673,0	5863,0	5673,0
	maximum value	10616,0	10616,0	16689,0	19314,0	16689,0
	standard deviation	910,1	914,9	1525,2	1690,3	1820,6
	variation coefficient	0,127	0,128	0,163	0,172	0,177
	skewness	0,388	0,398	0,506	0,627	0,374
	kurtosis	0,323	0,331	0,479	0,943	-0,084

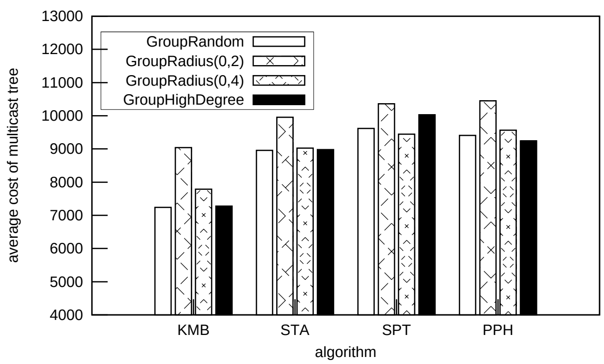


Figure 9. Value of the average cost of multicast tree for different methods for distribution of receiving nodes and Barabási-Albert model ($n = 100$, $m = 20$, $D_{av} = 4$)

Waxman method and for the Barabási-Albert method. The differences in the costs of trees within the same algorithm are 8–30%.

The third stage of the research investigation was conducted for 1,000 networks and the costs of the obtained trees were averaged. The implementation of descriptive statistics as standard deviation, minimum and maximum value, variation coefficient, skewness and kurtosis, allows to confirm the efficacy of this approach (Table III).

The standard deviation is the lowest for the KMB and the DDMC algorithms. It has a slightly higher value for the networks generated with the application of the Barabási-Albert model. The values of the variation coefficient show that the results of KMB, DDMC and STA are least differentiated. The analysis of the skewness parameter indicates an asymmetry (positive skew) of the distribution of the algorithms results. The asymmetry is maximum for the STA algorithm. In the case of the Waxman model, the distributions of algorithms results are close to a normal distribution. The implementation of the Barabási-Albert model leads to a leptokurtic distribution (similar to gamma distribution).

Kurtosis defines a relative measure of the concentration and the flatness of the dispersion of the results for costs of trees. In the case of the Waxman model, the result dispersion has a shape similar to normal dispersion (mesokurtic). This observation applies to all algorithms under scrutiny in the present study. The implementation of the Barabási-Albert model is followed by large values of kurtosis – the dispersion graph becomes more slender than that for normal dispersion (except for the PPH algorithm).

A comparison of the algorithm requires a construction of confidence intervals for average values. The construction of confidence intervals makes it possible to examine whether it is correct to compare the algorithms results only on the basis of the average obtained from 1,000 results. The data presented in Table IV are arranged according to the ascending value of the average cost of trees c_T. The analysis of the ordered data in the table allows us to observe that also the edges of the confidence intervals for both generative methods are arranged in the ascending order. Moreover, except the KMB and DDMC algorithms – for which the values are comparable, the confidence intervals of the remaining algorithms do not overlap. For the given network parameters and algorithm parameters, the following dependence is then correct:

$$c_{KMB} \approx c_{DDMC} < c_{STA} < c_{SPT} < c_{PPH}. \quad (6)$$

VII. CONCLUSIONS

The article proposes a novel multicast routing algorithm without constraints and introduces the group members arrangement as a new parameter for analyzing multicast routing algorithms finding multicast trees. In the article, the results of the proposed STA algorithm are compared with the representative algorithms without constrains.

The article extends the existing methodology for the evaluation of multicast routing algorithms. Analyzing group

Table IV
CONFIDENCE INTERVALS FOR AVERAGE VALUES OF COSTS OF TREES FOR GROUP TRANSMISSION CONSTRUCTED BY STUDIED ALGORITHMS WITHOUT CONSTRAINT FOR 1,000 NETWORKS ($n = 200$, $m = 20$, $D_{av} = 4$)

algorithm	Waxman model			Barabási-Albert model		
	$\overline{c_T} - u_\alpha \frac{S_N}{\sqrt{N}}$	$\overline{c_T}$	$\overline{c_T} + u_\alpha \frac{S_N}{\sqrt{N}}$	$\overline{c_T} - u_\alpha \frac{S_N}{\sqrt{N}}$	$\overline{c_T}$	$\overline{c_T} + u_\alpha \frac{S_N}{\sqrt{N}}$
KMB	4698,53	4731,20	4763,88	7073,68	7130,08	7186,48
DDMC	4708,04	4741,01	4773,98	7077,45	7134,10	7190,75
STA	6975,04	7037,78	7100,51	9234,30	9328,84	9423,37
SPT	7397,94	7471,65	7545,37	9675,09	9779,85	9884,62
PPH	7692,86	7773,99	7855,12	10147,39	10260,24	10373,08

members arrangement constitutes an essential input in the research methodology. This kind of approach has not been considered in relevant literature so far. The article examines unconstrained routing algorithms for multicast connections emphasizing the quality of the network model (the accuracy of the illustration of a real Internet topology) and presents a new proposal.

The research has been conducted by the authors for several years. Initially, the studies focused on the accuracy of multicast routing algorithms in relation to the exact algorithm (MST) and were provided for networks consisting of several nodes [34]. The next stage of research work evaluated the inffuence of the parameters describing graphs (that represents real topologies) on the costs of trees constructed by examined algorithms [29], [30]. The studies are unique because they analyze the algorithms in wide range of network sizes (from several to ten thousand nodes) [31], [33]. Analyses are conducted for networks generated as random graphs with an implementation of Waxman and Barabási-Albert method and with an application of Inet heuristic generator. Separate track of the research analyzed the inffuence of network topology parameters on the cost of multicast tree constructed by selected genetic algorithms [35].

The research results show that obtaining of a tree with the lower 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. Although the Barabási-Albert method returns trees with higher costs, it reflects the real topology of the Internet in the most accurate way. The study also shows that the selection of the generative method should depend on the size of the network (the number of the nodes) and the size of the multicast group.

Finally, authors have proposed the simulation methodology that reflect some network parameters proposed by authors and examine these parameters influence on the costs of multicast trees constructed by multicast routing algorithms.

REFERENCES

[1] B. Waxmann, “Routing of multipoint connections,” *IEEE Journal on Selected Area in Communications*, vol. 6, pp. 1617–1622, 1988.

[2] L. Kou, G. Markowsky, and L. Berman, “A fast algorithm for Steiner trees,” *Acta Informatica*, no. 15, pp. 141–145, 1981.

[3] J. S. Crawford and A.G. Waters, “Heuristics for ATM Multicast Routing,” *Computer Science*, University of Kent at Canterbury, 1998.

[4] Q. Sun and H. Langendoerfer, “Efficient Multicast Routing for Delay-Sensitive Applications,” *Proceedings of the 2-nd Workshop on Protocols for Multimedia Systems (PROMS’95)*, October 1995, pp. 452–458.

[5] R. Bellman, “On a routing problem,” *Quarterly of Applied Mathematics*, vol. 16, no. 1, pp. 87–90, 1958.

[6] Q. Zhu, M. Parsa, and J. J. Garcia-Luna-Aceves, “A source-based algorithm for delay-constrained minimum-cost multicasting,” in *INFOCOM ’95: Proceedings of the Fourteenth Annual Joint Conference of the IEEE Computer and Communication Societies (Vol. 1)-Volume*. IEEE Computer Society, 1995, p. 377.

[7] A. Shaikh and K. G. Shin, “Destination-Driven Routing for Low-Cost Multicast.” *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 3, pp. 373–381, 1997.

[8] M. Doar and I. Leslie, “How bad is naive multicast routing?” *IEEE INFOCOM’ 1993*, pp. 82–89, 1993.

[9] S. Hakimi, “Steiner’s problem in graphs and its implications,” *Networks*, vol. 1, pp. 113–133, 1971.

[10] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee, “How to Model an Internetwork,” *IEEE INFOCOM ’96*, 1996.

[11] R. Karp, “Reducibility among combinatorial problems,” *Complexity of Computer Computations*, pp. 85–104, 1972.

[12] E. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.

[13] F. Cheng and R. Chang, “A tree switching protocol for multicast state reduction,” in *IEEE Symposium on Computers and Communications ISCC 2000*, 2000, pp. 672–677.

[14] L. Wei and D. Estrin, "The trade-offs of multicast trees and algorithms," in *Proceedings of ICCCN'94*. IEEE, 1994, pp. 55–64.

[15] G. Rouskas and I. Baldine, "Multicast Routing with End-to-End Delay and Delay Variation Constraints," *IEEE Journal on Selected Areas in Communications*, no. 15, pp. 346–356, 1997.

[16] G. Liu and K. Ramakrishnan, "A*Prune: An Algorithm for Finding K Shortest Paths Subject to Multiple Constraints," in *INFOCOM 2001 Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies*, 2001, pp. 743–749.

[17] E. Gelenbe, A. Ghanwani, and V. Srinivasan, "Improved neural heuristics for multicast routing," *IEEE Journal on Selected Areas in Communications*, no. 15, pp. 147–155, 1997.

[18] D. Thaler and C. Ravishankar, "Distributed center-location algorithms," *IEEE Journal on Selected Areas in Communications*, no. 15, pp. 291–303, 1997.

[19] M. Piechowiak, P. Zwierzykowski, and T. Bartczak, "An Application of the Switched Tree Mechanism in the Multicast Routing Algorithms," in *1-st Interdisciplinary Technical Conference of Young Scientists InterTech 2008*, 2008, pp. 282–286 (the best paper award).

[20] M. Newman, A. Barabási, and D. Watts, Eds., *The Structure and Dynamics of Networks*, ser. Princeton Studies in Complexity. Princeton University Press, 2007.

[21] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, pp. 509–512, 1999.

[22] R. Kruskal, "Mininum Spanning Tree," *Proceedings of the American Mathematical Society*, pp. 48–50, 1956.

[23] R. Prim, "Shortest connection networks and some generalizations," *Bell Systems Tech J.*, vol. 36, 1957.

[24] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'smallworld' networks," *Nature*, vol. 12, no. 393, pp. 440–442, 1998.

[25] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-Law Relationships of the Internet Topology," *ACM Computer Communication Review, Cambridge, MA*, pp. 111-122, 1999.

[26] T. Bu and D. Towsley, "On distinguishing between Internet power law topology generators," *Proceedings of INFOCOM 2002*, 2002.

[27] A. Medina, A. Lakhina, I. Matta, and J.Byers, "BRITE: An Approach to Universal Topology Generation," *IEEE/ACM MASCOTS*, pp. 346–356, 2001.

[28] M. Piechowiak and M. Stasiak and P. Zwierzykowski, "Analysis of the Influence of Group Members Arrangement on the Multicast Tree Cost," in *Proceedings of The 5-th Advanced International Conference on Telecommunications*, Venice, Italy, May 2009 (the best paper award).

[29] M. Piechowiak and P. Zwierzykowski, "The Influence of Network Topology on the Efficiency of Multicast Heuristic Algorithms," in *Proceedings of The 5-th International Symposium - Communication Systems, Networks and Digital Signal Processing*, M. Logothetis, Ed., July 2006, pp. 115–119.

[30] M. Piechowiak and P. Zwierzykowski, "The Application of Network Generation Methods in the Study of Multicast Routing Algorithms," in *Proceedings of 4th International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks HET-NETs 2006*. Ilkley, UK: Networks UK Publishers, September 2006, pp. P24/1–8.

[31] M. Piechowiak and P. Zwierzykowski, "Performance of Fast Multicast Algorithms in Real Networks," in *Proceedings of EUROCON 2007 The International Conference on: Computer as a tool*, IEEE. Warsaw, Poland: IEEE, September 2007, pp. 956–961.

[32] M. Piechowiak and P. Zwierzykowski, "Heuristic Algorithm for Multicast Connections in Packet Networks," in *Proceedings of EUROCON 2007 The International Conference on: Computer as a tool*, IEEE. Warsaw, Poland: IEEE, September 2007, pp. 948–955.

[33] M. Piechowiak and P. Zwierzykowski, "Efficiency Analysis of Multicast Routing Algorithms in Large Networks," in *Proceedings of The Third International Conference on Networking and Services ICNS 2007*, IEEE. Athens, Greece: IEEE, June 2007, pp. 101–106 (the best paper award).

[34] M. Piechowiak and P. Zwierzykowski, "Analiza efektywności algorytmów heurystycznych dla połączeń rozgałęźnych w sieciach pakietowych,"*Krajowe Sympozjum Telekomunikacji*, pp.192-200, September, 2005 (in polish).

[35] A. Chojnacki, M. Piechowiak, and P. Zwierzykowski, "Genetic Routing Algorithm for Multicast Connections in Packet Networks," *International Journal of Image Processing & Communications*, vol. 13, no. 1-2, pp. 13–20, 2008.