

Algorithmic Support of Optimization of Multicast Data Transmission in Networks with Dynamic Routing

Anatoliy Pavlovich Nyrkov¹, Andrey Sergeevich Belousov¹ & Sergey Sergeevich Sokolov¹

¹Department of Complex provision of information security, Admiral Makarov State University of Maritime and Inland Shipping, Saint Petersburg, Russian Federation

Correspondence: Sergey Sergeevich Sokolov, Department of Complex provision of information security, Admiral Makarov State University of Maritime and Inland Shipping, Saint Petersburg, Russian Federation. Tel: 7-812-748-9720. E-mail: SokolovSS@gumrf.ru

Received: March 5, 2015

Accepted: March 15, 2015

Online Published: March 25, 2015

doi:10.5539/mas.v9n5p162

URL: <http://dx.doi.org/10.5539/mas.v9n5p162>

Abstract

This article describes the optimization process for the multicast tree in the networks with dynamic traffic routing, which is based on the channel state and evaluates each channel. The problem of constructing a multicast connections tree on graphs is formalized. The possibility of improving the existing algorithms by selecting certain "rendezvous points", as well as the algorithm for reducing the "cost" of the multicast routing tree are analysed. Simulation of tree construction was made in accordance with the family of Protocol-Independent Multicast protocols and the proposed algorithm. A critical analysis of the simulation results and comparative evaluation of uniformity of loading data channel at multicast routing are carried out.

Keywords: data networks, dynamic routing, multicast routing, data transfer job, algorithms for constructing trees, multicast tree

1. Introduction

The global transition from an industrial society to an informational one necessitates timely and quality information management, the growth of which has long been in a non-linear function of time. Integrating spheres of management of the Russian Federation into the international information space sets high requirements and standards for the transmission and processing of information. The spread of information technology has defined the problem of centralization and creation of large information systems of regional, departmental and national scale - this is confirmed by the adoption of the federal target program "Information Society (2011-2020)" (Governmental Decree N 313 as of April 15, 2014).

In view of the processing of mostly restricted information in the state systems, which is not related to state secrets (e.g., personal data, financial information, etc.), and taking into account the importance of the formation and management of centralized data accounting, the technology of "next generation", communication networks became widespread, because it provides a user with the access to definable set of services within a single information-computing space. "The current state of telecommunications is characterized by a variety of different services on a single infrastructure" (Luzgachev & Samuilov, 2011).

Recently there is an increase in price of equipment for building networks and organizational technologies for communication and access channels, loads to the key network nodes increase, as well as the number and complexity of the structure in characteristics of the transmitted information. "The reliability of network functioning is compromised by rapid growth in the number of telecommunication services, characterized by an unpredictable amount of traffic transmitted over a network and high performance requirements to the network nodes and servers for different purposes" (Samuilov et al., 2014). In this regard, one of the important and relevant information technology and socio-economic problems is data transmission in modern networks. It is based on the construction of the optimal structure of the data transmission between participants and services, taking into account the need to ensure continuous monitoring and integrity and availability of the transmitted data (Naumov et al., 2007).

The task of information streams routing in data networks in any environments is the major for the network operation (Olifer, 2010). With the increasing number of devices in the network, numerous problems reduce the

effectiveness and, in some cases, block data transmission completely. To ensure the reliability of the network, some of the segments are reserved by additional communications links, which increase the number of possible ways of information communication (Katorin et al., 2013).

One possible solution to optimize the use of network resources is to combine the transmitted information into multicast streams. The advantage of a multicast stream is the ability to deliver one copy of the information packet to multiple recipients without repeating the transmission process. Thus, it is possible to create a group of recipients and transfer the shared data without the repetitive process of sending unicast packets (Deart et al., 2007).

The information can be separated into a multicast stream only in the case if no individual data is required for each recipient. Such requirements are not satisfied by all the categories of the transmitted data. Multicast streams in most cases are represented by a continuous information sequence, at that the necessary bandwidth is many times greater than the network requirements when using unicast connections (Naumov et al., 2007).

Multicast traffic routing problem is significantly different from the one for unicast. In addition to bandwidth requirements, the need to build complex ways of information distribution across the network shall be noted. Such connections can be classified as "point-to-user group" connection (Hwa-Chun Lin and Hsiu-Ming Yang, 2014). Distribution route for a multicast stream is recorded as a set consisting of vertices (source) and branches (recipients/transit nodes), and represents a tree (Casteigts et al, 2012).

Networks with dynamic routing use special, pre-defined criteria for the communication links. This routing method assigns each set of links with a quantitative characteristic, and the selected criterion allows choosing the desired channel. When using unicast connections, then finding the optimal route is sufficient with finding the shortest path to one of the known algorithms (Dasgupta et al., 2014; Takha, 2007). When using data multicast data, building the optimal set for the description reduces to finding the minimum spanning tree from the set of graph's vertices (Nyrkov et al., 2014). This is the formulation of the Steiner's problem on graphs.

The Steiner's problem on graphs is NP-complete. Great practical achievements as for speed, accuracy, etc. have not been reached yet. There are attempts to solve the problem of distributed computing because of the increased processing power of modern technologies, so the possibility of decomposition is also relevant (Du et al., 2009); (Casteigts et al, 2012); (Marcelo, 2013).

The solution to the Steiner's problem is used partially or fully in many branches of science and industry: in the manufacture of integrated circuits, in the design of cable communications systems when finding the optimum travel route, etc. However, we shall note a difference in the baseline at considering data networks with multicast connections. Due to the high data transmission rate, the ability to change the data transmission conditions dynamically, as well as changes in the state of channels, when solving the Steiner's problem in data networks, the speed of finding solutions to reduce the costs of organizing data is the main one, while high accuracy can be neglected.

Multicast routing protocols

Today, there are several protocols for multicast traffic, such as group of protocols: Protocol-Independent Multicast (PIM); Multicast Open Shortest Path First (MOSPF); Distance Vector Multicast Routing Protocol (DVMRP) and other. One of the most commonly used protocols is the one of the PIM group. (Hwa-Chun Lin and Hsiu-Ming Yang, 2014)

PIM protocol group is engaged in multicast traffic routing. Despite the fact that the protocols of the group have the general principles of operation, each of them individually has its own characteristics. The name of "Protocol Independent" suggests that PIM can work with routing table regardless of the way for its completion. The routing table can be built by any dynamic routing protocol or composed of static routes. PIM protocols do not transmit data about routes and do not build own routing table, they build a multicast tree based on the unicast routing table.

The multicast tree is built on routers based on entries in the following format: (Source (S), group (G)). Entries are of the two types: with a known source (S, G) and without a known source, when the tree comes from some general point (*, G). In the router for each entry, one can find information about how the local interface goes up along the route to the source and which local interfaces go down the route to the recipient or to the next router in the path. In constructing the multicast tree, Reverse Path Forwarding (RPF) of the traffic distribution path is performed to ensure the absence of loops and rings. (Nyrkov et al., 2014)

The two main types of PIM protocol shall be considered. The first is PIM Dense Mode (PIM-DM) protocol. A special feature of this type is broadcast on all sources of information throughout the network (domain). Wherein

sending information about the source will be produced even if there are no recipients at the source of the moment. After sending a broadcast, every router in the network has the remained information about the multicast group, and in the case if recipients appear, any router has the information about the route to the source of this group. A router receives a packet about the source only on interfaces that have been RPF-tested, i.e. there is no formation of rings or loops.

The second type is PIM Sparse Mode (PIM-SM) protocol. The main difference from PIM-DM is that the messages of all sources are not broadcasted under the condition of their appearance. In order to make the router start getting multicast traffic, it shall send a command to connect to the group. This occurs when recipients appear on the router or a lower router requests traffic of the group. Since the information on the sources and associated groups is not sent preliminarily to all routers, they cannot determine the correct route to the source.

It is necessary to return to the previously mentioned "common point", because it performs exactly the rendezvous place of all streams of multicast traffic in networks using PIM-SM. It is called rendezvous point (RP) and one originally designated router performs its role. Information about RP shall be known to all other routers, and the router that serves as RP shall keep the way to all sources.

Summarizing, we can say that RP is a summation router for multicast traffic from any source in the network. In this case, connecting routers to RP is the same as connecting to the source of the recipients in PIM-DM protocol.

Using network resources for multicast

The main advantage of using multicast is to reduce the number of copies because the traffic is combined in channels that are on the way of the route to multiple recipients.

In constructing the paths from the source to each receiver, there may be a situation where the traffic is transmitted over two "parallel" paths though having the possibility of combining traffic via one of the paths with adding of several edges. (Nyrkov et al., 2014)

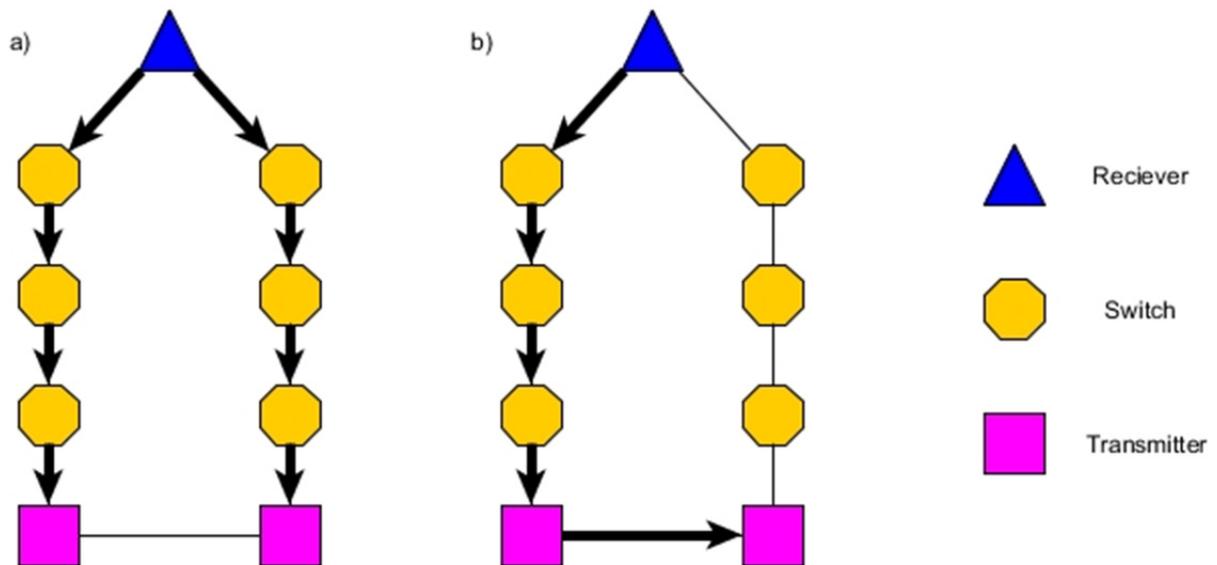


Figure 1. Summation of multicast traffic in the channels for multiple recipients. a) - "parallel" transmission, b) - "joint transmission"

Figure 1 shows a fragment of a network with one source and two recipients, as well as communication links and routers. Figure 1a shows the case where each source builds the path directly to the source and these paths do not intersect. Figure 1b presents an opportunity to make "summation" of multicast traffic and to transmit information for multiple recipients via one path.

In constructing the multicast tree for a large number of recipients, it is possible to reduce data transmission costs significantly the expense of "summation" of information streams. Thus, the tree with many branches will be less efficient than tree with a large trunk and short branches (Mingfang Ni et al., 2011).

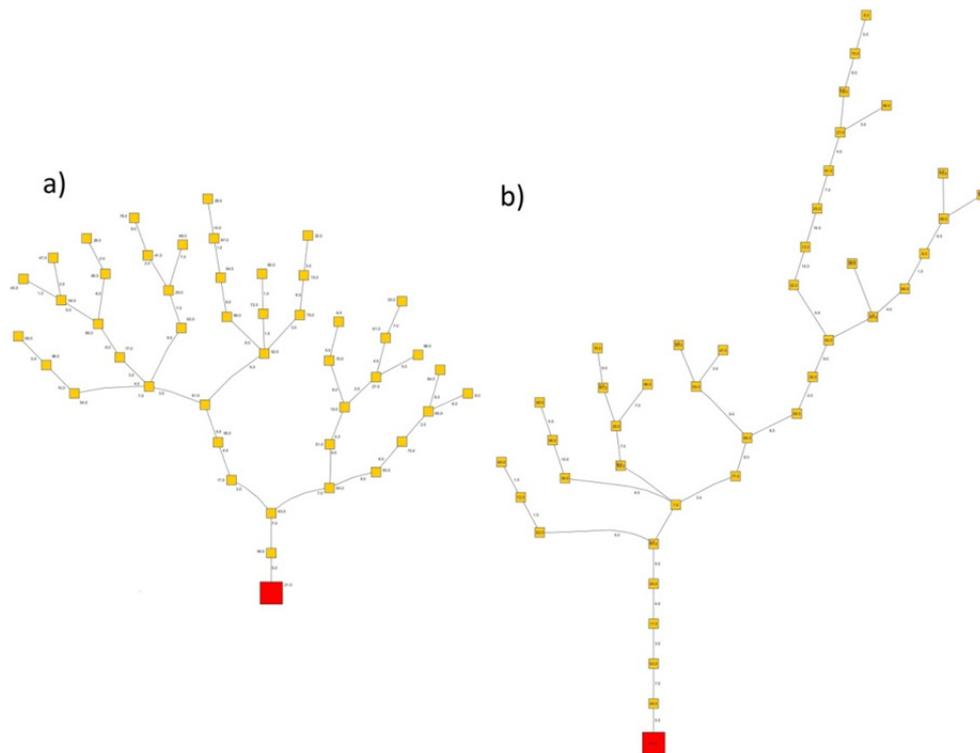


Figure 2. Example of multicast trees

Figure 2 shows an example of two multicast trees for the same set of "source-recipient". The tree was built by different methods. Wherein a) one can see a large number of short branches on the tree from the left side, while on the right, conversely, a small number of longer branches. Consumption network resources in case of b) the tree will be lower since fewer branches transmit traffic for a single recipient only, and the average number of branches of the recipients will be higher than those of the left tree.

Since the PIM protocols family does not directly process the routing of multicast traffic, but only uses the existing unicast traffic routing table, network resources are consumed in suboptimal way in a situation. In such a situation, one of the main advantages of using multicast information is lost: the ability to transmit the same information packets over the same communication link to different destinations without duplication.

When using a dynamic routing protocol based on channel state, for example, Open Shortest Path First (OSPF), each router finds the shortest path to all other routers in the network. This occurs based on the assessment associated with each communication link, and then the Dejkstra algorithm is used to find the shortest path (Cormen et al, 2009). Since the PIM protocols family requires that recipients receive multicast traffic directly from the source or RP, and private routing shall not be performed, though the routing table is used for unicast forwarding instead, then the multicast tree is built only based on channel conditions and location of "source-recipient" pairs.

In this situation, the information transmission over a communication link to multiple participants is the "coincidence" that depends only on the relative position of recipients and the source in the network. Multicast tree will have an arbitrary form. Deliberate "summation" of multicast traffic in communication links is not performed.

2. Method

2.1 Cost Reduction for the Multicast Tree

If the PIM-DM protocol is used, then influencing what will the multicast tree look like, and, respectively, and the costs of network resources for the given parameters: connections state between routers in the network, locations of the source and the recipient, is impossible. All recipients will build their own path to the source. When using the PIM-SM protocol, there is the possibility to specify a specific router as RP.

When selecting a specific RP, part of the tree, namely, from the source to the RP, will be already built, and the

stream for all recipients will be transmitted using this path. At that, the increase in the number of participants does not affect the resources consumption in the network for the information delivery to the RP. Recipients of multicast traffic will connect directly to RP. Thus, when choosing a different RP, resource consumption may be different on both sides and will depend on many parameters.

Assessing the cost for the multicast tree functioning requires deployment of the system for assessing each of the trees. As this is impossible to deploy on a functioning network due to a number of technological peculiarities, it is necessary to create a mathematical model of a data transmission network, in which it will be possible to reconstruct the processes taking place at the organization of the multicast traffic transmission.

The mathematical model shall take into account factors that may make restrictions on some actions, as well as have a number of parameters specific for this type of traffic. This model shall be scalable to allow increasing the number of network elements, depending on the verification assessment and the number of participants.

2.2 The Statement of Multicast Stream Transmission Problem

Let us view the network as an undirected graph $G(V, E)$, where the set of vertices V corresponds to routers on the network, and the set of edges E corresponds to the network segments.

To describe all the processes of data transmission in the network, it is required to determine a number of different additional parameters that will characterize a particular type of connection to transmit information. However, some parameters necessary to describe the processes will be constant, i.e., will not change depending on the established connections. These parameters in the model on the graph include the following variables:

$w(x, y)$ - the total bandwidth of the edge (x, y) of the graph G ;

$\varphi(x, y)$ - the total stream on the edge (x, y) .

Working with multicast connections is significantly different from the "point to point" mode. Interaction of recipients and source of multicast signal is not built on the "point to point" principle, i.e. the recipient does not need to connect namely to the source, it can receive a signal from any router through which or to which the multicast traffic is transmitted. (Basharin et al., 2008)

Instead of the routes in the case of multicast connections, a multicast tree will be built. The root of the tree is the source of the traffic. Further, the branches will transmit traffic to all the "leaves-recipients" of the tree. Branches may contain routers that act as transit hubs.

We denote the set of vertices of the tree as

$$M_V = \{v_S, v_1, \dots, v_{N_V}\},$$

where N_V is the number of nodes through which traffic is transmitted from the source v_S .

The set M_V includes all the vertices that transmit multicast traffic, including both transit and final destination vertices. Let us introduce the set M_{VE} , which includes only the vertex of the source and the destination vertex/vertices:

$$M_{VE} = \{v_S, v_i, \dots, v_{i_{N_D}}\},$$

where N_D is the number of vertices of recipients.

Obviously, $M_{VE} \subseteq M_V$. Since at such a record of the tree, there may be situations where it is impossible to determine unambiguously the final form of the multicast tree, let us introduce additional notation for the edges. Let us consider the set $M_E \subseteq E$. It includes all the edges used for transmission of information from the source. The set consists of pairs of vertices included in the set M_V :

$$M_E = \{(v_S, v_1), \dots, (v_{N_E-1}, v_{N_E})\},$$

where N_E is the number of edges transmitting multicast traffic.

Additional notation to construct a model of multicast traffic is the following:

$f_M(s)$ - the stream from the source vertex s to all of the recipients;

$b_M(s)$ - the requirement to the network bandwidth for the source s .

The data transmission network may simultaneously use various kinds of connection, wherein the amount of each connection is not limited.

Each edge E of the graph G shall be associated with w - the full bandwidth of the edge. The total stream

from all sources over all the edges to all recipients for each edge shall satisfy the following condition:

$$0 \leq \varphi(x, y) \leq w(x, y).$$

The following condition shall be performed to all edges included in the set M_E :

$$w(x, y) - \varphi(x, y) \geq b_M(s),$$

where $(x, y) \in E$.

The initial data for the problem of a particular network includes the graph $G(V, E)$ and bandwidth capacities of edges $w(\square)$ of the graph G (Basharin G. P., 2008). In multicast connections, each information source shall have the following specified:

- set of vertices M_V used to transmit multicast traffic;
- set of vertices of the source and recipients M_{VE} ;
- set of edges M_E used to transmit multicast traffic;
- $b_M(s)$ – the requirement to the network bandwidth for the multicast stream source s .

At that, the following conditions shall be met:

$$\varphi(x, y) = \sum_{j=1; (x,y) \in M_E}^{N_S} b_M(j),$$

where N_S is the number of multicast traffic source.

2.3 Data Transmission Cost: The Main Terminology

Statement of the problem allows describing the processes of data transmission in networks with multicast connections. In order to be able to compare the results of using different methods for constructing a multicast tree, we introduce a quantitative characteristic of "cost" for each edge of the graph $G(V, E)$.

The concept of "cost" is conditional and can take different meanings depending on the goals of the problem. One of the main meanings of the "cost" of data transmission via an edge can be "physical cost of data transmission". For example, the link may be rented from providers or data services suppliers, and this will be charged in cash. Some links can be provided in a variety of conditions, such as communication media, physical environment of installation and laying of communication links, geographical location, and others. With various combinations of these and other factors, it may happen that the organization and maintenance of two identical links will have a completely different cost. In this case, the problem of the proper traffic distribution may be economic in nature, and cost reduction will directly affect organizing the financial cost for network functioning (Malikov et al., 2009).

In other cases, the term "cost" can mean the reliability of data transmission channels. Thus, if the channel is highly reliable and has a full-backup communication channel, one can say that with a high probability, the information transmitted via this path will be delivered to the recipient in full and without errors. On the contrary, considering the communication channel, which is organized with the help of wireless communications in strong external interference, one can say that the probability of loss or errors in such a channel is much larger, and the information can be obtained or received with errors or it can be incomplete. With such a formulation of the problem, the criterion may be the reliability of assessment, and it will be placed in each edge of the graph $G(V, E)$ (Katorin et al., 2013).

In some situations, it is convenient to use the term of "cost" to define the capacity of the edges or the remaining unused part. With this approach, one can solve the problem of routing and load balancing on each edge of the graph. The problem of the information stream distribution on the network segments is the most important one for its efficient functioning. When overloading of communication channels and switching nodes, interruption in the data transmission may occur. They can lead to partial loss of packets due to buffer over stream of the switching equipment or complete blockage of service at an overload of communication channels and equipment.

When the term of the "cost" means the bandwidth, it should be pointed out that in the dynamic routing protocol, which is based on the channel state (based on the Link State Algorithm (LSA), the communication channel assessment is applied to their bandwidth. These algorithms may include the Open shortest path first (OSPF) algorithm and the Intermediate system to intermediate system (IS-IS) algorithm. It is the channel assessment that helps to determine the correct load distribution over the network, thus the network is loaded evenly, and preventing overload of communication links when others are not used (Lun et al., 2006).

In this paper, the "cost" means channel bandwidth: the higher the cost, the lower bandwidth a channel has. This assessment corresponds to channel assessment in dynamic routing protocols based on the channel status. Using

the channel with the lowest cost is preferable compared to the same channel with the greater cost. Thus, the optimization problem will be reduced to the cost of the multicast tree.

This approach allows balancing the network load. When transmitting information through channels with higher bandwidth, inefficient channels will not be overloaded with multicast traffic. Multicast traffic in most cases is represented by the information stream of a large amount, for example, a video stream. (Antonov et al., 2003)

In the case of multicast traffic, it is necessary to calculate the total cost of all links used for transmission of information. At that, the cost will be affected not only by the assessment of each channel, but by the number of channels used in general. Therefore, the optimization problem must take into account these two important parameters.

Multicast data is different in that there is no need to transfer own copy of the information to each recipient, so when calculating the cost of the multicast tree, the cost of using the link will be counted only once, no matter how many recipients will receive the information transmitted via this link. Therefore, one of the main ways to optimize the multicast tree is the process of "summation" of paths to recipients (Chu & Wong, 2005).

Considering the network within the previously described mathematical model, which is represented by an undirected graph with many vertices corresponding to routers, and the set of edges corresponding to the communication links, as well as to the estimated cost associated with each edge, the problem of finding the minimum multicast tree for a given set of "source-recipients" will be reduced to finding the minimum spanning tree of the graph for the given node. Such a problem is called "the Steiner's problem on graph."

This article examines the applied case of the Steiner's problem, namely, a data network with multicast connections. In this case, one of the distinguishing factors is that the information streams in the network may change rapidly in a short period. For example, the multicast recipient may stop receiving traffic information, and the part of the tree will become unwanted. Therefore, the most accurate calculation of the optimal tree cannot be justified as a part of the application. Computation time for the tree can be significantly longer than the data transmission via this tree. In connection with these conditions, it will be sufficient to reduce the cost with minimal time spent in such a way that the computation time of the improved tree would be no more than the total data transmission time via this tree.

2.4 Simulation of the Process of Constructing the Multicast Tree

When using the PIM-DM protocol, the multicast tree will look like a set of routes between all pairs of "source-recipient". In addition, each path will be built as the path with the lowest cost found using the Dejkstra algorithm. Thus, in order to simulate the multicast tree on the graph according to the PIM-DM protocol, it suffices to construct a path with the minimum cost from the source to each destination, and the combination of these paths will be the multicast tree for a given set of "source-recipients". At the same time, finding a minimal path between the two vertices can be done in various ways, but this article will consider the case where the minimum paths are found based on the Dejkstra algorithm.

Let us consider an algorithm that will build a multicast tree through the certain vertex, while the path from the source to this vertex will be built in advance, and all recipients will connect with this vertex. Let us call that vertex as RP, by analogy with the PIM-SM protocol. Let us assume random coinciding of paths or their parts from any of the recipients with the path from the source to RP, but in this situation the multicast stream to these destinations will come from the vertex, which is the entry point into the path from the source to RP, i.e. not required to take the stream namely from RP.

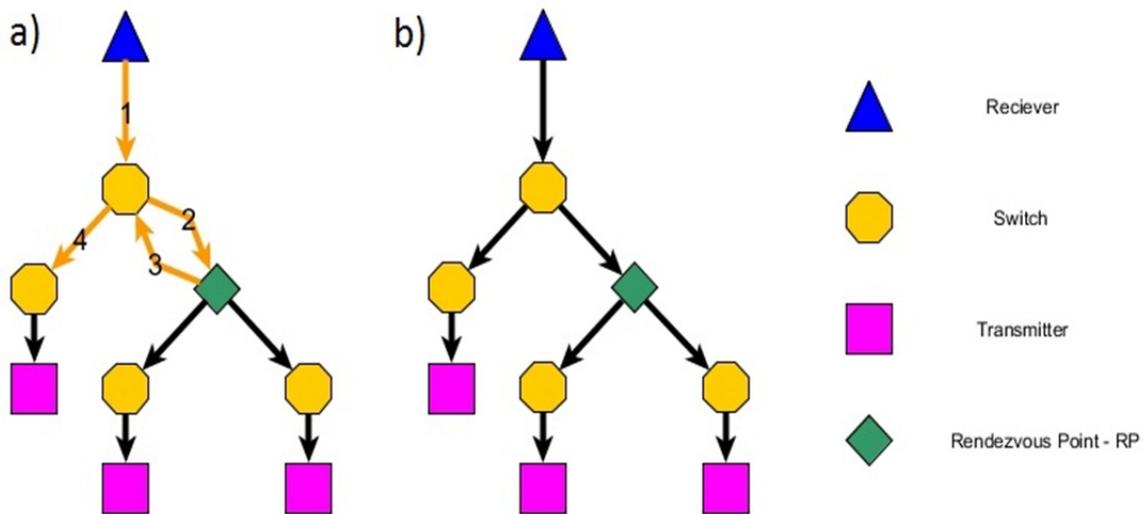


Figure 3. An example of the construction of the tree to RP

Figure 3 shows two options of constructing a multicast tree through RP. Figure 3a shows an option, where the recipient is forced to take traffic from RP, while the path from the recipient to RP coincides with a part of the path from the source to RP, and traffic can be obtained directly from the first router in the path. Figure 3b shows the case where the traffic can be received from the first appeared router on the path from the source to RP. This article will consider the case where the traffic can be obtained from any location at the path from the source to RP.

An algorithm that takes a stream from the certain vertex (RP) corresponds to the PIM-SM protocol. In this algorithm, as well as algorithm of the PIM-DM protocol, the recipients all have to take on a predetermined stream point, i.e. algorithms are the same in meaning except that the point of intersection of all streams can be selected arbitrarily, and therefore the stream control is not performed as well.

Another proposed algorithm is aimed on improving the multicast tree in a decrease in its cost. The algorithm works in two modes with direct connection to the source or connected via a predefined router (by analogy with the previous algorithm let us call it RP). An algorithm for constructing the multicast tree shall be as follows:

Step 1. All recipients make a connection one by one, starting with the recipient, who can connect to the source or RP at a lower cost.

Step 2. If two recipients have the same cost of connecting, the one with fewer edges will connect first.

Step 3. When connecting, the recipient can connect to any point in the network, which already has multicast traffic. Since the task is to reduce the cost of the tree, then the connection will be made to the vertex with the path with minimum cost.

The algorithms, in which there is a possibility to choose RP, will have a complete search among possible vertices to become RP. At the same time, the multicast tree will be the tree having a chosen RP with the least cost of the tree.

For convenience, later in this article the algorithm for constructing the tree will be have the following numbers: 1, 2, 3 and 4 in accordance with the order of their description.

3. Results

Let us consider a random undirected graph of a network consisting of 75 nodes and 330 edges. Each edge in any form will get the valuation from 1 to 10. The cost will mean channel bandwidth: the smaller the cost, the greater the channel bandwidth. Let us randomly choose one vertex, which will be the multicast source. All the recipients will receive stream from it. These conditions are the initial data for the simulation.

For each vertex, let us build a table of shortest paths of minimal cost to any other vertex based on the Dejkstra algorithm (Dasgupta et al., 2014), (Takha, 2007). By analogy with the existing network with dynamic routing, based on the channel state, each router keeps exactly the same table.

Recipients of the multicast stream will appear in random order. Number of recipients will vary from 1 to 74, one vertex (source) less than the total number of recipients.

As recipients are added, a multicast tree will be built in accordance with algorithms 1, 2, 3 and 4.

Let us consider the results of the test.

Figure 4 shows the results of simulation. As it can be seen from the graphs, the cost is reduced by the use of algorithms. In this case, the biggest benefit comes from the use of algorithms 3 and 4. Note that the algorithm 4 gives almost identical results with the algorithm 3.

Figure 5 shows a graph of cost reduction in percentage depending on the number of recipients. As it can be seen from the graph, the benefit from using algorithms based on the simple brute force method, constitutes more than 20% at the number of recipients more than 5. At that, algorithms 3 and 4 provide the benefits of more than 30% at the number of recipients over 7. The maximum benefit is 50.6% at using the algorithm 4 at number of the recipients equal to 13.

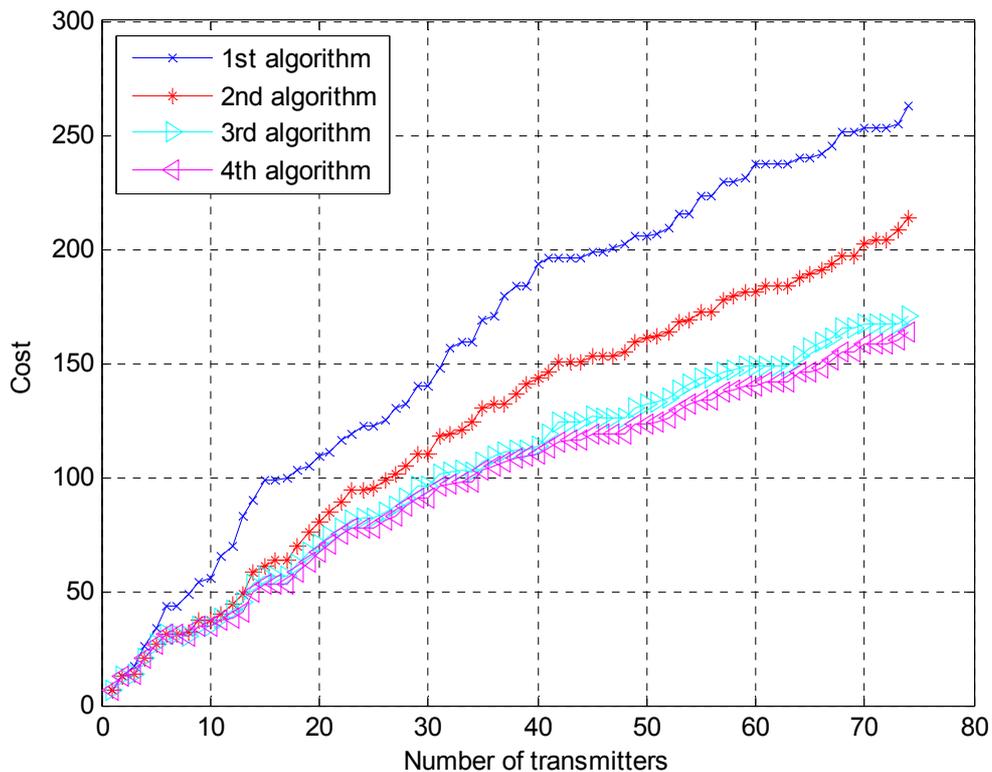


Figure 4. Graph of the cost dependency on the number of recipients

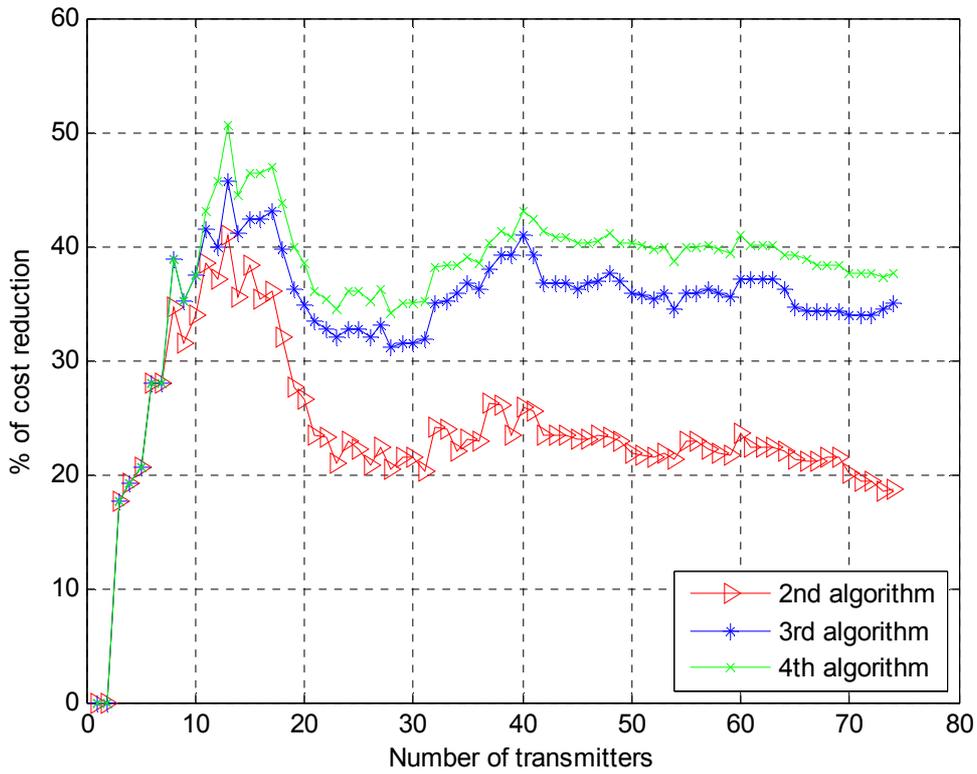


Figure 5. Benefit in percentage of the cost at using algorithms

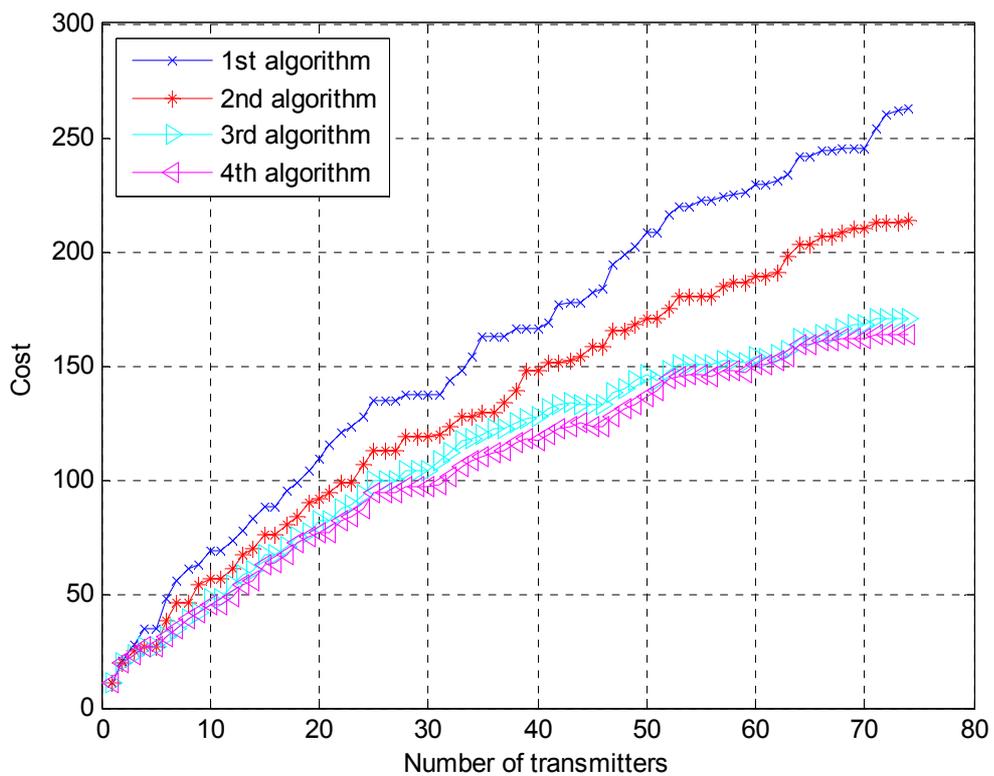


Figure 6. Graph of the cost dependency on the number of recipients

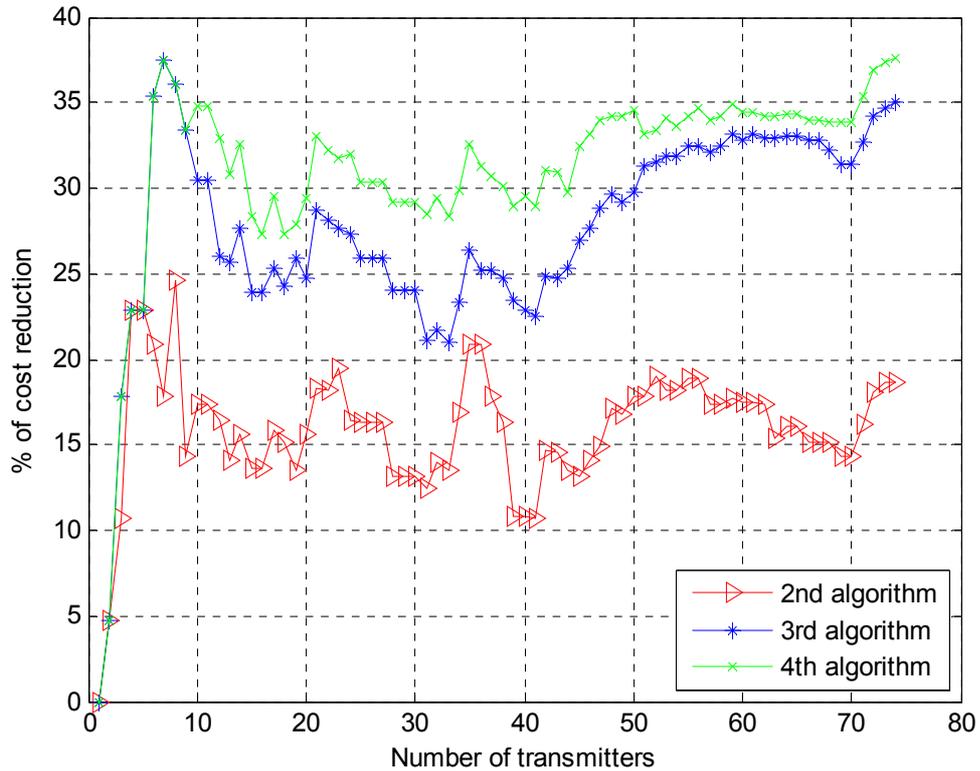


Figure 7. Benefit in percentage of the cost at using algorithms

Figures 6 and 7 illustrate simulation results for the multicast trees with the same input parameters; the only difference is in the sequence of occurrence of the multicast recipients. These graphs differ slightly in shape from the graphs in Figures 4 and 5. The cost is incremented in almost a linear function of the recipients' number. However, one shall notice the difference of graphs 5 and 7. There is a significant reduction in the benefits of algorithms. In addition, the graphs of benefit have sharp jumps to a greater or a smaller side. The reason is that when recipients appear randomly, they can be located in either a favourable place, i.e. "close" to the already running branches, or "far" from the operating units. In this case, the behaviour patterns of graphs at using different algorithms are identical. Of course, the order in which the recipients appear is important and can often dramatically change the results of the experiment. By increasing the number of experiments, it is possible to determine the dependencies on the order of recipients' appearing.

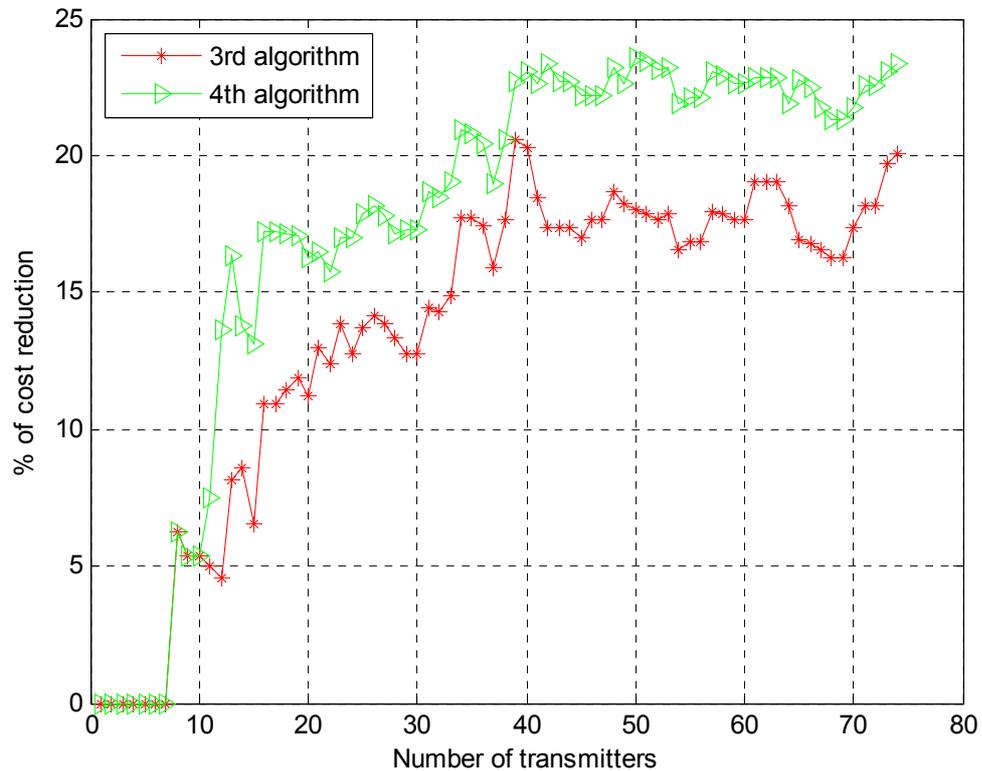


Figure 8. Benefit as a percentage of the cost of using algorithms No. 3 and 4 compared with the algorithm No. 2

The considered algorithms No. 1 and 2 are presented for comparative analysis of the algorithms No. 3 and 4 proposed in this article. As seen from the results, algorithms No. 3 and 4 are superior more than 20% compared with the algorithms No. 1 and more than by 10-15% in comparison with the algorithm No. 2 (Figure 8).

Comparing algorithms No. 3 and 4 between each other, we note that the algorithm No. 4 is somewhat more difficult than the algorithm No. 3, as it performs an exhaustive search of all possible RP's and, respectively, each time a multicast tree is constructed. The difference between the cost of the tree and these algorithms cannot exceed 10-15%. In addition, a great difference is observed only when the number of recipients is less than $N/3$, where N is the number of nodes in the network. By increasing the number of recipients, the benefit from the use of exhaustive RP search is reduced.

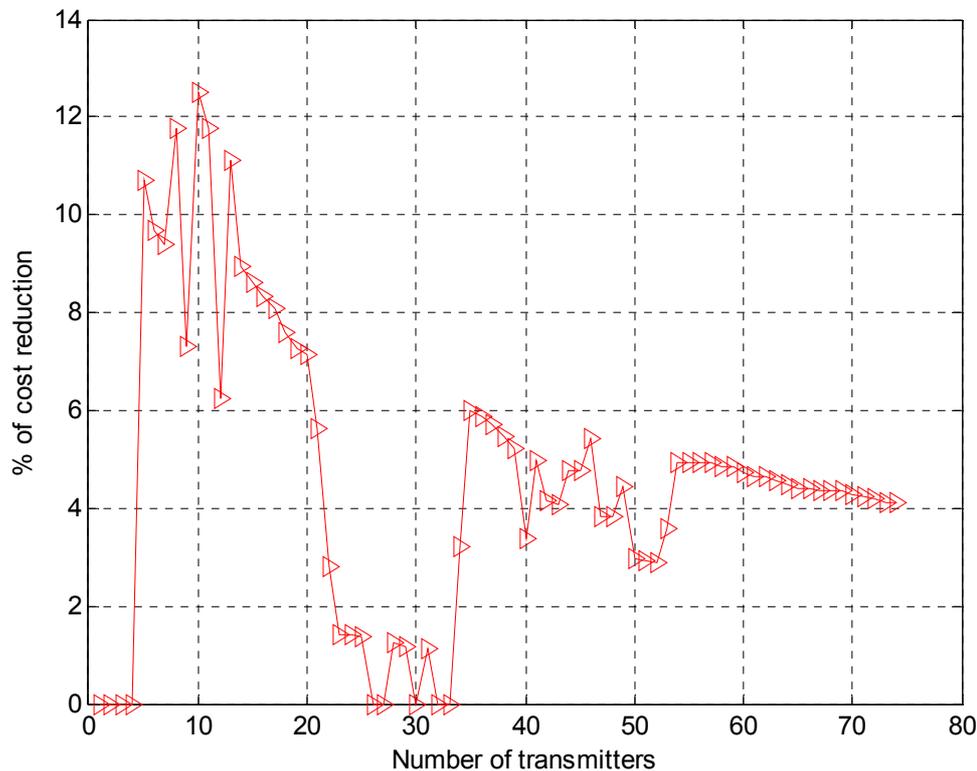


Figure 9. The benefit in percentage of using algorithm No. 4, compared with the algorithm No. 3 depending on the number of recipients

4. Discussion

The obtained results indicate that the tree constructed according to the methods used in PIM protocols, is not optimal. It can be improved by the application of new algorithms for multicast trees constructing.

In the case of tree constructed simply as a set of optimal paths from the source of traffic to each recipient, the main factor is the relative positions of the communication links, i.e., if two recipients are about at the same area of the graph and the optimal path to the source coincides partially or completely, in such a situation, the cost will decline sharply. If the recipients are relatively far from each other, i.e., the optimal paths to the source of each recipient do not intersect, in this situation the cost of connecting these recipients may be reduced. It shall also be noted that in order to find ways to use the minimum cost of the Dijkstra algorithm, which refers to the "greedy" algorithms, because of this we cannot say that in the event of two or more paths to the source of equal cost there will be chosen the one that will reduce the overall cost of the multicast tree.

The tree, which is built according to PIM-SM protocol, i.e. through a pre-selected point, even though improves the previous tree, but also may not be optimal. After brute force finding the vertex through which the cost of the constructed multicast tree is minimal, we can say that we have a case where the source is replaced by a given vertex. It is possible to cut off part of the tree's vertices that were connected to the source without passing through the selected point and re-apply the algorithm recursively as long as the split-off set will have one single vertex, which will be connected by the shortest path found. The complexity of this algorithm is very high.

The algorithms proposed in this paper can build a path to the vertices, through which multicast is already streamed. However, it does not matter if the vertex is a transit one or a recipient, the stream can be picked up from any of the connected vertices. In this case, the path uniting will be non-random; the stream will be connected directly to the vertex, the path to which is the cheapest.

It can be seen that the total cost of the multicast tree when adding new recipients either remains the same or increases (Figures 4 and 6). This is true for each of the algorithms. This situation is easily explained by the fact that when a recipient appears to the vertex, which at the moment is transit and sends traffic to other recipients,

the cost of the tree does not change. Accordingly, the case, when the cost goes up, comes at the time when the multicast traffic shall be transferred to new vertices. The case of reducing the cost of the tree when connecting new subscribers is not possible.

The growth rate of the cost of the tree, as seen from the graph, may be different. This can be explained as follows. As recipients appear in a completely random order, there may be cases where the recipients are located in close proximity to the connected vertices, or vice versa. Thus, it is possible to trace the correlation of the operation cost at different algorithms applied. I.e. improving and maintaining the cost of the multicast tree occurs equally for each algorithm.

From the graphs of interest benefits from applying algorithms to reduce the cost of the multicast tree, it follows that an increase in the number of recipients, the benefit is reduced. This situation is explained by the fact that an increasing number of already connected vertices and the best paths will coincide.

This article only partially addresses the processes occurring in the network with multicast connections. When changing the number of recipients, the tree always was rebuilt, but in a real situations such expenditure of computing resources may be unnecessary and it is necessary to consider cases in which new recipients are advantageous to attach to an existing tree. In addition, the article did not consider the case when it is beneficial to use a more costly algorithm No. 4 in terms of the computing resources in comparison with the fast algorithm No. 3.

In formulating the problem, we used random graph of the network and no dependencies on its type have been investigated. Probably, in certain situations, dependencies will occur on the graph's type.

This work only shows the fact that the application of the PIM family algorithms can have non-optimal nature and improvement of the results is possible in terms of consumption of network resources.

5. Conclusions

This article describes the ways for optimization data multicast tree of a network with dynamic traffic routing.

The analysis showed that the data transmission tree in the network, built using traditional PIM-DM and PIM-SM protocols is not optimal and can be improved in terms of reducing the cost of transmission and integrity of the transmitted information.

To fulfil the said requirement, the article shows that a tree, which includes the optimal route from the source to the destination point, if the transmission paths match (at least partially), then it can reduce the cost provided the availability of the data multicast over a single channel.

Finding the best ways to transfer was carried out using the Deijkstra algorithm.

Multicast tree was built for data multicast using the proposed algorithms that can build paths to the vertices, which are already included in the multicast stream. At that, it does not matter if the vertex is transit one or a recipient, the stream can be taken from any of the connected vertex. Selecting the vertex to connect is made from the standpoint of the least cost of a path to it.

Based on the developed algorithmic support, computational experiments were carried out that showed the cost benefit of the obtained solutions.

The results of the experiment demonstrated that the use of PIM family algorithms, supported by the algorithms proposed in the article, will reduce the cost of data transmission.

Further development of the problem is the study of the types of the obtained trees and their resulting impact on the key characteristics of the data, as well as simulations of real networks containing sections of different physical nature and switching equipment with different characteristics.

References

- Antonov, S. V. (2003). Information technologies of modelling and dynamic management in multi-level channel switching networks. *High-end technologies*, 4, 70-78.
- Basharin, G. P. (2008). Models for the analysis of the quality of service in the next generation communication networks. M.: Publishing House of PFUR, 111.
- Basharin, G. P. (2008). Quality management and operation of probabilistic models of next generation networks. M.: PFUR, 157.
- Casteigts, A., Flocchini, P., Quattrociocchi, W., & Santoro, N. (2012). Time-varying graphs and dynamic networks. *Int. Journal of Parallel, Emergent and Distributed Systems*, 27(5), 387-408.

- Chu, C., & Wong, Y. C. (2005). Fast and accurate rectilinear Steiner minimal tree algorithm for vlsi design. In Proc. International Symposium on Physical Design, pages 28–25, San Francisco, CA.
- Cormen, T., Leiserson, C., Rivest, R., & Stein, C. (2009). *Introduction to Algorithms* (3rd ed.). MIT Press.
- Dasgupta, S. (2014). Algorithms / S. Dasgupta, H. Papadimitriou, W. Wazirani; Translation from English edited by A. Shenya. MCME, 320.
- Deart, V. Yu. (2007). Multi-service communication networks. M.: Insvyazizdat, 166.
- Du, D. Z., Lu, B., Ngo, H. Q., & Pardalos, P. M. (2009). Steiner tree problems. *Encyclopedia of optimization*, 5, 227-290.
- Governmental Decree N 313 as of April 15, 2014 "On approval of the state program of the Russian Federation . Information Society (2011 - 2020). Retrieved February 15, 2015, from <http://base.garant.ru/70644220/#ixzz3S5IqbBCp> (access date)
- Katorin, Yu. F., Razumovskiy, A. V., & Spivak, A. I. (2013). Protection of information by technical means. SPb: ITMO R&D, 468.
- Lin, H. C., & Yang, H. M. (2014). An Approximation Algorithm for Constructing Degree-Dependent Node-Weighted Multicast Trees. *IEEE Transactions on Parallel & Distributed Systems*, 25(8), 1976-1985. <http://dx.doi.org/10.1109/TPDS.2013.108>
- Lun, D. S., Ratnakar, N., Medard, M., Koetter, R., Karger, D. R., Ho, T., Ahmed, E., & Zhao, F. (2006). Minimum-cost multicast over coded packet networks. *IEEE Trans. Info. Theory* 52, 6(June), 2608–2623
- Luzgachiov, M. V., & Samuilov, K. E. (2011). Analytical solution of the resource sharing problem for a single link of a multi-service network with elastic traffic. *T-Comm: Telecommunications and transportation*, 5(7), 102-105
- Malikov, S. N. (2009). Research and evaluation methods of reliability of complex multi-functional systems. *Scientific Journal*, 1(56), 93-101.
- Marcelo, Z. N. (2015). An Interactive Programme for Steiner trees. Retrieved from <http://arxiv.org/abs/1210.7788>
- Mingfang, N., Xinrong, W., Zhanke, Y., & Bin, G. (2011). A Simple and Fast Algorithm for Restricted Shortest Path Problem in Proc. *2011 Fourth International Joint Conference on Computational Sciences and Optimization*, 99-102,
- Naumov, V. A., Samuilov, K. E., & Yarkina, N. V. (2007). Teletraffic theory for multiservice networks. M.: Publishing House of PFUR, 192.
- Nyrkov, A. P. (2014). Ensuring safe operation of multi-service network of transport industry. *Reports by TUSUR*, 2(32), 143-149.
- Nyrkov, A. P. (2014). Multiservice network of transport industry. *Bulletin of computer and information technologies*, 4, 33-38.
- Olifera, V. G., & Olifera, N. A. (2006). Computer networks. *Principles, technologies, protocols* (4th ed.). SPb: Piter, 958.
- Samuilov, K. E., Abaiev, P. O., Gaidamaka, Yu. V., Pechinkin, A. V., Razumchik, R. V., Sopin, E. S., & Shorgin, S. Ya. (2014). Analytical and simulation models to assess indicators of the sip-server under overload. *T-Comm: Telecommunications and transportation*, 8(8), 83-88.
- Takha, H. A. (2007). *Introduction into operations research*. M.: Williams, 912.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).