MULTICRITERION ESTIMATION OF EFFICIENCY OF MOBILE NETWORK CLUSTERING

Yu. Kulakov¹, V. Vorotnikov², O. Boychenko²

¹ National Technical University of Ukraine "Kyiv Polytechnic Institute"
² Zhytomyr Military Institute of S.P. Korolev of the State University of Telecommunications, 22, Myru Ave., Zhytomyr 10004 Ukraine
vvvorotnik@ukr.net

Abstract. The article is sacred to the questions of estimation of efficiency of division of networks on clusters with the use of methods of nonlinear chart of compromises. For the estimation of efficiency of method of clusterization of network an objective function, is used that is scalar chart of private criteria. As private criteria it is expedient to use the next parameters of network: reliability of network, fluence of energy and volume of service information. An example of calculation of efficiency of division of networks on clusters is made at the use of three private criteries for network from 10 knots. It is shown that for the estimation of efficiency of division of networks on clusters it is necessary to use nonlinear chart of compromises. They got results are actual at the decision of intricate problems.

Keywords: Mobile ad hoc networks, network topology, clusters, reliability of network, fluence of energy, volume of service information, efficiency of division, nonlinear chart of compromises

Introduction

The development of information and communication technologies has led to a widespread use of wireless mobile networks with access to the global information network. Such networks are known as Mobile Ad hoc NETWORKs (MANET) do not have a permanent structure, are mobile, and have different quality of technical composition (2, 22, 24).

Communicating over a MANET networks are regulated by the routing protocols which can be divided into three main groups (6, 14, 17 and 20): proactive routing protocols, reactive routing protocols and hybrid protocols.

Hybrid protocols combine methods of proactive and reactive protocols on different hierarchical levels, determining, among route search method, the method of partitioning the network into hierarchical structure or clusters. The disadvantage of hybrid protocols is the comparative complexity of the implementation which is connected with the necessity of clustering (19, 25-27).

The efficiency of network clustering depends on many parameters (8, 14): the number of nodes, the distance between nodes (19), the probability of failure-free work of communication channels between nodes (4, 5), the energy efficiency of information exchange between nodes (11, 16), connectedness (reliability) of network (7, 23) and others. Estimation of the efficiency of the method of network clustering is multicriteria problem (1, 3, 13, 15 and 23). Difficulties of criteria choice, that allow building an adequate objective function, are faced in solving this class of problems (27). The problem of constructing an objective function is that the criteria that characterize the network parameters have their limitations.

Currently, there are many methods for solving multiobjective problems, which are based on different principles of heuristic regularization or extending the definition of formulation the problem [19, 21 and 26).

Another method of solving multiojective problems is connected with the regularization of incorrect task and reducing it to a task that has only solution (25-27). Such methods do not allow solving the problem of multipath routing, which improves the efficiency of a computer network usage without increasing energy costs (8). Furthermore, methods of solving multicriterion problems, which do not take into account the system of restrictions on private quality criteria, are not suitable to solve routing problems (26).

Most of these methods of solving multicriterion problems (25-27) are based on reducing them to a one-criterion and consists in selection the main criterion of quality among the private criteria, followed by optimizing it with appropriate restrictions. This problem belongs to the class of mathematical programming problems, which computational complexity depends exponentially on the dimension of the input data (19).
That is why these methods cannot be used for routing in computer networks with large dimension. Another disadvantage of this method is the need to solve the optimization problem for the majority of private quality criteria at their maximum permissible values. Any change in the mode of a computer network, the load on it can lead to going beyond the limits of partial criteria of quality.

This fact prevents the use of linear convolution of private quality criteria for resolving routing problem in computer networks with large dimension. In addition, linear convolution leads to multieextremum of generalized quality criterion. Thereby, there is the problem of finding the global extremum in multieextremal problem. As a result, the computational complexity of the optimization of the generalized criterion of quality increases even more.

Thus, the definition of multicriteria decision is compromise on its nature and is based on the hypothesis of the existence of a utility function.

The aim of this study is development of a method for estimation the efficiency of the clustering of the mobile network using a nonlinear scheme of compromises.

Method

Let us suppose that there is a network of \( N \) mobile nodes. The network is set by the graph \( G(V, E) \), where \( V \) – set of graph’s nodes \{ \( V_1, V_2, ..., V_N \} \); \( E \) – set of graph’s edges. The network is divided into \( m \) clusters by one of the ways of clustering (3, 15, 18 and 23).

To estimate the efficiency of mobile network clustering it is used the objective function \( E^* \), which is a scalar convolution of private criteria. The following network parameters should be used as special criteria: the reliability of the network, energy costs for information transfer (energy flux density) and the amount of service information.

Reliability of the network

By reliability it is understood the probability that the network operates at a different probabilities of all the of communication channels operability. Statement of the problem of finding the network reliability is as follows: for each link in the network it is given the probability that it is in working condition, it is necessary to calculate the reliability of the route in the network.

Network reliability measures are either certain probabilities of particular random events, or expected values of random variables that depend on the network structure, distance and bandwidth (7, 9, 13).

During calculating the reliability it is necessary to consider not only the topology of the network, but data streams in it. Therefore, the analysis of their complexity includes considerations that are related, but distinct from the mechanism used in dealing with problems of recognition and optimization: classes \( P \), \( NP \) and \( NP \) – complete problems.

In accordance with (13) reliability polynomial can be written as:

\[
\text{Rel}(G) = \sum_{i=0}^{m} F_i p^{m-1} (1 - p)^i,
\]

where \( m \) – the number of components in the network.

The general term in the reliability polynomial \( F_i p^{m-1} (1 - p)^i \) represents the probability that exactly \( m-i \) components of the network work and the network functions in general. This shows that the problem of determining the coefficients \( F_i \) is the combinatorial problem.

One of such problems is the problem of complete counting of states in which initially all the paths \( P_1...P_n \) of the graph \( G \) are known. There is a suggestion that \( E_i \) – the event, when all edges of the route \( P_i \) are workable. Then the reliability equals the probability that one (or more) events \( \{E_i\} \) happened. The events \( \{E_i\} \) are not independent. Probability of two events \( \text{Pr}[E_1 \text{ or } E_2] \) equals \( \text{Pr}[E_1] + \text{Pr}[E_2] - \text{Pr}[E_1 \text{ and } E_2] \). Then the reliability polynomial \( \text{Rel}(G) = \text{Pr}[E_1 \text{ or } E_2 \text{ or ... or } E_j] \) generally defined as:

\[
\text{Rel}(G) = \sum_{i=1}^{m} (-1)^{i+1} \sum_{i=0}^{m} \text{Pr}[E_i],
\]

where \( E_i \) is the event when all paths \( P_i \) are operable. It is a standard extension of inclusion-exclusion [8].

In the presence of minimal paths, the probability of operability of each multipath route is calculated.
Energy flux density

The energy flux density - the amount of energy expended in information transfer and for the ideal conditions of signal propagation is inversely proportional to the square of the distance between nodes (6, 11, 16 and 28-30):

\[ S(G) = \frac{P_T}{4\pi \sum_{i=1}^{m} r_{(1 \rightarrow N_m)}^2} \]

where: \( P_T \) - transmitter power; \( r_{(1 \rightarrow N_m)} \) - the distance between the transmitting and receiving nodes; \( N_m \) - controller of \( m \)-th cluster.

Amount of service information

The amount of service information in general network traffic is one of the important integral indicators characterizing the costs of maintaining the network and determining its capacity. Thus, for the network of IEEE 802.11.n standard with bit rate of 108 bit/s, the theoretical capacity is not more than 51% (22).

To estimate the amount of service information required to maintain the current network topology, let us consider the following topology:

1. Full connected network - each node is connected each (mesh topology).
2. Centralized network - one cluster includes all nodes, node having the minimum total distance \( S_i \) is the cluster controller (star topology).
3. Distributed network - a network is divided into clusters (hybrid topology).

For the full connected network the connectivity degree of nodes is defined as

\[ S = N - 1 \]

It is used \( N \) data channels, on which nodes are connected to each other, for service information transmission to \( N \) nodes. Then the number of packets of service traffic for a full connected network is defined as the sum of the number of packets transmitted on \( N \) channels:

\[ V_s = \sum_{i=1}^{N} N \cdot S = N^2 \cdot S = N^2 \cdot (N - 1) \]

For a centralized network, where information about all nodes is stored in one node, it is necessary to consider two possible cases. In the first case, the location of the central node is changed. In another case, the location of one of the nodes is changed. The degree of connectivity of nodes in a centralized network is equal to one, because they are connected only with a central node. Then the amount of service information in case of changing the location of nodes in centralized network:

\[ V_s = \sum_{i=1}^{N-1} S = (N - 1) \cdot S = N - 1 \]

The degree of connectivity of the central node in the network:

\[ S = N - 1 \]

Considering that the central nod must transmit information about of its location changing to all nodes of the network then the amount of service information in this case is calculated as:

\[ V_s = (N - 1) \cdot S = (N - 1) \cdot (N - 1) = N^2 - 2 \cdot N + 1 \]

Thus, the total amount of service information for a centralized network:

\[ V_s = N^2 - 2 \cdot N + 1 + N - 1 = N^2 - N \]

In practice full connected networks (mesh network) are very rare and even if they are used they require large energy costs. On the other hand, centralized networks have a major disadvantage: the need of providing the powerful computing means for a central node to deal with congestions.

The only way to reduce the relative service traffic in a network is clustering (18). When the network is divided into \( m \) clusters the number of nodes in each cluster is defined as \( n_m \). To find the amount of service information within the cluster it is used the expression (9). Since information about changing the location of any node in the cluster is transferred to cluster controllers, not to all network nodes, you must also take into account the amount of service information that circulates directly between clusters. Exchange of information
between the clusters occurs over separate channels, so the total number of channels will be one less than the number of clusters. Accordingly, the amount of service information in the network which is divided into \( m \) clusters is calculated as:

\[
V_s = \left( \sum_{k=1}^{m} (n_k^2 - n_k) \right) + (m - 1),
\]

where \( n_k \) – the number of nodes in \( k \)-th cluster.

**Objective function of work efficiency of the cluster network**

The objective function that characterizes the work efficiency of the cluster network, can be written as:

\[
E^* = \arg \{ \max (\text{Rel}(G)); \min (V_s(G), S(G)) \}. \tag{11}
\]

Using the scalar convolution of private criteria for nonlinear compromise scheme (26), the objective function can be written as:

\[
E^* = \arg \min \left\{ \frac{1}{1 - V_s(G)} + \frac{1}{1 - S(G)} + \frac{1}{\text{Rel}(G)} \right\}. \tag{12}
\]

**Method**

The efficiency of clustering is considered on an example of the network (Fig. 1), which is given by a matrix of distances \( D[i, j] \), which consists of \( N \) nodes, with limit: the probability of operability of communication channels in the network is the same.

Let us consider a centralized network with a central node 5 (Fig. 1a).

\[ \text{a) centralized control b) distributed control (4 clusters)} \]

**Fig. 1** Graph of mobile network

The amount of service information in this network (9):

\[
V_s = N^2 - N = 100 - 10 = 90.
\]

Centralized network can be considered as a network which consists of a single cluster. Polynomial of reliability of this network (1-2):

\[
\text{Rel}(G) = 2p - p^2.
\]

Density of the energy flux (3) for transferring data from node 1 to the node 10 via the central node 5 in a centralized network is written as:

\[
S(G) = \frac{PT}{4\pi \sum_{i=1}^{r} \left( \frac{1}{N_m} \right)^2} = \frac{PT}{4 \cdot \pi \cdot \left( r_1^2 \cdot r_5^2 + r_5^2 \cdot r_{10}^2 \right)} = 0,5198.
\]

**Fig. 2a** shows the dependence of reliability of the network with a centralized management of probability of communication channels operability.

Based on the above, for the investigated network, divided into 2, 3 and 4 clusters, the values of particular indicators of quality is determined.

For 2 clusters case:

\[
V_s = \left( \sum_{k=1}^{m} (n_k^2 - n_k) \right) + (m - 1) = n_1^2 - n_1 + n_2^2 - n_2 + 2 - 1 = 41;
\]
\[ \text{Rel}(G) = 3p - 3p^2 + p^3, \quad 0 \leq p \leq 1; \]
\[ S(G) = \frac{P_T}{4\pi \sum_{i=1}^{m} (1 \rightarrow N_m)^2} = \frac{P_T}{4 \cdot \pi \cdot (r_{1 \rightarrow 5}^2 + r_{5 \rightarrow 6}^2 + r_{6 \rightarrow 10}^2)} = 0.4328^* P_T. \]

For 3 clusters case:
\[ V_s = (\sum_{k=1}^{m} (n_k^2 - n_k)) + (m - 1) = n_1^2 - n_1 + n_2^2 - n_2 + n_3^2 - n_3 + 3 - 1 = 26; \]
\[ \text{Rel}(G) = 4p - 6p^2 + 4p^3 - p^4, \quad 0 \leq p \leq 1; \]
\[ S(G) = \frac{P_T}{4\pi \sum_{i=1}^{m} (1 \rightarrow N_m)^2} = \frac{P_T}{4 \cdot \pi \cdot (r_{1 \rightarrow 3}^2 + r_{3 \rightarrow 5}^2 + r_{5 \rightarrow 8}^2 + r_{8 \rightarrow 10}^2)} = 0.2847^* P_T. \]

For 4 clusters case (Fig.1b):
\[ V_s = (\sum_{k=1}^{m} (n_k^2 - n_k)) + (m - 1) = n_1^2 - n_1 + n_2^2 - n_2 + n_3^2 - n_3 + n_4^2 - n_4 + 4 - 1 = 13. \]
\[ \text{Rel}(G) = 5p - 10p^2 + 10p^3 - 5p^4 + p^5, \quad 0 \leq p \leq 1; \]
\[ S(G) = \frac{P_T}{4\pi \sum_{i=1}^{m} (1 \rightarrow N_m)^2} = \frac{P_T}{4 \cdot \pi \cdot (r_{1 \rightarrow 2}^2 + r_{2 \rightarrow 5}^2 + r_{5 \rightarrow 6}^2 + r_{6 \rightarrow 8}^2 + r_{8 \rightarrow 10}^2)} = 0.03413^* P_T. \]

Generalized results of calculations of private quality indexes are shown in Fig. 2.

The value of the objective function for different values of the probability of failure-free work of all nodes in the network (12) is shown in Fig. 2d.
Fig. 2 Indexes of quality of network functioning (ending)

Results

Analysis of the results shows that for a network of 10 nodes which is divided into 2 clusters it is possible to reduce the amount of service information almost two times; with the dividing into 3 clusters - almost three times; network of 4 clusters - 6 times. In addition, taking into account the use of reserved paths between the cluster controllers, the average value of reliability of multipath route is increased by 16%.

Discussion

Thus, network clustering allows: reducing the total amount of service information due to transmission of the signaling data within its cluster; improving the reliability of the network due to the possibility of using multipath routing between the cluster controllers; reducing the energy costs of information transmitting due to hierarchical method of creating the route.

Estimation of the efficiency of clustering using nonlinear scheme of compromises combines the quality criteria that characterize the structure of the network, energy costs and the amount of service information. The objective function values (12) decrease while the number of clusters increases, that demonstrates the efficiency of the clustering.

References


http://dx.doi.org/10.1016/j.adhoc.2009.12.005


