Routing Procedure In Network With Nodes Dynamic Usage Rate

Pavel Pustovoitov
Department of Information systems, National Technical University «Kharkiv Polytechnic Institute»,
Kharkiv, Ukraine.
p.pustovoitov@gmail.com

Galina Sokol
Department of Computer Engineering,
Poltava National Technical University,
Poltava, Ukraine.
Sokolgalin_a@mail.ru

Rvachova Natalia
OOO R&D company "RadICS",
Poltava, Ukraine.

***Abstract*— In an article it was proposed a computational procedure that provides a solution to the problem of finding a route in a network that minimizes the total duration of the message transmission between nodes of the network, taking into account the nodes dynamics usage. The described procedure is easily realized, its complexity depends linearly on the complexity (number of nodes) of the network.**

Keywords*— *routing, dynamic nodes usage, computer networks, transmission criteria.

I. Introduction

One of the central problems of organizing traffic in branched computer networks is routing. In a situation where each node of a computer network is both a source and a receiver of data streams, due to the uncontrolled nature of the intensity of these flows, their non-stationarity, prerequisites arise of uncontrolled occurrence of bottlenecks, leading to delays or even loss of information. A natural, real way to mitigate the negative impact of congestion in the nodes of the network is the rational routing of the data transfer process. The corresponding task is to indicate for each source-consumer node pair the sequence of intermediate points in such a way that the determined route is the best from the point of view of the chosen criterion. The standard set of criteria traditionally used in solving the routing problem includes [1-2]:

- minimization of the total data transmission delay during the route;
- minimization of total costs in case of differences in payment for transmission over different lines of the network;
- minimization of the number of intermediate points;
- minimize the probability of failures when passing a route, taking into account differences in the reliability of network lines.

In practice, other criteria or a combination of the above may also be used. Regardless of which criterion will be chosen, the problem is solved on the assumption that all the necessary information regarding the state of each of the nodes of the network at each instant of time is known. In particular, when solving a task on the first of the listed criteria, information about the length of the queue of messages awaiting the start of maintenance for each of the nodes of the network is of particular importance.

II. Statement of the Problem and Analysis of Literature

For the solution of the formulated problem, the well-known algorithms for finding the shortest paths in the network are traditionally used: Bellman [3], Dijkstra [4], Land and Doig [5], Ford-Falkerson [6], branch and bounds [7]. However, all of them have a fundamental disadvantage, the essence of which is the constructive use of the assumption that the usage of network nodes during the whole time of the implementation of data transfer from the source point of the network to the destination does not change. In reality, the delay in the message when passing through the next intermediate point may turn out to be such that the busyness of the next node on the selected route due to the waiting queue of messages waiting to be processed by this time will prove to be unacceptably large [8-10], which will lead to the inadvisability of its use. Thus, ignoring unmanaged changes in the state of network nodes can lead to a significant degradation in the efficiency of the entire process of data transmission in the network.

In this regard, the topic of the article is routing control in a computer network taking into account the dynamics of the usage rate of nodes.

III. Formulation of the Problem

A branched computer network unites a large number of nodes. Therefore, for each pair of correspondent nodes, it is possible to build a very large number of different routes containing a different number of intermediate nodes.Suppose that to transfer information from the *i*-th point of the network to the *j*-th, a route containing s intermediate nodes - $(l_1, l_2, ..., l_s)$ is selected. At the same time, the total delay of data transmission when it is implemented along the specified route will be equal to the sum of the total transit time through the appropriate relay links and delays at intermediate nodes. Therefore, if a value equal to the total delay is chosen as the criterion for the efficiency of the route, then for each node it is necessary to calculate the duration of the delay of the message arriving at the input of this node at a given moment in time, taking into account the dynamics of its usage rate. This data should then be used to select a route that minimizes the criterion. Let's consider the methodology of solving the problem.

IV. MODDELING

Let there be given a network consisting of a set of nodes, some pairs of which are connected by oriented arcs. We introduce the set of $N\{(i,j)\}$ pairs of nodes such that the node i is directly connected with j. In addition, we assign to each pair $(i,j)\in N$ a number c_{ij} ,that specifies the length of the path along the arc emanating from the node i and entering the node j. The value of c_{ij} cannot be measured in units of length, it can be the cost or time of moving from i to j. Now we introduce the set of indicators $\{x_{ij}\}$ as follows:

$$\{x_{ij}\}=\begin{cases}1, \text{if } (\mathrm{i},\mathrm{j})\in N \text{ and arc } (\mathrm{i},\mathrm{j}) \text{ is included to the search path,}\\ \text{otherwise is } 0.\end{cases}$$

Let it be necessary to find the shortest route from the node s to the node r. The mathematical formulation of the corresponding problem has the form: find a boolean set $\{x_{ij}\}$, that minimizes:

$$L(\{x_{ij}\})=\sum_{(i,j)\in N}\sum c_{ij}x_{ij}$$

and satisfies the constraints

$$\sum_j x_{sj}-\sum_i x_{is}=1,$$

$$\sum_j x_{rj}-\sum_i x_{ir}=-1,$$

$$\sum_{(k,j)\in N} x_{kj}-\sum_{(i,k)\in N} x_{ik}=0,\ k\neq s,\ k\neq r.$$

The solution of the problem is achieved by passing to the dual: minimize y_s-y_r under constraints $y_i-y_j\le c_{ij}$, $(i,j)\in N$.

The algorithm of solving the problem consists of a preliminary step and a sequence of similar iterations.

A. Preliminary step.

We set $y_r=0$, $y_k=\infty$, $k\neq r$.

B. Main iteration.

If there is at least one arc $(i,j)\in N$ in the network such that $y_i>c_{ij}+y_j$, the value of y_i changes to $c_{ij}+y_j$. If there is no such arc then stop.

The obvious drawback of the described method is determined by its exhaustive character. The complexity of the solution of the problem grows rapidly with its dimension.

More effective is another method for solving the problem of finding the shortest route, also based on the transition to the dual problem.

In the preliminary step of the method, the network nodes are numbered in this way, for all $(i,j)\in N$ there was a $i>j$. In this case, it is clear that the end node will be number 1, and the initial one – number m, where m - total number of nodes in the network. The algorithm for solving the problem contains the first step and the sequence of the same iterative steps.

A. Step 1.

We set $y_1=0$.

B. Iterative step.

We carry out $y_i=\min\limits_{(i,j)\in N}(c_{ij}+y_j)$, $i=2,3,\ldots,m$.

This algorithm is faster than the previous one, however, it is also unacceptably slow for real-dimensional problems.

Let's consider one more method of searching for the shortest route in the network. The method uses a special operation on the matrices. In this case, two matrices $A(\dim A=m\times p)$ and $B(\dim B=p\times q)$ are associated with a matrix $C=A\otimes B$ $(\dim C=m\times q)$, which elements are defined by the formula

$$c_{ij}=\min(a_{i1}+b_{1j},a_{i2}+b_{2j},\ldots,a_{ip}+b_{pj}), i=1,2,\ldots,m,\ j=1,2,\ldots,q.$$

In the particular case when $m=p=q$, this rule takes the form:

$$c_{ij}=\min_k(a_{ik}+b_{kj}),\ i=1,2,\ldots,m,\ j=1,2,\ldots,m.$$

To implement the method in the problem of finding the shortest path, we introduce the matrix of lengths of one-link paths $R_1=(r_{ij}^{(1)})$, $(i,j)\in N$, between nodes of the network. In this case, if $(i,j)\notin N$, then the corresponding value is $r_{ij}^{(1)}=\infty$. Now calculate the matrix $R_2 = R_1 \otimes R_1$.

It is clear that the element $r_{ij}^{(2)} \in R_2$ determines the length of the shortest path from i to j. The matrix has two important properties. First, in this matrix there can appear two-link paths between some nodes i and j such that $(i,j)\notin N$. Secondly, which is much more important, for some pairs of nodes there can be two-link paths for which $r_{ij}^{(2)} < r_{ij}^{(1)}$, that is, for a pair of nodes (i,j) there can be an intermediate node k such that $r_{ik}^{(1)} + r_{kj}^{(1)} < r_{ij}^{(1)}$.

Continuing further, we obtain the matrix $R_3 = R_2 \otimes R_1$, which determines the best three-link paths. The described procedure is performed until at the next step it turns out that $R_l = R_{l-1}$. Moreover, it is clear that the matrix R_{l-1} determines the shortest paths for any pairs of nodes of the network.

The described procedure is very effective and can be used to solve routing problems of real dimension. However, it does not allow directly to take into account the dynamics of the usage of nodes. Let's consider the method of finding the shortest route in the network, which minimizes the total duration of message transmission, taking into account the dynamics of the usage of intermediate nodes.

We introduce the function $T_k(t)$, which specifies the average delay of the message arriving at the input of the k-the node at the time t. Then, when implementing the route $\{i,l_1,l_2,...,l_s,j\}$, used to transmit the message from point i to j, the average values of the delay values for each of the nodes route.

Let τ_{kl} - be the transmission time of the message over the communication channel between the k-th and l-th nodes (it is equal to infinity if the k-th и l-th nodes are not directly connected by the communication channel).We obtain a set of recurrence relations for calculation of delays in nodes when a message is traversed along the selected route. We have:

$T_{l_1} = T_{l_1}(\tau_{il_1})$ - delay in the l_1 node of the message arriving at the node's input l_1 at the time $t=\tau_{il_1}$;

$T_{l_1l_2} = T_{l_2}(\tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2})$ - delay in the l_2 node of the message arriving at the node's input l_2 at the time $\tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2}$;

$T_{l_1l_2l_3} = T_{l_3}(\tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2} + T_{l_2}(\tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2}) + \tau_{l_2l_3})$ - -delay in the message l_3, node of the message arriving at the node's input l_3 at the time

$t = \tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2} + T_{l_2}(\tau_{il_1} + T_{l_1}(\tau_{il_1}) + \tau_{l_1l_2}) + \tau_{l_2l_3}$.

It is clear that this sequence of formulas can be continued.

Then the total delay $T_{l_1l_2,...,l_s}^{\Sigma}$, arising from the use of the route $l_1,l_2,...,l_s$, will be:

$$T_{l_1l_2,...,l_s}^{\Sigma} = \tau_{il_1} + \tau_{l_sj} + \sum_{k=1}^{s-1}\tau_{l_kl_{k+1}} + T_{l_1} + T_{l_1l_2} + T_{l_1l_2l_3} + \ldots + T_{l_1l_2\ldots l_sj}. \quad (1)$$

Thus, the problem is reduced to the following: for a pair of nodes (i,j) find the route $\{i,l_1,l_2,...,l_s,j\}$, connecting them, minimizing (1) and satisfying the constraints:

$$\delta(i,l_1) = \delta(l_1,l_2) = \ldots = \delta(l_s,j) = 1, \quad (2)$$

where

$$\delta(k,l) = \begin{cases} 1, \text{ if } k\text{-th and } l\text{-th nodes} \\ \text{directly connected by a communication channel,} \\ \text{otherwise is } 0. \end{cases}$$

It is clear that to obtain a solution it is necessary to have an analytic description of the function $T_i(t)$. We obtain it by adopting the following assumptions. First, we assume that the input of the k-th node of the network receives a superposition of fluxes from different sources, forming in aggregate a non-stationary Poisson flow with intensity $\lambda_k(t)$. Secondly, suppose that the random service time of messages in the k-th node of the system is distributed exponentially with an intensity of $\mu_k(t)$. Let $n_k(0)$ - be the length of the queue in the k-th node at the time $t=0$ begins the message transmission from the node i. Finally, we assume that the average time for a message to be transmitted from one node to another via a communication channel in a real network is less than the time for a significant change in the network state.

Then the average length of the queue in the k-thnode at time t will be:

$$n_k(t) = n_k(0) + \int_0^t \lambda_k(t)dt - t\mu_k. \quad (3)$$

Moreover, for the description of $\lambda_k(t)$, for example, in the form of a regression polynomial, one can use a sequence of values of the number of messages arriving at the input of this node at some interval preceding the moment of sending a message from the node i. Then, taking into account (3), the average delay of the requirement received at the input of the k-th node at time t is equal to:

$$T_k(t) = \frac{1 + n_k(t)}{\mu_k}. \quad (4)$$

Let us turn to the technology of finding the optimal route.

We choose the time interval $[0, T_{\max}]$, the length of which is certainly longer than the transmission time of the message in the network and split it into q sub-intervals $[0,\Delta),[\Delta,2\Delta),...,\ [(s-1)\Delta, s\Delta),\ ...,[(q-1)\Delta, T_{\max}]$, of length $\Delta = \frac{T_{\max}}{q}$. Now, for each subinterval $s = 1,2,...,q$ we calculate the delay matrix for implementing one-path paths between each pair of network nodes:

$$T^{(1)}(s\Delta) = \left(T_{kl}^{(1)}(s\Delta)\right);$$

$$T_{kl}^{(1)}(s\Delta) = \left(\tau_{kl} + T_l(\tau_{kl} + s\Delta)\right), \qquad s = 1,2,...,q \ . \qquad (5)$$

Now, taking into account (5), we obtain a set of matrices that specify the minimum delay in realizing the shortest (in sense of delay) two-step paths:

$$T^{(2)}(s\Delta) = \left(T_{kl}^{(2)}(s\Delta)\right);$$

$$T_{kl}^{(2)} = \min_r \left\{\left(\tau_{kr} + T_r(\tau_{kr} + s\Delta)\right) + \left(\tau_{rl} + T_l\left(\tau_{kl} + T_r(\tau_{kr} + s\Delta) + \tau_{kr}\right)\right)\right\} =$$

$$= \min_r \left\{\left(\tau_{kr} + T_r(\tau_{kr} + s\Delta)\right) + \left(\tau_{rl} + T_l\left(\tau_{rl} + T_{kr}^{(1)}(s\Delta)\right)\right)\right\} =$$

$$= \min_r \left\{T_{kr}^{(1)}(s\Delta) + T_{rl}^{(1)}\left(T_{kr}^{(1)}(s\Delta)\right)\right\}, s = 1,2,...,q \ . \qquad (6)$$

This recurrence relation has a very clear meaning. In order to obtain the optimal two-way path between any pair of nodes (k,l) it is necessary to select such an allowable intermediate node r, that the sum of the delay in the first step and the delay in the second step, depending on the choice in the first step, be the smallest. It is clear that such an optimal route will be found if for pair (k,l) there is at least one node r such that $\delta_{kr} = \delta_{rl} = 1$.

Further, the obtained set of $T^{(2)}(s\Delta)$ matrices is used to calculate the matrices of optimal three-step paths:

$$T^{(3)}(s\Delta) = \left(T_{kl}^{(3)}(s\Delta)\right);$$

$$T_{kl}^{(3)} = \min_r \left\{\left(\tau_{kr} + T_r(\tau_{kr} + s\Delta)\right) + T_{rl}^{(2)}\left(T_{rl}^{(1)}(s\Delta)\right)\right\} =$$

$$= \min_r \left\{T_{kr}^{(1)}(s\Delta) + T_{rl}^{(2)}(s\Delta)\right\}, \qquad s = 1,2,...,q \ . \qquad (7)$$

And this ratio is easily interpreted: the optimal three-step path is determined by the best choice of the first intermediate node, taking into account the optimal two-step continuation. Acting similarly, for an arbitrary p-th step we have $T^{(p)}(s\Delta) = \left(T_{kl}^{(p)}(s\Delta)\right);$

$$T_{kl}^{(p)} = \min_r \left\{T_{kr}^{(1)}(s\Delta) + T_{rl}^{(p-1)}(s\Delta)\right\}, \ s = 1,2,...,q \ . \qquad (8)$$

The described procedure is continued until, at some next iteration, the resulting matrices for $s = 1,2,...,q$ optimal paths do not coincide with the corresponding matrices calculated in the previous step. Further calculations can be stopped.

It should be noted that the computational complexity of the procedure does not depend on the step number: at each iteration, the first step is selected taking into account the already prepared optimal continuation. The accuracy of the resulting solution is determined by the length of the subinterval Δ, which determines the discreteness with which the continuation matrix is selected after calculating the values $T_{kr}^{(1)}(s\Delta)$, $s = 1,2,...,q$.

V. CONCLUSIONS

Thus, a computational procedure is proposed that provides a solution to the problem of finding a route in a network that minimizes the total duration of the message transmission between nodes of the network taking into account the dynamics of the usage of intermediate nodes. The described procedure is easily realized, its complexity depends linearly on the complexity (number of nodes) of the network.

REFERENCES

[1] E. Tanenbaum. *Komuternie seti* [*Comuter neworks*]. St Peterburg, Russia: Piter, 2010. (In Russian).

[2] P. Pustovoitov. “Flow management in routers by criteria of average queue length”, *Proceedings of the XIth International Conference of Modern problems of radio engineering, telecommunications and computer science*, TCSET’2012, Lviv – Slavske, 2012, P.363.

[3] S. Vegeshna. *Kachestvo obslujivania v setiah IP. [IP Quality of Service]*. Moscow, Russia:Williams, 2003. (In Russian).

[4] H.A. Taha. *Vvedenie v issledovanie operacij* [*Operations Research: An Introduction*]. Moscow, Russia: Williams, 2005. (In Russian).

[5] A. H. Land and A. G. Doig. *An automatic method of solving discrete programming problems*, pp. 497-520.

[6] B. Stallings. *Sovremennie komputernie seti* [*Modern computer networks*]. St Peterburg, Russia: Piter, 2003. (In Russian).

[7] T. Korman *Algorithmi: postroenie I analiz [Introduction to algorithms]*. Moscow, Russia: Williams, 2006. (In Russian).

[8] N.V. Rvachova and G.V. Sokol. *“Mathematical Model Of The Process Of Communicating Real Time Traffic In Telecommunication Network Channel”* IEEE Second International ScientificPractical Conference «Problems of Infocommunications. Science and Technology» (PIC S&T 2016). Ukraine, Kharkiv, 2016, pp. 196-198.

[9] P. Pustovoitov, P. and G. Sokol *The model of dual channel network node with shared memory*, Proceedings of 2nd International Conference on Advanced Information and Communication Technologies, AICT 2017 - pp. 266-269.

[10] Pustovoitov, P.Ye and Raskin, L.G. *Technique of analysis of semi-Markovian system of high dimension*, Journal of Automation and Information Sciences, 2013.