ADAPTIVE ROUTING TECHNIQUE AND SIMULATION OF COMMUNICATION NETWORKS

Alexander V. Butrimenko
Institute of Information Transmission Problems, USSR Academy of Sciences
Moscow, USSR

ABSTRACT

The report describes decentralized distribution of the information flows in a switching connecting network, when the controlling devices have incomplete information on the network state. Some results of computer simulation of the network allow to assess the effectiveness of this method in different conditions.

Modern systems of wire communication, covering vast territories and having connecting units located far from one another, require effective control.

Since statistical characteristics of the network are, as a rule, not known in advance and may change in the process of operation, the control system must be able to redistribute information flows in accordance with the altered traffic capacity of various parts of the network and traffic, i.e. to be adaptive.

Besides, the control system must possess great survivability, for disruptions of various parts of the network not to affect the control system as a whole.

All these features stipulate for the application of the decentralized or distributed method which transfers the control function to each connecting unit.

The controlling devices in various units do not repeat one another and none of them gets complete information on the network state. Such a distribution of control functions among various parts of the data system, none of which knows the task as a whole and uses only some data on the adjacent elements, resembles an automata game in which the automaton knows neither the payment functions nor the action of its partners. M.L. Tsetlin showed that a group of automata operating in similar conditions exhibited advisable behaviour. The possibility of application of automata games for the purpose of organizing adaptive and survivable systems of decentralized control over the networks was proposed by V.G. Lazarev.

Let us describe the decentralized control method using the example of a switching long-distance telephone network. The connecting units are assumed to be receivers and sources of information.
If unit $V_i$ is to be connected with unit $V_j$, the control system is to find a path between these two, consisting of working trunk lines, each containing at least one empty channel. Usually, the number of intermediate transit connecting units determines the communication quality and is one of the parameters of the optimal control. The number of transducers determines another criterion of the network operation - the probability of successful termination of a conversation, i.e. The conversation is not interrupted in case of a trunk line disruption. The main parameter of the control quality is the probability of a refusal for establishing a call, which is due to one of the three reasons:

a) a trunk line disruption disintegrates the network and the units to be connected belong to different subnetworks;

b) there is no path due to an "uneconomical" occupation of the channels by previous demands, in other words the call could be established if earlier demands were distributed in a different way;

c) there is a path but it cannot be found by the control system either due to the lack of data or incorrect data on the network. Evidently, the former reason is determined only by the reliability of the system and does not depend on a control method. Though the probability of a network disintegration is not great, the optimal control inside each newly formed subnetwork becomes of particular importance and makes the use of centralized control systems impossible.

To control a network, algorithms described in 1, 2 could be used, but their direct application encounters certain technical difficulties.

Let us see how a network can be controlled, whose trunk lines and units do not fail and possess unlimited traffic capacity. In this case it is enough to indicate the shortest paths connecting each pair of units, so that, in accordance with decentralization, neither the initial unit nor any transit one would know the list of further transit units, and only one trunk line leading to adjacent units would be selected, i.e. the route matrixes should be pointed out.

Let us indicate the shortest paths leading to unit $V_i$ (Fig.1) and introduce integers which would denote the shortest path from a respective unit to $V_i$. The shortest path is the minimum number of transit units. Thus $V_i$ will be denoted by 0, the adjacent units by 1, etc. If the above integers are assigned to the network units, the network possesses $V_i$-relief and the integers will be called heights and denoted as $v_{ij}$.

The second index is a unit to which the height is assigned, and the first one is the relief index. All the other reliefs are constructed similarly, and their number is equal to the number of network units. Thus each unit has a set of $N$ heights $v_{ij}$ ($j=1,2,\ldots,N$), $v_{ij}=0$. Each unit $V_j$ knows the heights of the adjacent units. This information is presented as matrix $M(V_j)$ consisting of $N$ rows and having a number of columns equal to the number of units adjacent to $V_j$. The lines' indexes $V_1, V_2,\ldots,V_N$ correspond to the unit numbers. The column indexes correspond to the numbers of units adjacent to $V_j$. The matrix element $f_{jkl}$ characterizes the minimum number of transit trunk lines leading to unit $V_k$, going from $V_j$ to adjacent $V_l$.

The method of selecting the outgoing trunk line in each initial and transit unit naturally results from the relief definition. If a message is to be transmitted from $V_j$ to $V_i$, the minimum element is selected in row $V_j$ of matrix $M(V_j)$. The element's second index determines the next transit unit. In the next unit again the minimum element is selected in row $V_i$, and so on. Thus all the units are connected and the message reaches the address.

So, the knowledge of the structure is not necessary to transmit information along the shortest ways in such an ideal network. The formation of the relief requires no complete data on the network either. Let each unit $V_i$ have a device selecting the minimum in each row of matrix $M(V_i)$.
Each of the number obtained is increased by a unit and written down in the column with index $V_i$ of all the adjacent units. This set of numbers is heights of all the reliefs of unit $V_i$ increased by a unit.

Thus the finite number of steps depending on the capacity and competency of the graph denotes the correct relief, i.e. a relief ensuring transmission of messages along the shortest paths. The algorithm of relief formation operating after the correct relief was established does not alter the heights.

Of course, for the idealized network under consideration the very notion of control has no sense.

In case of serious trunk line disruptures and disruptures of units, requiring, as a rule, long restoration time, the search of a path in a network with constant relief does not secure a good quality of connection. If a trunk line fails, the graph relief will be wrong. Let trunk line $l_{12}$ be out of order (Fig.1); then, $h_{v_1v_2} = 1$ is not the correct height, since unit $V_2$ is at least three trunk lines apart from $V_1$. If the relief-changing device continues to operate, the relief will be altered. For a certain period of time, depending on the rate of the control device operation, a correct relief will be formed on the network, corresponding to its new structure. The minimum is, of course, selected from the elements of the columns corresponding to the functioning trunk lines only. If the relief is not yet formed, the path may not be found at all and the message will pass from one unit to another for a long time and occupy new channels. To avoid this, each message is given, besides the address, a special label pointing to the number of passed units $T_r$. If this number exceeds the allowed level, $T_r \text{ max}$, the message will be turned down and all the channels occupied by it will be freed. In the example with trunk line $l_{12}$ disruption, the selection of the minimum in unit $V_2$ will use only a set of heights of column $V_3$, matrix $M(V_2)$, and no information of column $V_1$ will be used. A new correct $V_1$ relief is represented by bracketed numbers (Fig.1). As soon as trunk line $l_{12}$ is restored, the relief is altered. It is insignificant what numbers are contained in column $V_1$ of matrix $M(V_2)$ at the moment of putting trunk line $l_{12}$ into operation. It may be assumed that the trunk line has just been constructed and put down both in column $V_1$, matrix $M(V_1)$, and column $B_2$, matrix $M(V_2)$. The instruction to change the relief is transmitted along the same trunk lines as ordinary messages. Assuming the traffic capacity of the trunk lines is unlimited, the control system learns how to find the shortest paths in the altered network. The less time is required to exchange control information and the more accurate the account of changes in the network is, the less is the number of call losses.

Relatively rare disruptures of trunk lines are easy to be taken into account and, if the operation rate of the controlling device is high enough, the relief is, as a rule, correct. The account of too heavy traffic of the trunklines represents more difficult problem, since the frequency of occupation and liberation of trunk lines is comparable with the controlling device operation rate, and the relief has not enough time to adapt itself to the quickly changing situation. Let us exemplify this case. In each second let all the channels of trunk line $l_{15}$ (Fig.1) be occupied with probability $P_{oc}$, and the next instant at least one of the channels is freed with probability $P_{free}$; $h_{v_1v_5} = 3$ - the unit height at moment $t$ in the conditions of complete occupation of the trunk line, appears to be wrong by moment $t + 1$ with probability $P_{free}$; $h_{v_1v_5} = 1$, the height, calculated in the conditions of the availability of free channels, is wrong at a randomly taken moment with probability $P_{oc}$. If the mean time of the trunk line complete restoration and the time at which at least one channel is free are of the same order and do not differ greatly from the period of the controlling device operation, then the probability that the measured height carries wrong information is great.

In order to take into account the network statistical properties, it is expedient to introduce "inertia" into the control system to consider a "mean" height of the units.

In real networks the stationarity of flows arriving at one and the same trunk line is usual-
ly not observed. Hence, for a system to take into account the statistical changes of calls, some optimal inertia is required. The averaging of the height through steps depends on additional memory. To spare the memory, it is expedient to introduce inertia, with the height given by

\[ h_{v_k}(t-m+1) = \frac{h_{v_k}(t-1) \cdot m + h_{v_k}(t)}{m+1} \] (1)

In this case the height at instant \( t \) depends only on the "averaged" height at instant \( t-1 \) and the instantly measured height at moment \( t \). Parameter \( m \) plays the role of inertia and does not allow the heights to "oscillate" under the influence of arbitrary short-lived changes of traffic.

Since the analytical study of a network, controlled as described above, is practically impossible, its features were studied with the help of computer modelling.

The programme is compiled so that the system could be studied in different conditions.

The traffic of the network is given by gravity matrix \( T = t_{ij} \), whose element denotes the average number of calls for connecting unit \( V_i \) with \( V_j \). Conversation time distribution is given by a table. The changes in gravitation and conversation time matrices allow to follow up the effectiveness of the traffic control system.

The network reliability is given by the total time of the trunk line work for refusal and mean time of its renewal \( \tau_{\text{ref.}} \), \( \tau_{\text{ren.}} \), these time distributions are also given by tables. The changes of these parameters reveal the dependence of the network operation quality on the reliability.

Fig. 2 shows a network being simulated. The elements of matrix \( t_{ij} \) were given by approximately proportional powers of units \( V_i \) and \( V_j \).

The number of demands per hour was 1000. The average conversation time was the same for all the units and equalled 5 min. All the trunk lines were considered equally reliable and failed once a day, the average renewal time was 1.25 hours. The number of channels in the general variant in each trunk line direction was 6. The maximum number of transits for each signal was assumed to be 10 (\( T_{\text{max}} = 10 \)).

Fig. 2

Simulated first was a network, with the flows distributed in accordance with matrixes formulated in advance and not changing during the operation, i.e. the flows were distributed along the shortest paths with available by-passes in case of extremely heavy traffic or disruption of the main direction, but the order of selection of by-passed did not change in the process of operation.

For the time corresponding to one hour of the network operation in real time, the following data were obtained:

a) \( P_{\text{ref.}} \), the mean frequency of refusals;

b) \( 1 \), the mean length of the path travelled by a message;

c) \( P_{\text{suc.term.}} \), the average amount of successfully terminating conversations.

Experiment 1: \( P_{\text{ref.}} = 0.157; P_{\text{suc.term.}} = 0.92; 1 = 3.05 \). The refusals may be due to both a trunk line disruption and limited traffic capacity. It is interesting to know the number of refusals caused by each of the reasons. For this purpose two experiments were carried out. In experiment 2 the \( \tau_{\text{ref.}} \) was assumed to be \( \tau_{\text{ref.}} \) (Table 1).

In experiment 3 the trunk line traffic capacity was increased 5 times (\( \beta = 30 \)), the reliability and traffic was the same as in the experiment 1. It may be assumed that in experiment 3 all the refusals were due to a trunk line disruption.
This is confirmed by experiment 4, carried out in the same conditions as experiment 3, but the number of the channels in the trunk line was less (β = 12). The coincidence of results of experiments 3 and 4 allows to think that a further increase in the trunk line traffic capacity does not improve the quality of the network operation.

Comparing the results of the experiments with infinite (3) and limited (experiment 1) trunk line traffic capacity, we come to a conclusion that about 6.4% out of 15.7% of all the refusals are due to the insufficient traffic capacity, while the frequency of a trunk line disruption is the same. At the same time comparison of the results of experiments 3 and 1 shows that the lack of traffic capacity is due to a decrease of the traffic capacity of the network as a whole when the trunk lines are out of order.

The results of the experiments were used to evaluate the effectiveness of the dynamic adaptive control of a relief-employing network.

When simulating relief changes, it was assumed that the controlling devices in a real network operate simultaneously. This means that the calculation of a new height by formula (1) and exchange of the control information between the units occur in the network at one and the same time. The units exchange control information once in half a minute. At first it may seem that the synchronization of the controlling devices is a limitation, but in fact this is not so. Let 0.5 min be a period of operation of a slowest device. The information coming from this device is important for all the network, and a faster device cannot be employed. Synchronizing the operation of all the devices by the greatest period possible, we shall get a worse quality in the model than in the dynamically controlled real network.

It is natural to assume that the operation rate of a controlling device within a decrease in the number of trunk lines outgoing from this unit. A high-rate device is to be placed in a unit in which a trunk line disruption may greatly change the relief, and this is a unit with the minimum number of outgoing trunk lines.

The results of the experiments carried out in the same conditions of the traffic, traffic capacity and reliability as experiment 1 and at different values of the inertia parameter m, are given in Table 2.

### Table 1

<table>
<thead>
<tr>
<th>exp</th>
<th>β</th>
<th>Pref.</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6</td>
<td>0.006</td>
<td>2.56</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0.093</td>
<td>2.99</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.090</td>
<td>2.96</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>m</th>
<th>Pref.</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>3.05</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3.06</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>3.06</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>3.03</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>3.09</td>
</tr>
</tbody>
</table>

The minimum refusal probability is 0.06.

Comparing this result with the refusal probability obtained in experiment 2 at $\beta = 30$ we see that in this case no increase of the trunk line traffic capacity ensures the quality of operation reached when the proposed method of dynamic control is applied.

The qualitative dependence may be explained in the following way. At low values of m the system considers quickly and well trunk line disruptions but at the same time is unstable in respect to random changes in the traffic. At high values of m the control system reacts to trunk line disruptions too slowly though it distributes the flows well, providing a constant network structure, since it makes a good statistical consideration of the traffic in different directions. The consideration of these two processes requires different inertia. It may be assumed that with a decrease of the traffic the optimal value of m decreases, and with the decrease of the loss probability it increases. The existence of various processes, adaptation to each of which differentely depends on m, means a possibility of the existence of two minimums of function P_{ref}.

The optimal value of inertia parameter may be obtained only in experiments. Since the conditions in a real network may change in the process of operation, m worked out in advance may not be the best choice after some time. Therefore it is important to introduce inertia changing with altering conditions of the network operation.
When calculating some height $h_{v_i, v_j}$, the inertia must decrease with a decrease in the distance between the units $v_i$ and $v_j$. This becomes clear if one takes into account the fact that the less the distance between the units, the less is the number of possible paths (in respect to the relief) from $v_j$ to $v_i$, the sooner the data on these paths will reach $v_i$. On the contrary, the data on the changes of the network will reach a far-off unit later and reflect the real situation worse. Hence, in this case the averaging of the heights must be over a larger number of steps, and the inertia is to be greater. This allows to introduce the inertia parameter $m$ as a monotonously increasing function of the height, for instance

$$H(t) = \frac{H(t-1) + h(t)}{2}, 0 < \alpha$$

The results of Table 2 were obtained on the assumption that in each trunk line there is a special additional channel for transmitting control information. The data of experiment 10, carried out in the same conditions, as experiment 6 but without a special channel, shows the channel's significance.

Experiment 10. $P_{\text{ref.}} = 0.09; P_{\text{suc. term.}} = 0.92; \lambda = 3.05$. Hence, when designing a network in which heavy traffic is expected, it is expedient to single out a special channel for transmitting control information. For networks in which the main difficulty for information transmission is represented by frequent structural changes and no heavy traffic is expected, the absence of the special channel does not affect the operation quality. An important characteristic of the system operation is represented by the changes in the number of refusals, depending on the distance between the units. Table 3 shows that an increase in the distance between the units by 1 in an uncontrolled network leads to an increase in refusal probability by about 6%, while in a dynamically controlled system this increase is about 3%.

<table>
<thead>
<tr>
<th>Distance</th>
<th>Uncontrolled network</th>
<th>Controlled network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P_{\text{ref.}}$</td>
<td>$P_{\text{ref.}}$</td>
</tr>
<tr>
<td>1</td>
<td>8.1</td>
<td>3.0</td>
</tr>
<tr>
<td>2</td>
<td>14.8</td>
<td>5.5</td>
</tr>
<tr>
<td>3</td>
<td>18.7</td>
<td>8.4</td>
</tr>
<tr>
<td>4</td>
<td>25.6</td>
<td>11.2</td>
</tr>
</tbody>
</table>

Hence, the relative effectiveness of the described method grows with the network volume. The above methods were obtained with the help of a programme which allowed modelling of networks in which the number of units could not exceed 29, and the number of channels in each trunk line was 30. At present, experiments are being conducted by a programme allowing to simulate networks with the number of units reaching 120, the number of trunk lines not exceeding 1170 and the number of channels in each trunk line being not more than 255.

REFERENCES


