CYBERNETICS AND PHYSICS, VOL. 11, NO. 3, 2022, 145–150

# AN APPROACH FOR OBTAINING ESTIMATION OF STABILITY OF LARGE COMMUNICATION NETWORK TAKING INTO ACCOUNT ITS DEPENDENT PATHS

**Boris Melnikov**
Department of Computational Mathematics
and Cybernetics,
Shenzhen MSU – BIT University,
China
bormel@smbu.edu.cn, bormel@mail.ru

**Yulia Terentyeva**
Department of Communication Network,
Center for Information Technologies and Systems
for Executive Authorities,
Russia
terjul@mail.ru

Article history:
Received 08.08.2022, Accepted 25.10.2022

**Abstract**

The physical data layer transmits bits over physical communication channels, such as coaxial cable or twisted pair. That is, it is this level that directly transmits data. At this level, the characteristics of electrical signals that transmit discrete information are determined. After that it is necessary to consider the control of the communication network, its various algorithms.

When designing a communication network, a prerequisite is to calculate its stability, and in the case of large-scale communication networks, this is a big problem. The most common deterministic, as well as fairly fast approximate method, which is often implemented at the present time, is the method by which the stability of the communication direction which is estimated by analyzing independent paths only. The main disadvantage of this method is obtaining an understated estimate of stability due to unaccountable dependent routes of communication directions. And this leads to inefficient use of resources.

Our proposed methodology allows to take into not only independent paths, but also dependent ones, which is the basis for obtaining a significantly more correct estimate. It is based on an algorithm for checking the presence of a certain path and, based on it, an algorithm for obtaining an exact assessment of stability. The paper also provides analytical and statistical analysis of the considered algorithms. In particular, a special parameter was introduced that characterizes the probability of a failure event of communication lines, in which the number of failed communication lines lies in a certain specified range; after which a study of the function describing this parameter was carried out.

**Key words**

Communication network, stability, estimation, algorithms, heuristics, information direction.

## 1 Introduction

The physical data layer transmits bits over physical communication channels, such as coaxial cable or twisted pair. That is, it is this level that directly transmits data. At this level, the characteristics of electrical signals that transmit discrete information are determined, for example, the type of encoding, data transfer rate, etc. And after that it is necessary to consider the control of the communication network, i.e., its various algorithms.

When designing a communication network, the calculation of its stability is mandatory requirement. In the case of large-scale communication networks, this becomes a problematic issue, which is caused by the NP complexity of algorithms for obtaining the correct stability estimate.

The problems of obtaining an assessment of the stability of the information direction of communication stimulated the development of correct and approximate methods, see monographs [Trivedi, 2008; Nazarov and Sychyov, 2011] and papers [Al-Kuwaiti et al., 2009; Pokorni and Jankovic, 2011; Wang et al., 2013; Bailis and Kingsbury, 2014] etc.

- The *exact* methods, i.e., the following ones:
  - the brute force method, i.e., the full iteration of states (vertices of the graph of the communication network);
  - the decomposition of a Boolean function of a special relative element;

  – the calculation by a set of routes (taking into account the identification of intersecting fragments of the routes).
- The *approximate* methods:
  – the method for obtaining upper and lower estimates for a limited set of routes and cross-sections;
  – as well as a method of statistical tests.

It should be noted that for large-scale communication networks, the time required to obtain a stability assessment by exact methods becomes critical. Therefore, these methods require a significant increase in technical resources for their implementation by parallelizing the calculation process, as well as increasing the power of computing facilities. Due to these circumstances, such methods are not effective for operational comparison of various options for designing a communication network according to the network quality criterion determined by its stability. Therefore, in reality, as a rule, *approximate methods* of obtaining an estimate *are used*.

Currently, approaches to the approximation of this estimate can be differentiated into non-deterministic and deterministic. The method of statistical tests used Monte Carlo algorithm is *non-deterministic*. However, the cardinal disadvantage of using this method for large-scale communication networks is the fact that in order to obtain an adequate precision of the result, the more experiments must be carried out, the lower the operational readiness coefficient of communication lines. Besides, the growth rate of the required number of experiments will be exponential, which is a factor in a significant decrease in the effectiveness of this approach to the point that in this case, even the exact methods will have less complexity. Moreover, the estimation of amount obtained for a conditionally real number of experiments will not be valid. Finally, when using the Monte Carlo approach, there is always a possibility of getting an overestimate.

The most common *deterministic* and also fast approximate method, which is often implemented at the present time, is the method by which the stability of the communication direction is estimated by analyzing *independent paths only*, see [Trivedi, 2008; Al-Kuwaiti et al., 2009; Nazarov and Sychyov, 2011; Wang et al., 2013] etc. The main disadvantage of this method is obtaining the underestimated assessment of stability due to unaccountable *dependent* routes of communication directions. And this in turn leads to inefficient use of resources intended for the design and modernization of the communication network. This is due to the fact that when using the methodology for assessing the stability of a communication network based on the analysis of only independent routes, normative indicators of stability ([Nazarov and Sychyov, 2011], see also [Bailis and Kingsbury, 2014]) can be achieved with a larger number of communication lines than when using the methodology taking into account dependent routes. This is an economic factor and is accompanied by costs for the construction and operation of these lines connections.

Thus, when considering the stability of a communication network as a criterion of its quality, tools are needed to obtain as an exact assessment of it as possible in order to optimize the costs of construction and modernization. However, the methodology proposed in this paper considers almost the same definitions of the reliability of the communication network, but also allows to take into account *dependent routes*, which is the basis for obtaining a significantly more exact estimate.

In further publications on this topic, we propose, firstly, to continue the analytical and statistical analysis of the method proposed here, and, secondly, to transfer the theoretical and practical results considered here to those situations where it is necessary to analyze the presence of several independent paths. Often, the number of required independent paths is set in advance, but sometimes it can also be obtained based on some other parameters of the communication network.

Here are the contents of the paper by sections.

Section 2 includes designations. In it, we present not only the standard notation of graph theory that we use, but also our interpretation of these notations, i.e., the specific variants of their writing in this paper.

The contents of Section 3 it is clear from its title: we define the states of operability of the network as elementary events and also, for them, the complete group of events. And using the full probability formula, we obtain the probability of connectivity.

In section 4, we present our interpretation of the algorithm for checking the existence of a path in the given graph.

In section 5, we present an algorithm for obtaining an exact assessment of the stability of the information direction of communication based on the constructed complete group of events. It is also important to remark that we informally use the probability space; we are going to define it strictly in the following paper.

Section 6 is the conclusion. In it, we announce the directions of furthers work and publications on this topic.

## 2 Preliminaries

This section includes designations. We present not only the standard notation of the graph theory that we use [Harary, 1969; Diestel, 1997; Gera et al., 2016; Karpov, 2017; Gera et al., 2018], but also our interpretation of these notations, i.e., the specific variants of their using in this paper.

For each $n \in \mathbb{N}$, we shall denote

$$\overline{1,n} = \{1, 2, \ldots, n\}$$

Thus, let us consider an arbitrary communication network, which is defined by the graph

$$G = (V, E), \tag{1}$$

where

- $V$ (usually $V = \overline{1,n}$) is the set of vertices of the graph (network nodes), $n \in \mathbb{N}$ is the number of vertices;
- $E$ is the set of edges of the graph (communication lines),

$$E \subseteq \bigl\{ \{ v_i, v_j \} \bigm| 1 \leqslant v_i \leqslant n,\ 1 \leqslant v_j \leqslant n,\ i \neq j \bigr\}.$$

We shall sometimes use other notations, i.e. assume, that all the elements of the set $E$ are numbered, i.e.,

$$E = \Bigl\{ e_k = \bigl\{ v_{\alpha(k)}, v_{\beta(k)} \bigr\} \Bigm| k \in \overline{1,m} \Bigr\}, \qquad (2)$$

where $m \in \mathbb{N}$ is the number of edges. From the context, it will always be clear which designations are used, there will be no contradictions.

We shall also consider a given function on the edges of the graph

$$F_R : \overline{1,m} \to [0,1].$$

This function defined for each edge of the graph is considered as the readiness coefficient of the communication line of the network, the edge number according to (2). In practice, this readiness coefficient is usually calculated as the ratio of the average time of failure to the total average time of failure and recovery time.

As usual in the tasks of analyzing communication networks, we will understand by the *information direction* of the communication network *two its nodes*, one of which generates information intended for targeting transmission to another communication node. To obtain a *quantitative value* of the stability of the information direction of the communication network, two vertices of the graph are fixed, specifying the information direction of communication. These vertices will play the role of poles, relative to which stability will be calculated. Below, they will be designated $u', u'' \in V$, sub-scripted $v$ will usually not be used for them.

## 3 States of operability of the network as elementary events and the complete group of events

The contents of this section it is clear from its title: we define the states of operability of the network as elementary events and also, for them, the complete group of events. After that using the full probability formula we obtain the probability of connectivity.

It is natural to define an event as the operability of some subset of the set of communication lines:

$$B_i = \Bigl\{ \bigl( b_1^{(i)}, b_2^{(i)}, \ldots, b_m^{(i)} \bigr) \Bigm| b_k^{(i)} \in \{0,1\} \text{ for } k \in \overline{1,m} \Bigr\}. \qquad (3)$$

Here, each $B_i$ characterizes the state of operability of the objects of communication network, and

$$i \in \overline{1, 2^m}.$$

Let us note about this the following things.

- In fact, we can say that we consider the *probability space* here, despite we do not strictly define it. The corresponding strict definitions are more or less obvious, they are entirely based on the notation given here, and perhaps we shall give them in the continuation paper. However, we shall use this term in the rest of this paper, and understanding should not cause difficulties.
- Because the concepts "the state of operability of the network" and "the elementary event" are almost the same, we shall often call $B_i$ simply *"a state"*.

The probability of each state is calculated as the product of the probabilities of the states of the objects of communication network based on the probability theorem of the product of events. The totality of all the possible states $B_i$ covers all the possible states of the objects of network.

Let us explain this thing in other words. For a specific event corresponding to a state $i$:

- the numbers of communication lines that are *in working condition*, coincide with $k$'s (as before, $k \in \overline{1,m}$), for which $b_k^{(i)} = 0$ according to the entered numbering (2);
- vice versa, a state, which is *in not-working condition*, coincides with $k$, for which $b_k^{(i)} = 1$.

As it is easy to see, the set of events forms *a complete group of events*, since they are incompatible in pairs and the appearance of one and only one of them is a reliable event, see [Trivedi, 2008; Blitzstein and Hwang, 2019] etc.

As we said before, the probability of a working state for each communication line is determined by the readiness coefficient $F_R(k)$, where $k \in 1, 2, \ldots, m$ is the number of the edge (communication line) according to the numbering (2). From here, it is easy to get the probability of each event $B_i$ defined by (3), i.e.,

$$P(B_i) = \prod_{k=1}^{m} F'_R(k), \qquad (4)$$

where

$$F'_R(k) = \begin{cases} F(k), & \text{if } b_k^{(i)} = 0, \\ 1 - F(k), & \text{if } b_k^{(i)} = 1. \end{cases} \qquad (5)$$

Let $A_{u'u''}$ (or simply $A$) be an event consisting in the fact that there will be connectivity between the given poles $u'$ and $u''$. Then, according to the full probability formula, the probability of connectivity, which in this case is, will be equal to

$$P_A(G, u', u'') = \sum_{i=1}^{2^m} P(A|B_i) \cdot P(B_i), \qquad (6)$$

or, applying usual probability-theoretic notation, simply $P(A)$.

Certainly, $P(A|B_i)$ is the probability of connectivity between the poles, provided the event $B_i$ occurs. According to the meaning of the concepts introduced earlier, this probability is as follows:

$$P(A|B_i) = \begin{cases} 1, & \text{if there exists at least one path} \\ & \text{between the poles, provided that} \\ & \text{the following edges are excluded} \\ & \text{from the connection graph } (V,E): \\ & \text{according to the numbering (2),} \\ & \text{their numbers are equal} \\ & \text{to an index } k \ (k \in \overline{1,m}), \text{ such that} \\ & \text{in the vector } \left(b_1^{(i)}, b_2^{(i)}, \dots, b_m^{(i)}\right), \\ & \text{we have } b_k^{(i)} = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Let us note in addition, that it is possible to reduce the dimension of the probability space based on:

- the topology of the communication network defined by the graph $(V,E)$,
- and also the fact that there exist nodes and communication lines, which, due to this topology, cannot be part of the path connecting the considered poles $u'$ and $u''$.

Such nodes are, for example, vertices that have the only adjacent vertex (with the exception of the considered poles). It is advisable to carry out the procedure of reducing the dimension of the corresponding probability space for each information direction of the communication.

## 4 Algorithm for checking the existence of a path

In this section, we present our interpretation of such an algorithm.

It would not be a big overstatement to say that the material in this section is simple and has been known for a long time. However, the following comments can be made to this.

- Firstly, it was on the basis of the algorithm proposed in the section that we built real computer programs included in the complex to assess the reliability of communication systems. We are going to write about such program complex in the following paper.
- Secondly, the algorithm proposed here is "tailored" to our situation, i.e., when the "good" edges are unknown in advance.
- Thirdly, even with such a simple algorithm, not everything is completely clear, see some consideration in [Melnikov et al., 2022].
- Fourth, it is the variants of algorithms considered in [Melnikov et al., 2022] for a completely different task can be adapted to the case of the problem discussed in this paper.
- Fifth, there is no need to find the path itself, but only to establish the fact of its existence; this is what is necessary for the theory described above. This fact reduces the complexity of the algorithm compared to classical algorithms for finding paths between the vertices of the graph
- Sixth, based on the algorithm proposed in this section, we shall describe similar algorithms in the following publications that check for the presence of a predetermined number of paths (more than one)...

Thus, let us consider the algorithm.

**Algorithm 4.1.** *Determining whether there exists at least one path in the graph between the specified vertices $u'$ and $u''$ if an event $B_i$ occurs.*

*Input:*

- graph $G = (V,E)$ (1);
- its poles $u', u'' \in V$;
- event $B_i$; as before, we assume that $B_i$ is described by (3).

*Output:* "yes" (i.e., there exists the required path), or "no" (otherwise).

*Notation:* Boolean result of this algorithm will be denoted by

$$\mathcal{A}(G, u', u'', B_i).$$

*Method.*

*Step 1.* From the graph $(V,E)$, we exclude edges whose numbers, according to numbering (2), coincide with the indices $k \in \overline{1,m}$, such that $b_k^{(i)} = 1$.

*Step 2.* Define two variable sets of vertices $V_1, V_2 \subseteq V$; their initial assignments will be made according to the following formulas:

$$V_1 := \{u'\}, \qquad V_2 := V \setminus \{u'\}.$$

*Step 3.* If for all vertices included in $V_1$, there are no adjacent vertices in $V_2$, then exit (output "no").

*Step 4.* In $V_2$, we find adjacent vertices for all vertices included in $V_1$; let these vertices forms the set $V'$. Then we set

$$V_1 \mathrel{+}= V', \qquad V_2 \mathrel{-}= V'.$$

*Step 5.* If $u'' \in V_1$, then exit (output "yes"), else go to Step 3.

*End of the algorithm description.*

## 5 Algorithm for obtaining the correct stability estimate

Now, let us present an algorithm for obtaining an exact assessment of the stability of the information direction of communication based on the constructed complete group of events (formulas (2)–(6); let us also remark that we informally use the probability space that will be defined strictly in the following paper).

We note in advance, that this algorithm demonstrates an approach to obtaining an estimate that will be used in the future, i.e., when constructing an algorithm to obtain an estimate of the stability of large-scale communication networks (not only their information direction).

**Algorithm 5.1.** *Calculation of the assessment of the stability of the information direction of the communication network determined by the poles* $u'$ *and* $u''$.

*Input:*

- graph $(V, E)$ (1);
- its poles $u', u'' \in V$.

Without loss of generality, we assume that for each vertex of the set $V \setminus \{u', u''\}$, the number of the adjacent vertices is 2 or more.

*Output:* the assessment of the stability.

*Auxiliary variables:*

- vector
$$W = (w_1, w_2, \ldots, w_m) \in \{0, 1\}^m ,$$
where, as before, $m$ is the number of edges;
- real value $p$.

*Method.*

*Step 1.* Let
$$p := 0, \qquad W := (0, 0, \ldots, 0).$$

*Step 2.* Consider graph $G' = (V', E')$, which is obtained from the original one (1) by removing the following edges. According to (2), the removed edges correspond to some numbers $e_k$ (where $k \in \overline{1, m}$), for which the condition $w_k = 1$ holds.

*Step 3.* If $\mathcal{A}(G', u', u'', W)$, then
$$p \mathrel{+}= P_A(G', u', u'').$$
(the necessary designations were introduced above, see (6) and Algorithm 4.1).

*Step 4.* We form the next value of the vector $W$. This can be done, for example, by adding 1 to the number for which this vector is a binary notation. If it succeeded (i.e., if we have not looked at all possible vectors yet), then go to Step 2.

*Step 5.* Exit, output value $p$.

*End of the algorithm description.*

It should be noted that the undoubted advantage of this algorithm is that it gives an exact estimate taking into account all dependent paths connecting the specified vertices. However, it will be effective for communication networks of small scale only. For large-scale communication networks with modern computing capacities of PCs (we shall not consider parallelization effects here), this algorithm will be unacceptable due to incommensurable time indicators. But it can be optimized based on the properties of the distribution of a random variable equal to the number of failed communication lines, as well as on a specially constructed structure of clusters of probability space; in the following publications, we are going to return to such possible optimizing of the algorithm.

## 6 Conclusion

Thus, in this paper, a method is proposed for finding the reliability of a communication network taking into account *dependent* paths, which allows for a more exact assessment compared to widely used methods based on the search for independent paths. The relevance of obtaining a more exact assessment is dictated by the concept of resource savings in the design and / or modernization of communication networks, especially large-scale ones. Since in order to achieve the reliability indicators of the required values, it is necessary to build new communication lines, which is very expensive in large areas with different climatic conditions.

For now, we have described only a possible approach to the definition of stability of large communication network taking into account is dependent paths (as it was said in the title). We can say that the subject of the paper is the formulation of its main algorithm (i.e., Algorithm 5.1), and not the results of its work. Simplifying it even more, we can say that the subject of the paper is the formula (6) of the full probability, i.e., its application in our field. Before, only the set of independent paths was used.

Despite all that was said in the previous paragraph, computational experiments, of course, have been carried out and are currently being carried out. Their description, as well as their detailed analysis (first of all, the analysis of Algorithm 5.1, will be given in the following publications.

In addition to those possible continuations of the subject of this paper, which have already been mentioned above, we shall focus on the following. In Algorithm 5.1, the next vector of 0's and 1's is determined "simply", i.e., by the previous one, without some its clarification; more precisely, in the description of the algorithm, we simply suggest to add 1 to the previous binary number. But if we use Gray's algorithm ([Reingold et al., 1977] etc.), then it is possible to save the information obtained by auxiliary algorithms (algorithm

$\mathcal{A}(G, u', u'', B_i)$, but not only by it) for use in the next iterations. Apparently, it will be possible to use most of this information, which should significantly save the time of the overall operation of the algorithm.

Another advantage of the proposed method is that it is an *anytime algorithm* [Melnikov et al., 2009; Melnikov and Sayfullina, 2013; Melnikov et al., 2018; Grandcolas and Pain-Barre, 2022; Yu and Oh, 2022]. Moreover, it provides for setting the precision with which the reliability assessment will be obtained.

In further publications on this topic, we are going to present the results of computational experiments. Besides, as we said in introduction, we propose:

- firstly, to continue the analytical and statistical analysis of the method proposed here;
- and, secondly, to transfer the theoretical and practical results considered here to those situations where it is necessary to analyze the presence of several independent paths.

We especially note that there is no contradiction in the last item:

- we really offer a method for analyzing the considered problems using *dependent* paths,
- but at the same time, *independent* paths may appear in various *formulations* of these problems.

## References


Al-Kuwaiti, M., Kyriakopoulos, N., and Hussein, S. (2009). A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability. *Communications Surveys & Tutorials*, **11** (2), pp. 106–124.

Bailis, P. and Kingsbury, K. (2014). The network is reliable. *ACM Queue*, **12** (7), pp. 1–20.

Blitzstein, J. K. and Hwang, J. (2019). *Introduction to Probability (2nd Edition)*. Chapman and Hall / CRC, London.

Diestel, R. (1997). *Graph theory*. Springer-Verlag, Heidelberg.

Gera, R., Hedetniemi, S., and Larson, C. (2016). *Graph theory. Favorite Conjectures and Open Problems – 1*. Springer, Berlin.

Gera, R., Hedetniemi, S., and Larson, C. (2018). *Graph theory. Favorite Conjectures and Open Problems – 2*. Springer, Berlin.

Grandcolas, S. and Pain-Barre, C. (2022). A hybrid metaheuristic for the two-dimensional strip packing problem. *Annals of Operations Research*, **309** (1), pp. 79–102.

Harary, F. (1969). *Graph theory*. Addison-Wesley Publ., Massachusetts.

Karpov, D. (2017). *Graph theory*. Publishing House of the St. Petersburg Branch Mathematical Institute named after V. A. Steklov of Russian Academy of Sciences, Saint Petersburg (in Russian).

Melnikov, B. F., Tsyganov, A. V., and Bulychov, O. I. (2009). A multi-heuristic algorithmic skeleton for hard combinatorial optimization problems. *In: Proceedings of the 2009 International Joint Conference on Computational Sciences and Optimization, CSO-2009*, pp. 33–36, 5193637.

Melnikov, B. F. and Sayfullina, E. F. (2013). Application of a multiheuristic approach for the random generation of a graph with a given degree vector. *News of higher educational institutions. Volga region. Physical and mathematical sciences*, 3 (27), pp. 70–83 (in Russian).

Melnikov, B. F., Melnikova, E. A., Pivneva, S. V., Dudnikov, V. A., and Davydova, E. V. (2018). A multi-heuristic algorithmic skeleton for hard combinatorial optimization problems. *In: CEUR Workshop Proceedings, 2212*, pp. 312–321.

Melnikov, B. F., Starikov, P. P., and Terentyeva, Yu. Yu. (2022). About one problem of analyzing the topology of communication networks. *International Journal of Open Information Technologies*, **10** (6), pp. 1–8 (in Russian).

Nazarov, A. N. and Sychyov, K. I. (2011). *Models and methods for calculating the quality indicators of the functioning of node equipment and structural and network parameters of next-generation communication networks*. OOO Polikom Ed., Krasnoyarsk (in Russian).

Pokorni, S. and Jankovic, R. (2011). Reliability estimation of a complex communication network by simulation. *In: IEEE Telecommunications Forum*, pp. 226–229.

Reingold, E. M., Nievergelt, J., and Deo, N. (1977). *Combinatorial Algorithms: Theory and Practice*. Pearson College Ed., London.

Trivedi, K. S. (2008). *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*. John Wiley & Sons, New York.

Wang, Zh., Yan, X., Ge, W., and Zhang, X. (2013). Analysis of network reliability on intelligent substation. *In: IEEE International Conference on Intelligent Networks and Intelligent Systems*, pp. 147–150.

Yu, W. and Oh, J. (2022). Anytime 3D object reconstruction using multi-modal variational autoencoder. *IEEE Robotics and Automation Letters*, **7** (2), pp. 2162–2169.