

Article

Control of Telecommunication Network Parameters under Conditions of Uncertainty of the Impact of Destabilizing Factors

Valerii Kozlovskiy ^{1,*}, Ivan Shvets ¹, Yurii Lysetskiy ¹, Mikolaj Karpinski ^{2,3,*} , Aigul Shaikhanova ⁴  and Gulmira Shangytbayeva ⁵

¹ Department of Information Protection System, National Aviation University, 1 Liubomyra Huzara Ave., 03058 Kyiv, Ukraine; ivan.shvets@intellecteu.com (I.S.); yurii.lysetskiy@snt.ua (Y.L.)

² Department of Computer Science, University of Applied Sciences in Nowy Sacz, 33-300 Nowy Sacz, Poland

³ Department of Cyber Security, Faculty of Computer Information Systems and Software Engineering, Ternopil Ivan Puluj National Technical University, 56 Ruska St., 46001 Ternopil, Ukraine

⁴ Department of Information Security, L.N. Gumilyov Eurasian National University, 2 Satpayev St., Astana 010008, Kazakhstan; shaikhanova_ak@enu.kz

⁵ Department of Computer Science and Information Technology, Faculty of Physics and Mathematics, K. Zhubanov Aktobe Regional University, 34 A. Moldagulova Prospect, Aktobe 030000, Kazakhstan; gshangytbayeva@zhubanov.edu.kz

* Correspondence: vvkzeos@gmail.com (V.K.); mpkarpinski@gmail.com (M.K.)

Abstract: The classification of the natural and anthropogenic destabilizing factors of a telecommunications network as a complex system is presented herein. This research shows that to evaluate the parameters of a telecommunications network in the presence of destabilizing factors, it is necessary to modify classical linear methods to reduce their sensitivity to the incompleteness of a priori information. Using generalized linear models of multiple regression, a combined method was developed for assessing and predicting the survivability of a telecommunications network under conditions of uncertainty regarding the influence of destabilizing factors. The method consists of accumulating current information about the parameters and state of the network, the statistical analysis and processing of information, and the extraction of sufficient sample statistics. The basis of the developed method was balancing multiple correlation–regression analysis with the number of regression equations and the observed results. Various methods of estimating the mathematical expectation and correlation matrix of the observed results under the conditions of random loss of part of the observed data (for example, removing incomplete sample elements, substituting the average, pairwise crossing out, and substituting the regression) were analyzed. It was established that a shift in the obtained estimates takes place under the conditions of a priori uncertainty of the statistics of the observed data. Given these circumstances, recommendations are given for the correct removal of sample elements and variables with missing values. It is shown that with significant unsteadiness of the parameters and state of the network under study and a noticeable imbalance in the number of regression equations and observed results, it is advisable to use stepwise regression methods.

Keywords: telecommunications network; system survivability; destabilizing factors; correlation and regression analysis; stepwise regression



Citation: Kozlovskiy, V.; Shvets, I.; Lysetskiy, Y.; Karpinski, M.; Shaikhanova, A.; Shangytbayeva, G. Control of Telecommunication Network Parameters under Conditions of Uncertainty of the Impact of Destabilizing Factors. *Information* **2024**, *15*, 69. <https://doi.org/10.3390/info15020069>

Academic Editors: Zoran H. Perić, Agbotiname Lucky Imoize, Cheng-Chi Lee and Webert Montlouis

Received: 9 November 2023

Revised: 18 December 2023

Accepted: 16 January 2024

Published: 23 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The fundamental difference between modern telecommunications networks—next-generation networks, future networks, and traditional networks—is, according to the standards of open systems, free access to resources. On the one hand, connecting to existing information and communication systems is greatly simplified, but on the other hand, there are new problems in protecting information from unauthorized access. This is even more critical for wireless networks.

Destabilizing factors have a detrimental effect on the stability and survivability of a network as a complex technical system. The statistical approach is the most informative approach to quantifying the impact of destabilizing factors on the characteristics of a telecommunications network. The research presented in this article used a method consisting of analyzing key performance indicators widely used for statistical analysis and forecasting the parameters and conditions of technical, economic, and social systems.

In the framework of the classical theory of system analysis, an abstract telecommunications network is considered a complex system, i.e., a system in the model in which there is not enough information for effective management. A study was conducted on various methods of accumulating information about destabilizing factors, and a combined method was developed that is suitable for monitoring the state and parameters of a telecommunications network under conditions of a priori uncertainty of the statistical characteristics of destabilizing factors.

The purpose of this research is, therefore, to manage the survivability of a new-generation telecommunications network in the presence of destabilizing factors with a priori unknown statistical characteristics.

The organization of the rest of this paper is as follows: In Section 2, similar significant works by other authors are presented. The background of the analysis of the mutual impacts of destabilizing factors and their effects on the survivability and stability of a telecommunications network is presented in Section 3. The analysis is based on correlation and multiple regression theory. In Section 4, the concept and combined statistical method of the search for missing values are described. In Section 5, the methodology and results of the experiment are presented concerning a corporate network, and the concept of key performance indicators (KPIs) is applied. Section 6 presents the conclusions and future work on the subject.

2. Related Works

Previous studies were devoted to solving the problems of ensuring the survivability, sustainability, and fault tolerance of complex systems. A.G. Dodonov, S.I. Otrokh, F. Hellmann, S.E. Dreyfus, P. Csermely, and others [1–14] have made important contributions to the theory in this field.

These problems are usually solved under conditions of insufficient (and sometimes absent) a priori information about the state and current parameters of the system under study, as well as the characteristics of useful signals, noise, interference, destabilizing factors, etc. Analysis with complete, or at least sufficient, a priori information is an unattainable ideal; in reality, it is necessary to work without a priori information [2,3]. Methods of ensuring the resistance of complex systems to the influence of destabilizing factors were considered in Dodonov et al.'s monograph [4], wherein they considered the impact of uncertainties as the main destabilizing factor deriving from the activity of malicious actors.

The scientific articles in [12,13] presented the results of methods developed to increase the structural survivability of telecommunications networks.

A model of distributed information processing under the conditions of the influence of destabilizing factors on a telecommunications network was developed in [14]. The model is partial and based on the method of functional redundancy, considering time characteristics and computational resources. Although the keywords mention destabilizing actions, the article does not specify their characteristics and results. The study by Tolubko et al. [15], a review devoted to describing the methods of assessing the sustainability of a telecommunications network under the influence of destabilizing factors, has the same limitation. Phase portraits of the model of destabilizing factors are developed. It is argued that they can be used to obtain asymptotic estimates of network stability, although practical methods and estimation algorithms are not provided.

The monograph in [16] focuses on optimizing the allocation of network resources to ensure the sustainable operation of access control protocols to the environment. A

conclusion is drawn about the stability of the whole network, although there is no proof of this statement.

The report in [17] is an International Telecommunication Union (ITU-T) standard for dealing with natural disasters to improve the stability and speed up the recovery of networks. This standard is a practical recommendation that can be followed in all cases.

The article in [18] describes the development of a method for calculating the survivability of telecommunications network channels, although the main content of the work is the operator method (Laplace method) for solving ordinary differential equations with constant coefficients. The condition of stability of the operator equation of an abstract closed system is accepted as a condition of the survivability of a network. The operator equation is stable if the poles of the function (the roots of the denominator polynomial) are in the left half-plane of the p -plane of the complex variable.

The scientific article in [19] presents the results of developing methods for increasing the structural survivability of new-generation telecommunication networks.

In conclusion, it is important to underscore that the purpose of this section was to provide an overview of the existing work on methods and tools of information security management under the uncertainty of destabilizing factors and the authors' views on the principles and main contents of these methods. The references to the literature are far from systematic.

3. Background

To understand the problem statement, the mutual impacts of destabilizing factors and their effects on the survivability and stability of a telecommunications network must be analyzed.

The problem of ensuring the stability of complex technical systems, i.e., their ability to maintain normal functioning during and after exposure to destabilizing factors (radiation, thermopower, vibration, etc., including human factors), is relevant to various applications.

The stability of a system means its ability to perform the assigned functions with the specified quality indicators under the influence of internal and external destabilizing factors. The mutual influences of destabilizing factors on each other and the stability of the system are complex.

Ensuring the resilience of complex technical systems against the effects of destabilizing factors is a long, time-consuming, and iterative process associated with different types of uncertainties.

First, this process covers all major stages of the system's life cycle, which can take decades. Secondly, determining the levels of sustainability indicators is a difficult task, much more complex than, for example, determining the levels of reliability indicators. Third, assessing the resilience of a complex technical system against the effects of destabilizing factors requires extensive knowledge in various scientific and technical fields. This is due to the need to determine the physical mechanisms of the effects of destabilizing factors on the object, for knowledge of the functioning of the complex technical system, and for the ability to predict the response of the functioning system to the impact, accompanied by different physical processes. Finally, the assessment of resilience is usually hampered by the lack of data required.

Telecommunications networks are designed to provide communication between enterprises, organizations, and all structural units of the economy and transport per the requirements and rules of technical operation.

For the sustainable operation and provision of subscribers with a given quality of telecommunication services, the state of the network and its elements is constantly monitored, and the quality of services is maintained at a given level. However, as experience shows, abnormal situations caused by the impact of destabilizing factors (DFs) and emergencies (EMs) not only reduce the quality of services provided but also negatively affect customer service.

External destabilizing factors that reduce the stability of functioning include the following:

- Emergency and fan power outages;
- Multipath interference (for wireless networks);
- Viruses and hacker attacks.
- Internal destabilizing factors that reduce the stability of systems include the following:
- Failures, including technical failures;
- Software errors;
- Unsuccessful architectural solutions;
- Inconsistencies due to the variety of characteristics of the installed equipment, not considered at the stages of design and deployment;
- Conflicts and deadlocks due to the incorrect distribution of system resources, some selected mechanisms of the organization of information and computational processes, and the architectural features of system components.

Considering the mutual influences of destabilizing factors is a non-trivial task. The statistical approach is the most accessible and suitable approach for practical applications [20,21], in particular, correlation–regression analysis.

Afifi and others applied the imputation method to replace missing data with some reasonable values, especially in the economic and financial spheres [22–25]. Mean substitution is the main method of imputation proposed for replacing missing data [22]. Afifi described this method in [21] in 1979. Imputation methods are not mathematically strict (e.g., they select the types of missing values based on missing-completely-at-random (MCAR), missing-at-random (MAR), and nonrandom missing data, and this selection is qualitative in a wide range). Therefore, multiple correlation–regression analysis with robust and resistant estimates of location and scale was used (median and bi-weight methods) in [26].

The general purpose of the multiple regression method is to analyze the relationship between several independent variables (also called regressors) and a dependent variable. Statistics represent a sample of the implementation of the values of random variables:

- The i -th sample of the numerical value of the result y_i , $i = 1, 2, \dots, N$;
- The j -th sample of the numerical value of the j -th factor x_j , $j = 1, 2, \dots, N$.

Using statistical data allows us to achieve optimal results by controlling the values of factors or to predict the possible value of the result with the formed values of factors.

The estimation is performed via observations of the input (the rows of the observation matrix X) and output (the elements of the response vector $\vec{y}$).

There is a stochastic (random) relationship between random variables, i.e., there is a correlation.

In the general case, the procedure for constructing multiple regression is to estimate the parameters of the linear equation. The functional dependence of the result on the factors is represented by the following regression equation:

$$E[\vec{y}] = X\vec{\theta}.$$

Matrices of multiple correlation coefficients and a system of multiple linear or polynomial regression equations are usually used as the main characteristics of statistical communication [18,19]. In addition, it is necessary to choose a method for approximating the repeatability curves of KPIs to automate measurements and calculations [24].

Let us consider the process of forecasting network parameters as a task of predicting the k -th dependent variable Y_k , $k = \overline{1, N}$, with M independent variables X_m , $m = \overline{1, 2, \dots, M}$; $m \neq k$. In the general case, $M \neq N$. When $m = 1$, there is a linear or polynomial regression equation of the independent variable X_m with the dependent variable Y_k . When $m > 1$, there is a system of multiple regression equations of variables $X_1, X_2, \dots, X_m$ with Y_k . (Functional rather than statistical dependence is indicated.) In

this problem, the independent variables $X_1, X_2, \dots, X_m$ are random variables that are not necessarily statistically independent.

The variable Y_k is approximated with a regression function $\psi(\cdot)$, which contains KPI estimates and unknown coefficients. The equation of the linear regression model of the independent variables $X_1, X_2, \dots, X_m$ with the dependent variable Y_k is written as follows:

$$Y_k = a_{0k} + a_{1k}X_1 + \dots + a_{mk}X_m + \varepsilon, \quad (1)$$

where ε is the approximation error.

Let $X_{1j} = X_1^j$. Then, the equation of polynomial regression is introduced in the form of

$$Y_k = a_{0k} + a_{1k}X_1 + a_{2k}X_1^2 + \dots + a_{mk}X_1^m + \varepsilon. \quad (2)$$

The parameters of the regression model are estimated by sampling the volume taken from a general population.

The sample is formed as follows. Based on the results of the network operation test, the first sample of independent variables $X_{11}, X_{12}, \dots, X_{1m}$ is recorded, and then the dependent variable Y_1 is calculated. Then, the second sample of independent variables $X_{21}, X_{22}, \dots, X_{2m}$ is fixed, and the dependent variable is calculated. The procedure continues until the N variables obtain a sample of observations.

$$\{Y_1 : X_{11}, X_{12}, \dots, X_{1m}\}, \{Y_2 : X_{21}, X_{22}, \dots, X_{2m}\}, \dots, \{Y_N : X_{N1}, X_{N2}, \dots, X_{Nm}\}.$$

The system of multiple linear regression equations takes the following form:

$$\left. \begin{aligned} Y_1 &= a_{01} + a_{11}X_{11} + \dots + a_{m1}X_{1m} + \varepsilon_1 \\ Y_2 &= a_{02} + a_{12}X_{21} + \dots + a_{m2}X_{2m} + \varepsilon_2 \\ &\dots \\ Y_k &= a_{0k} + a_{1k}X_{k1} + \dots + a_{mk}X_{km} + \varepsilon_k \\ &\dots \\ Y_N &= a_{0N} + a_{1N}X_{N1} + \dots + a_{mN}X_{Nm} + \varepsilon_N \end{aligned} \right\}, \quad (3)$$

where $\{a_{0k}, a_{1k}, \dots, a_{mk}\}$, $k = \overline{1, N}$, are unknown coefficients, and $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k, \dots, \varepsilon_N\}$ are random errors, which are logically considered normal and equally distributed with the parameters $\{0, \sigma_\varepsilon^2\}$.

To obtain estimates using the least squares method, it is necessary to minimize the sum S_k of squares of the deviations at each point. The best approximation corresponds to the minimum value of the following expression:

$$S_k = \sum_{k=1}^N (Y_k - a_{0k} - a_{1k}X_{k1} - \dots - a_{mk}X_{km})^2. \quad (4)$$

The value S_k is a measure of the error associated with the binding of available data to the selected regression model. The minimum S_k is achieved by differentiating the last expression by the coefficients $\{a_{0k}, a_{1k}, \dots, a_{mk}\}$, $k = \overline{1, N}$, equating the corresponding derivatives to zero, and solving a system of normal equations for $\{a_{0k}, a_{1k}, \dots, a_{mk}\}$. A system of equations for estimating the partial regression coefficients is obtained:

$$\left. \begin{aligned} \hat{Y}_1 &= \hat{\alpha}_{01} + \hat{\alpha}_{11}X_{11} + \dots + \hat{\alpha}_{m1}X_{1m} \\ \hat{Y}_2 &= \hat{\alpha}_{02} + \hat{\alpha}_{12}X_{21} + \dots + \hat{\alpha}_{m2}X_{2m} \\ &\dots \\ \hat{Y}_k &= \hat{\alpha}_{0k} + \hat{\alpha}_{1k}X_{k1} + \dots + \hat{\alpha}_{mk}X_{km} \\ &\dots \\ \hat{Y}_N &= \hat{\alpha}_{0N} + \hat{\alpha}_{1N}X_{N1} + \dots + \hat{\alpha}_{mN}X_{Nm} \end{aligned} \right\}. \quad (5)$$

where $\hat{\alpha}_{0k}, \hat{\alpha}_{1k}, \dots, \hat{\alpha}_{mk}$ are the estimates for $\{a_{0k}, a_{1k}, \dots, a_{mk}\}$.

Estimates are unbiased and effective, i.e., they have minimal variance for the samples $X_1, X_2, \dots, X_m$ among all linear estimates for predicting variables Y_k , $k = \overline{1, N}$.

Regression coefficients represent the contributions of each independent variable to the prediction of the dependent variable. Two opposite criteria are usually used to select the final regression equation.

To make the equation useful for prediction, the observer should seek to include as many independent variables as possible in the model so that the predicted values can be determined more reliably.

4. Description of Method

The concept of missing values is introduced in this section. When using one-dimensional methods of analysis (for example, the t-criterion [21]), the most reasonable action is to remove elements with no value of X (the analyzed variable) from the sample. However, the situation changes when using significantly multidimensional methods of analysis, i.e., when for each element of the sample, there are p observable variables $X_1, X_2, \dots, X_p$. If a sample element is missing, e.g., for a variable X_1 , it is not necessary to remove that sample element from the analysis because this results in the loss of information about the variables provided by that element. Since multiple linear regression analysis, as well as other multidimensional procedures, is based on the vector of means m and the covariance matrix S , this element can be left in the sample and the available measurements can be used to calculate the estimates of the vector of the means $\bar{x}$ and covariance matrix Σ .

Let us imagine a system of equations of the multiple linear regression model (3) in matrix form:

$$Y = \hat{X}B + E, \quad (6)$$

$$\text{where } \hat{X} = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1m} \\ 1 & X_{21} & \cdots & X_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{N1} & \cdots & X_{Nm} \end{pmatrix} \text{ is the so-called design matrix.}$$

For the regression model, it is a matrix of independent variables, supplemented with the first column of current observation weights. $B^T = \{\beta_0, \beta_1, \dots, \beta_m\}$ is the vector of regression equation parameters; $E^T = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$ is the vector of estimation errors with a multidimensional Gaussian distribution with zero vectors of mathematical expectations and a matrix of variances in the form of $\sigma^2 I$; I is a single matrix; and T is the transposition symbol. Then, Equations (4)–(6) can be combined and represented in matrix form as

$$S = (Y - \hat{X}B)^T (Y - \hat{X}B). \quad (7)$$

The least squares estimation vector is a solution of a system of normal equations,

$$(X^T \hat{X})B = XY \quad (8)$$

the solution of which has the following form:

$$B = (X^T \hat{X})^{-1} (XY). \quad (9)$$

Considering that the matrix of variances of the vector of estimation errors is described using the expression $\sigma^2 I$, the correlation matrix of the vector B is equal to $R_B = \sigma^2 (X^T \hat{X})^{-1}$.

In this case, instead of using the optimal estimate as the initial value, which is determined via the matrix equation, it is easier to look for the optimal estimate as a solution of the dual-difference equation. The coefficients of the difference equation are determined by the statistics of observations and interferences and, in the general case, are time-dependent variables. The advantage of this approach is that even if it is not possible to obtain an analytical solution of the difference equation, it is always possible to obtain its numerical solution

on a computer. Moreover, the solution can be obtained in real time, considering the newly received information about changes in the parameters of observations and interference.

Following [21], the iteration algorithm for solving Equation (6) in sequential form is constructed as follows:

$$\mathbf{F}\mathbf{B}[\hat{\mathbf{X}}(k) - \hat{\mathbf{X}}(k-1)] = \mathbf{G}[\mathbf{Y} - \mathbf{E} - \hat{\mathbf{X}}(k)\mathbf{B}], \quad (10)$$

where $\mathbf{F}$ and $\mathbf{G}$ are matrix multipliers with non-zero determinants or non-zero scalar multipliers.

These factors are chosen to ensure the maximum rate of convergence without losing the stability of Equation (10). For the optimal selection of values $\mathbf{F}$ and $\mathbf{G}$, the z-conversion operation is applied to (10):

$$\mathbf{F}\mathbf{B}[\hat{\mathbf{X}}(z)(1 - z^{-1})] = \mathbf{G}[\mathbf{Y} - \mathbf{E} - \hat{\mathbf{X}}(z)\mathbf{B}], \quad (11)$$

and the roots of the characteristic equation are calculated, which must have a modulo of less than one. Then, the general solution of Equation (11) with an unlimited increase in the number of iterations asymptotically converges to the exact solution of Equation (6). The rate of convergence depends on the value of the maximum modulus of the root of the characteristic equation. Given the value of the modulus of the relative error of the solution, the required number of iterations can be estimated as a local or non-local characteristic of the efficiency of finding a solution.

5. Methodology of Experiment and Results

For the simulation, a large corporate computer network with heterogeneous traffic [27] was used. The network had a mixed topology and consisted of more than 200 workstations. The diagram of the typical structure of the researched network is presented in Figure 1.

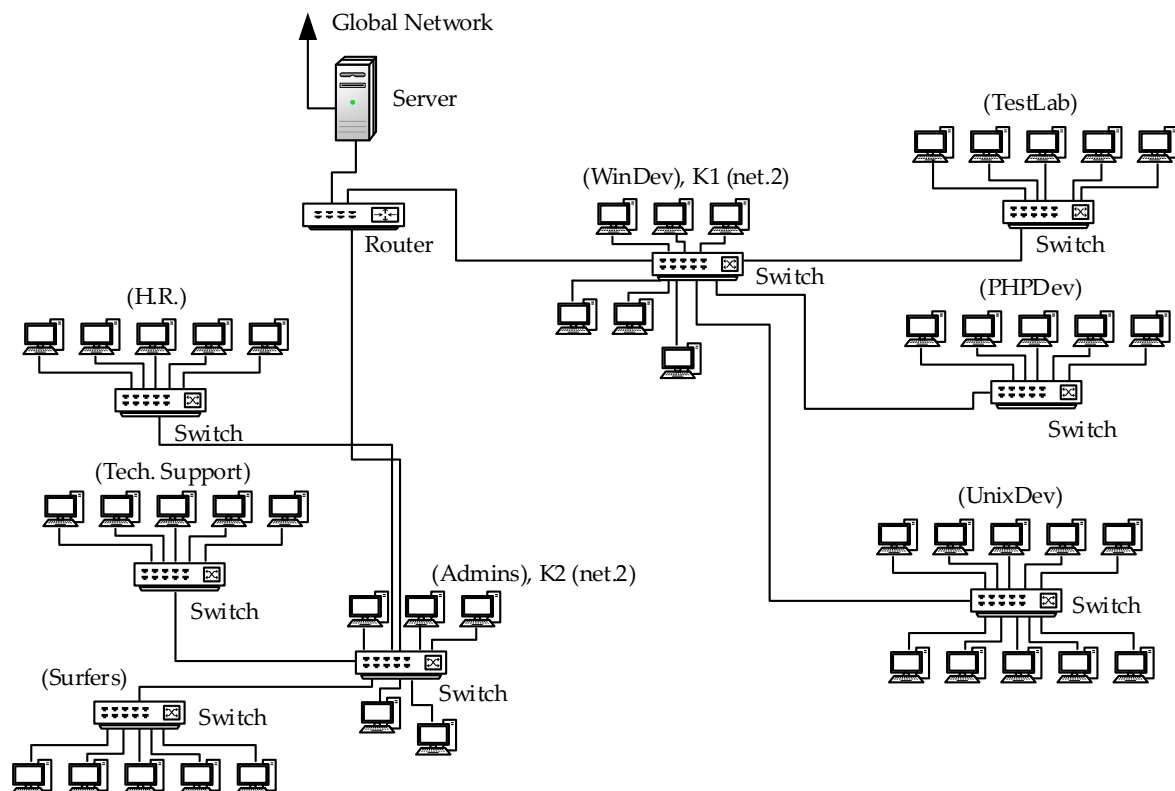


Figure 1. Typical structure of the corporate computer network.

A detailed description of all the traffic parameters is presented in [27]. Only the main ones are explained herein. The productivity and equipment parameters searched the service

output data. The time series represented the input traffic from the Internet to clients and the output traffic from clients to the Internet. The time interval was 5 min. The transformed input data (Table 1) and normalized KPIs (Table 2, featuring the main KPIs) were used.

Table 1. Traffic parameters.

Beginning of Work	Input Traffic Average Rate (bps)	Output Traffic Average Rate (bps)	Input Traffic Max. Rate (bps)	Output Traffic Max. Rate (bps)
05.11.04 12:25	1,305,583	231,727	1,583,073	246,461
05.11.04 12:20	789,773	204,819	952,414	212,975
05.11.04 12:15	574,802	184,979	582,777	194,439
05.11.04 12:10	596,443	184,760	636,903	199,803
05.11.04 12:05	618,268	190,052	636,903	199,803
05.11.04 12:00	633,440	191,657	682,937	209,495
05.11.04 11:55	934,657	203,777	1,255,028	209,495
05.11.04 11:50	964,489	196,501	1,255,028	196,501
05.11.04 11:45	581,864	184,419	594,714	196,501
05.11.04 11:40	573,807	178,597	584,367	190,760
05.11.04 11:35	548,584	173,477	584,367	190,760
05.11.04 11:30	526,192	168,952	555,655	191,190

Table 2. Normalized KPI data.

Sample Period (min)	Input Packet Loss (packet/s)	Output Packet Loss (packet/s)	Normalized Latency (Max. 0.27 s)	Normalized Throughput (Max. 0.27 s)
5	0	0	0.694	0.0102
5	0	0	0.724	0.0141
5	0.0033	0	0.810	0.3691
5	0	0.0033	0.827	0.4161
5	0.01	0	1.000	0.9065
5	0	0	0.942	1.0000
5	0	0.01	0.707	0.1209
5	0	0	0.724	0.2031
5	0.0066	0	0.738	0.2769
5	0	0	0.707	0.1230
5	0	0	0.724	0.2415
5	0	0	0.680	0.0117
5	0	0.0033	0.680	0.0112
5	0	0	0.680	0.0112
5	0	0	0.680	0.0127
5	0.0033	0.0066	0.680	0.0108
5	0	0	0.694	0.0100
5	0	0	0.724	0.0099
5	0	0	0.810	0.0100
5	0	0	0.827	0.0099

After concretizing the numerical values of the result y_i , $i = 1, 2, \dots, N$, the key performance indicators of the information and communication networks were analyzed.

The key parameters were transmission delay, bandwidth, packet loss, and security level. These parameters have the greatest impact on the resulting service quality. In [22], it was noted that the number of KPIs selected for analysis should be minimal, and in all cases, it is impractical to use more than 20 such indicators. These considerations were considered when specifying the set of KPIs.

The following were selected as the parameters of the optimized problem:

- Transmission delay τ ;
- Bandwidth C_p ;
- Packet loss L_p during data transmission;
- Level of security and data protection D_{sp} during network transmission;
- Web service quality;
- Quality of audio transmission (sound files and plain and IP-telephony);
- Speed and reliability of file exchange via FTP;
- Speed and reliability of e-mail;
- Video transmission quality.

A hypothetical IEEE 802.11 n WLAN network was considered, and the data for calculating its parameters were taken from [18]. The program of multiple correlation analysis, given in [23] and modified for this problem, was used for calculations.

Table 3 shows the partial correlation coefficients of the optimized parameters. Via these correlation coefficients, partial regression coefficients could be calculated later using Equations (1)–(3).

Table 3. Mutual correlation coefficients of optimized parameters.

Parameter	τ	C_p	L_p	D_{sp}	Web	Audio	FTP	E-mail	Video
τ	1.0	0.98	0.69	0.89	0.75	0.85	0.27	0.17	0.87
C_p	0.98	1.0	0.68	0.86	0.76	0.64	0.75	0.22	0.89
L_p	0.69	0.68	1.0	0.69	0.36	0.50	0.63	0.34	0.84
D_{sp}	0.89	0.86	0.69	1.0	0.77	0.56	0.61	0.78	0.82
Web	0.75	0.76	0.36	0.77	1.0	0.30	0.57	0.30	0.53
Audio	0.85	0.64	0.50	0.56	0.30	1.0	0.44	0.36	0.67
FTP	0.27	0.75	0.63	0.61	0.57	0.44	1.0	0.16	0.79
E-mail	0.17	0.22	0.34	0.78	0.30	0.36	0.16	1.0	0.30
Video	0.87	0.89	0.84	0.82	0.53	0.67	0.79	0.30	1.0

There was an essential correlation between the main key parameters. This is because they significantly impact sustainability requirements. The exception was e-mail, because, unlike streaming audio, video, web services, and file transfer via FTP, it is not critical for bandwidth or delivery delay. However, the level of security and data protection (parameter D_{sp}) was critical for almost all presented parameters because even for types of flexible traffic such as e-mail, data protection is an integral requirement for ensuring the quality of service (QoS).

The results of the correlation analysis are also a key indicator of the monitoring and regulation of streaming data and web services. This is necessary to ensure the secure transmission of information over the network, forecasting and preventing the congestion of the controlled network fragment. Thus, the current monitoring and management of the network survivability level, which are an integral part of the task of overall service quality management, can be successfully performed using statistical methods, particularly the correlation and regression analysis method.

In addition, the fully compiled calculation program takes up to 80 to 500 kilobytes in the memory of the computing device, depending on the scale of the network and the size of the sample. Since, currently, almost any network node is a specialized computer or even a multiprocessor system, the problem of the hardware implementation of the proposed method can be easily solved.

There are different methods for estimating μ and Σ (or, equivalently, the correlation matrices $\mathbf{R}$) when some values are missing. These methods are described in detail in [21]. Using those descriptions, some of the “best” equations that included the optimal number of independent variables were chosen. The step method (stepwise regression) was used to find the “best” regression equation.

If the number of independent variables is large, using such an approach to determine the best subset is practically not needed even when using a computer.

One solution is stepwise regression (direct), where independent variables are included one after another in a subset according to a predetermined criterion. At the same time, a variable can be replaced by another variable that is not included in the set or deleted from it. The set of criteria that determine which variables to include, replace, and delete is called a stepwise procedure. Using a stepwise procedure, the ordered list of predictors was found. For example, when $p = 5$, this list may have the form of X_2, X_5, X_1, X_4 , and X_3 . The variables were chosen so that for the definition of “the best” subset from the list of values $m \leq p$:

- (a) They could better predict $\mathbf{Y}$;
- (b) Their number t would be as small as possible.

In other words, the economy set consisted of a list of variables with the highest predictive power. In the example above, such a set could only consist of variables X_2 and X_3 if performing regression on them was almost as “qualitative” as performing regression on X_2, X_5, X_1, X_4 , and X_3 .

The procedure for determining the number t is called the stop rule. The essence of this study was to implement a systematic approach. The method of continuous network diagnostics, i.e., analyzing the impact of destabilizing factors on the overall efficiency of the network, involves dividing the process into the following interrelated stages:

1. In the first stage, diagnostics are performed at the physical level to eliminate errors and correctly interpret the results of further testing.
2. In the second stage, diagnostics of the terminal nodes of a network must be carried out by stress testing the network in two modes:
 - Calibration mode, where a network load is only applied to detect errors in the hardware and software implementation;
 - The mode where a network load is only applied to identify problems in station interactions, bottlenecks in the server, and communication channels.
3. The next step diagnoses the communication channels and servers using protocol analyzers and server analyzers. Joint processing and analysis of the speed characteristics, trends of the network traffic characteristics, and server counters obtained in the testing process are also carried out using statistical methods, which allows the causes of the malfunction of a communication channel (server) to be established and the impacts of internal and external destabilizing factors to be quantified.

6. Conclusions

This paper analyzed the system of key efficiency parameters and the features of their application for the quality management of a telecommunications network service as a complex information and communication system. It was shown that using the statistical approach made it possible to allocate dependencies between key network parameters, which allowed the service quality management system to be constructed and the most dangerous destabilizing factors to be considered. It was established that the matrices of coefficients of the normal equations for calculating the estimates of the minimum mean

square error had a structure close to a diagonally dominant matrix, which allowed the iterative search procedures for solutions to be accelerated and simplified.

Using the key performance parameters of a complex system with delays in signal and control information can provide a prediction of its status and solve the problem of real-time service quality management.

The scientific direction of future work is detailed as follows. Under the conditions of network operation with incomplete a priori certainty, it is expedient to investigate the problems of the prioritization of partial efficiency indicators, for example, using the method of hierarchy analysis, to optimize information systems with many, including contradictory, criteria. In addition, the analysis results may provide a good background for generating the parameters needed to use next-generation networks.

Author Contributions: General problem conceptualization: V.K.; methodology development: V.K., I.S. and M.K.; formal analysis of problem: I.S. and A.S.; general investigation: Y.L. and G.S.; writing—original draft preparation: V.K., Y.L. and G.S.; writing—review and editing: I.S. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. ITU. *Data Networks and Open System Communications—Open System Interconnection—Basic Reference Model: International Standard ITU-T Rec. X.200 (1994 E)*; ITU: Geneva, Switzerland, 1994; 59p.
2. Deb, A. *Control Systems Analysis and Identification*; CRC Press: Boca Raton, FL, USA, 2018; 364p.
3. Liu, Q.; Wang, Z.; He, X. *Stochastic Control and Filtering over Constrained Communication Networks*; Springer Nature Switzerland AG: Cham, Switzerland, 2019; 222p.
4. Dodonov, A.G.; Lande, D.V. *The Survivability of Information Systems*; Nauk: Dumka, India, 2011; 256p.
5. Diaz, O.; Xu, F.; Min-Allah, N.; Khodeir, M.; Peng, M.; Khan, S.; Ghani, N. Network Survivability for Multiple Probabilistic Failures. *IEEE Commun. Lett.* **2014**, *16*, 1320–1323. [\[CrossRef\]](#)
6. Csermely, P. *Weak Links the Universal Key to the Stability of Networks and Complex Systems*; Springer: Berlin/Heidelberg, Germany, 2009; 404p.
7. Dreyfus, S.E. *Some Types of Optimal Control of Stochastic Systems*; Rand Corp: Santa Monica, CA, USA, 1963; 29p.
8. Hellmann, F.; Schultz, P.; Grabow, C.; Heitzig, J.; Kurths, J. Survivability of Deterministic Dynamical Systems. [Electron Resource]. 2016. 12p. Available online: <https://www.nature.com/articles/srep29654.pdf> (accessed on 1 November 2023).
9. Karpinski, M. *Information Security*; Stowarzyszenie Inżynierów i Techników Mechaników Polskich: Warsaw, Poland, 2012; 280p. (In Polish)
10. Xie, L.; Heegaard, P.; Jiang, Y. Modeling and Quantifying the Survivability of Telecommunication Network Systems under Fault Propagation. In *EUNICE 2013 Advances in Communication Networking*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 25–36.
11. Agrawal, A.; Bhatia, V.; Prakash, S. Network and Risk Modeling for Disaster Survivability Analysis of Backbone Optical Communication Networks. *J. Light. Technol.* **2019**, *37*, 2352–2362. [\[CrossRef\]](#)
12. Knyazeva, N.A. A method for ensuring the structural survivability of a telecommunications network. *Int. J. Inf. Models Anal.* **2014**, *8*, 152–166.
13. Knyazeva, N.A.; Zimenko, L.N. Method of ensuring the structural survivability of the intellectual superstructure. *Inf. Control Syst. Railw. Transp.* **2016**, *6*, 23–29.
14. Otrokh, S.I.; Korshun, N.V.; Yarosh, V.A. Methodology for calculating the indicators of the survivability of modern telecommunication network channels. *Lucky Hijacker Inf.* **2016**, *4*, 52–57.
15. Tolubko, V.B.; Berkman, L.N.; Otrokh, S.I.; Yarosh, V.O. Methods for assessing the stability of telecommunication in the minds of the most unpredictable destabilizing factors. *Connectivity* **2016**, *5*, 3–7.
16. Zhang, R.; Cai, L.; Pan, J. *Resource Management for Multimedia Services in High Data Rate Wireless Networks*; Springer International Publishing AG: Cham, Switzerland, 2017; 140p.
17. Recommendation ITU-T L.392 (04/2016). Disaster Management for Improving Network Resilience and Recovery with Movable and Deployable Information and Communication Technology (ICT) Resource Units. [Electron Resource]. Available online: <http://handle.itu.int/> (accessed on 1 November 2023).

18. Vinogradov, N.A.; Lesnaya, N.N.; Savchenko, A.S. Investigation of the characteristics of useful throughput under changing load conditions. *Probl. Informatisation Manag.* **2009**, *4*, 28–31.
19. Vinogradov, N.A. Analysis of potential characteristics of devices for switching and network management of new generations. *Connectivity* **2004**, *4*, 10–17.
20. Bendat, J.; Piersol, A. *Engineering Applications of Correlation and Spectral Analysis*, 2nd ed.; Wiley-Interscience: Hoboken, NJ, USA, 2013; 472p.
21. Afifi, A.; Azen, S.P. *Statistical Analysis: A Computer Oriented Approach*, 2nd ed.; Academic Press: Cambridge, MA, USA, 1979; 442p.
22. Afifi, A.; May, S.; Donatello, R.; Clark, V.A. *Practical Multivariate Analysis*, 6th ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2020; 434p.
23. Zablotskyi, V.; Selepyna, Y.; Lyshuk, V.; Yakymchuk, N.; Tkachuk, A. Method for evaluation quality parameters of telecommunications services. *IT Autom. Meas. Econ. Environ. Prot.* **2022**, *12*, 30–33. [[CrossRef](#)]
24. Kreher, R. *UMTS Performance Measurement: A Practical Guide to KPIs for the UTRAN Environment*; John Wiley & Sons Ltd.: Chichester, UK; Hoboken, NJ, USA, 2006; 227p.
25. Nguyen, M. A Guide on Data Analysis. Bookdown. [Electron Resource]. 2020. Available online: https://bookdown.org/mike/data_analysis/ (accessed on 1 November 2023).
26. Mosteller, F.; Tukey, J. *Data Analysis and Regression: A Second Course in Statistics*, 1st ed.; Pearson: London, UK, 1977; 588p.
27. Savchenko, A.S. The Choice of Parameters of Switching Equipment on the Background of Analysis of Computer Networks Traffic. Ph.D. Thesis, National Aviation University, Kyiv, Ukraine, 2007; 144p.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.