UDC 621. 93

THE TASK OF THE TELECOMMUNICATION FLOW CONTROL SOLUTION BASED ON THE CHANNEL UTILIZATION MODEL

Oleksandr I. Romanov, Mykola. M. Nesterenro

Telecommunication Department Institute of Telecommunication Systems National Technical University of Ukraine "KPI" Kyiv, Ukraine

Victor I. Tikhonov

Telecommunication Networks Department A.S. Popov Odessa National Academy of Telecommunications Odessa, Ukraine

Abstract — The channel utilization factor is a typical QoS functional model of a telecommunication network. A common analytic formula of this factor includes non differentiable members that hinder functional model conversion. This paper introduces an approximation of the QoS functional model based on regression analysis and performed via differentiable functions. Recommendations are given to implement the findings for data flow control in telecommunications.
Keywords — channel utilization; approximation; regression analysis.

Introduction

An analytic functional presentation of a telecommunication network commonly results in rather complex math expressions. Therefore, a relevant practical solution is merely obtainable in respect to the telecommunication network control, design and quality of service estimation. For example, when searching permissible values of QoS on the branches of the network under given constraints on the quality of service, it is necessary to solve a system of $N(N-1)$ inequalities (where N - number of nodes in the network). The figures of service quality in the branches of the network are inequalities variables. Thus, any given branch is multiply used in flow dynamic routing for diverse data transmission directions.

Therefore, any change in the quality of service in one branch provokes a change of quality of service in a number of other communication directions. Moreover, any given network branch operates alternatively under different terms, e.g. requires more stringent QoS demands for a multiple hop path. Therefore, an accurate inequalities system solution based on the multi-direction QoS demands is practically impossible. To overcome this issue the channel utilization factor is treated in [1]. For instance, while forming a loading schedule under given inequalities system, the optimal data forwarding path is not unique. In this case, a recommended path presumes an increase of a total QoS factor. Ultimately, this allows minimizing the required number of channels and, as a consequence, the total expenses for network equipment. Though, a common analytic formula of this factor includes non differentiable members that hinder functional model conversion. Because of that, approximate analytic models are expedient to solve the spoken problem.

Statement of the Problem

The purpose of the work is to obtain an approximate differentiable analytical function of channel utilization dependent on channel number/load and packet loss probability which is relevant to the network data transfer processes.

The channel utilization factor (CUF) is one of the most commonly used indicators for evaluating the performance of telecommunications networks. This is due to its numerical value along with clear physical and cost meaning towards a network. For a circuit switched network branch the CUF can be represented as

$$K = \frac{Y(p)}{V}\Big|_{p\ =\ const},$$

where K is the channel utilization factor; $Y(p)$ is branch capacity; V is the number of channels; p is probability of connection reject. In case of the simplest jobs stream, the branch capacity is determined by the following expression:

$$Y(p) = Z(1-p) = Z\Big(1-\frac{Z^V/V!}{\sum_{i=0}^{V} Z^i/i!}\Big). \tag{1}$$

After simple transformations, the CUF formula evolves to

$$K = \frac{Z(\sum_{i=0}^{V} Z^i/i! - Z^V/V!)}{V\sum_{i=0}^{V} Z^i/i!}. \quad (2)$$

While solving network problems it is necessary to get an optimal correspondence between the number of channels, load and losses on the branches of the network, as well to ensure given QoS requirements in various link directions [2]. To do this, the following relationships are to be considered:

K = *f*(*V*), if *p* = *const*;

K = f(Z), if p = const;

K = f(p), if p = const etc.

The analysis of the CUF is much easier and more accurate if analytical expressions are differentiable in the entire range of values, and the dependence of one parameter from another can be represented explicitly. Equation (2) does not meet these requirements and its use in practical calculations has considerable difficulties.

It is therefore proposed an approximation of equation (2) using empirical differentiable functions. For this purpose it is proposed to use the mathematical apparatus of numerical analysis [3], which provides a high degree of adequacy of the results.

Task Solution

To construct an empirical formula two steps are to be made:

– determine general formula;

– find appropriate parameters.

For a given set of values $K_i,\ V_i)|_{p=const}$, $(K_i,\ Z_i)|_{p=const}$, and $(K_i,\ p_i)|_{V=const}$ an analytic function was taken:

$$y = \tilde{f}(x;\, a, b, c)\ , \quad (3)$$

where $a,\ b$ and c – constant coefficients (parameters). The analysis indicates the following power function is the most accurate [3]:

$$K = ax^b + c\ . \quad (4)$$

This result $k - c = ax^b$. When taking the logarithm we obtain: $g\left|k-c\right| = lg\left|a\right| + blg\ x$, $\left|k-c\right| \neq 0$, $\left|a\right| \neq 0$. Consider $lg\ |k-c| = K$ and $\lg x = X$ we get linear function

$$K = bX + \lg a\ . \quad (5)$$

Searching the formula (4) parameters we begin from c value. For this take geometric mean $x_s = \sqrt{x_1 x_n}$, where x_1 and x_n are marginal values. Then apply linear interpolation for x_s to find:

$$k_s = k_i + \frac{k_{i+1} - k_i}{x_{i+1} - x_i}(\bar{x}_s - x_i)\ ,$$

where x_i and x_{i+1} are intermediate values determining x_s $(x_i < \bar{x}_s < x_{i+1})$. Assume points $M_1(x_1,\ k_1)$, $M_s(x_s,\ k_s)$, $M_n(x_n,\ k_n)$ belong to curve (4), we result three inequalities:

$$k_1 = c + ax_1^b\ ,$$
$$k_s = c + ax_s^b\ ,$$
$$k_n = c + ax_n^b\ .$$

Raising the power $\bar{x}_s = \sqrt{x_1 x_n}$ and multiplying by a, we obtain $ax_s^b = \sqrt{ax_1^b\ ax_n^b}$, or

$$k_s - c = \sqrt{(k_1 - c)(k_n - c)}\ .$$

Solving the last equation in respect to c, we find

$$c = \frac{k_1 k_n - k_s^2}{k_1 + k_n - 2k_s}\ .$$

After c is defined we calculate points $N_i = (K_i, X_i)$, where $X_i = \lg x_i$, $K_i = lg\ (k_i - c)$, $(i = 1,\ 2,\ ....,\ n)$. As these points are located almost in a straight line, it confirms the correctness of the choice of (4).

Determining the best parameters $a,\ b$ in approximation function (4) is performed by the least squares method. The choice of this method stems from the fact that its use yields the smallest deviation from the baseline data by comparison with medium. Furthermore, it has another important advantage, namely if the sum S of squares of deviations is small, then these deviations themselves are also small in absolute value.

According to the least squares method (LSM) best coefficients are those for which the sum of squared deviations is minimal:

$$S(a_1,\ a_2,\ ...,\ a_m) = \\ = \sum_{i=1}^{n} [\tilde{K}(x_i,\ a_1,\ a_2,\ ...,\ a_m) - k_i)]^2 \to \min. \quad (6)$$

Hence, using the necessary conditions for a function extremum of several variables, we obtain the so-called normal system:

$$\frac{\partial}{\partial\ x_i} S(a_1,\ a_2,\ ...,\ a_m) = 0,\ i = 1,\ 2,\ ...,\ m \quad (7)$$

to find the parameters $a_1,\ a_2,\ ...,\ a_m$. The system is simplified if the empirical function is linear towards the parameters. Finally, a normal system would be:

$$\begin{cases} {}_0 n + a_1[x] + a_2[x^2] + ... + a_m[x^m] = [k], \\ a_0[x] + a_1[x^2] + ... + a_m[x^{m+1}] = [xk], \\ ...., \\ a_0[x^m] + a_1[x^{m+1}] + ... + a_m[x^{2m}] = [x^m k]. \end{cases} \quad (8)$$

As after taking logarithms, empirical formula (4) has a linear form (5), the system of equations can be written as:

$$\begin{cases} n \lg a + [X]b = [K]; \\ [X]\lg a + [X^2]b = [XK] \end{cases}, \quad (9)$$

where n is the number of given points; $[X] = \sum_{i=1}^{n} \lg x_i$; $[X^2] = \sum_{i=1}^{n} (\lg x_i)^2$; $[K] = \sum_{i=1}^{n} \lg|k_i - c|$; $[XK] = \sum_{i=1}^{n} (\lg x_i) \times (\lg|k_i - 1|), \ (i = 0, 1, 2, ..., n)$. Next we solve the system of linear normal equations (9) with a square symmetric matrix of coefficients on Cramer rule with determinants:

$$\begin{cases} \alpha \times \Delta = \Delta_\alpha \\ \beta \times \Delta = \Delta_\beta \end{cases}, \quad (10)$$

where $\Delta = \begin{vmatrix} n & [X] \\ [X] & [X^2] \end{vmatrix}$ is determinant of (9); $\Delta_\alpha = \begin{vmatrix} [Y] & [X] \\ [XY] & [X^2] \end{vmatrix}$; $\Delta_\beta = \begin{vmatrix} n & [K] \\ [X] & [XK] \end{vmatrix}$.

If $\Delta \neq 0$ the system (9) has a unique solution:

$$\alpha = \frac{\Delta_\alpha}{\Delta}; \quad \beta = \frac{\Delta_\beta}{\Delta}, \quad (11)$$

where $\alpha = \lg a$, $\beta = b$ are desired constants of the empirical analytical formula (4). The final chose of an empirical formula must satisfy the following relation:

$$\begin{cases} a \times (-1), \ \textit{если} \ c > k_i, \ |k_i - c| \neq 0; \\ a \times (+1), \ \textit{если} \ k_i > c, \ |k_i - c| \neq 0. \end{cases} \quad (12)$$

Based on the spoken approach the principal empirical relations for the channel utilization factor obtained towards load value, packet loss and the number of channels in the branches of a telecommunication network: $K = f(V_i)$ if $p = const$; $K = f(Z_i)$ if $p = const$; $K = f(p_i)$ if $V = const$; $K = f(p_i)$ if $Z = const$.

Next is a generalization of the analytical model obtained by smoothing of the statistical data. Assessment of the adequacy of the results can be made using the mathematical apparatus of the linear regression analysis. To carry out a statistical analysis of linear model the least squares method is applied. Let ε_i observation random error. Yet, we present the results as follows:

$$y_i = \beta_0 a_0(x_i) + \beta_1 a_1(x_i) + ... + \beta_{k-1} a_{k-1}(x_i) + \varepsilon_i,$$

Assume ε_1 not correlated and have zero expectations, i.e.

$$M[\varepsilon_i] = 0,$$

$$k_{ij} = \begin{cases} \sigma_\varepsilon^2, \ i = j, \\ 0, \ i \neq j, \end{cases}$$

where k_{ij} is covariation of random values ε_i and ε_j, $i, j = 1, 2, ..., n.$. It should be noted that LSM-estimations do not depend on the of the sample size (assuming $n \geq k$, where k is the number of estimated parameters). If the errors ε_i, $i, j = 1, 2, ..., n$ have normal distribution $N\ (0, \sigma_\varepsilon^2)$ and are not correlated, then LSM-estimations coincide with those obtained due to method of maximum likelihood.

Let $\tilde{\beta}_0, \tilde{\beta}_1, ..., \tilde{\beta}_{k-1}$ are estimations of linear model parameters. Let calculate the rest of square sum Q_e :

$$Q_e\ (\tilde{\beta}_0, \ \tilde{\beta}_1, \ ....\tilde{\beta}_{k-1}) =$$
$$\sum [y_i - \tilde{\beta}_0 a_0(x_i) - \tilde{\beta}_1 a_1(x_i) - ..... - \tilde{\beta}_{k-1} a_{k-1}(x_i)]^2,$$

or in matrix form:

$$Q_e = (Y - A\tilde{\beta})^T (Y - A\tilde{\beta}) = Y^T Y - \tilde{\beta}^T A^T Y \ .$$

In our case is more convenient to use:

$$|Q_e| = |Y^T Y - \tilde{\beta}^T A^T Y| \ . \quad (13)$$

The non-shifted estimation of error dispersion is:

$$\tilde{\sigma}_\varepsilon^2 = s_\varepsilon^2 = \frac{Q_e}{n-k} \ . \quad (14)$$

Herewith, the covariation matrix estimation is:

$$K_m = s_\varepsilon^2 (A^T A)^{-1}. \quad (15)$$

In the case where the observation errors are not correlated and have a normal distribution, the parameter estimations for analytical model β_j, $j = 0, 1, ..., k-1$, also have a normal distribution. Boundaries of confidence intervals in this case are defined as follows:

$$\tilde{\beta}_j \pm t_{1-\frac{\alpha}{2}}(n-k)s_{\varepsilon}\sqrt{a_{jj}} \quad , \tag{16}$$

where a_{jj} are the elements of matrix $(A^T A)^{-1}$, α is given significance level. Next, we find confidence intervals for the parameters of obtained analytical dependences. For practical calculations it is enough to take the level of significance $\alpha = 0,05$ with the use of t-distribution (or Student distribution). To solve the QoS analysis problems for a telecommunication network functioning, namely to estimate the channel bandwidth, is of interest the CUF change rate that is the first derivative of the found empirical function:

$$\tilde{K}' = abx^{b-1} \quad . \tag{17}$$

Using the dependence of the form (17) with the required values of the variables were obtained dependency graphs $\tilde{K}' = f(V_i)\big|_{p=const}$, $\tilde{K}' = f(Z_i)\big|_{p=const}$, $\tilde{K}' = f(p_i)\big|_{V=const}$, $\tilde{K}' = f(p_i)\big|_{Z=const}$ shown in Fig. 1 – Fig.4.

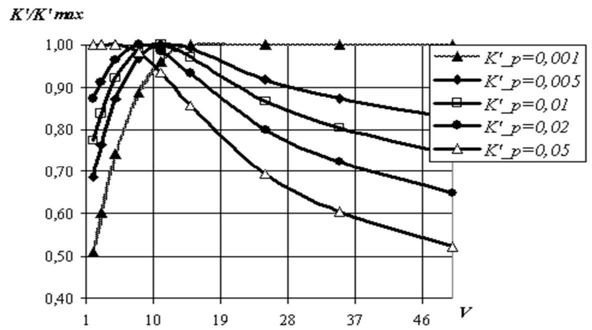

Fig. 1. Dependency graphs $\tilde{K}'/\tilde{K}'_{\max}$ of V with p = const

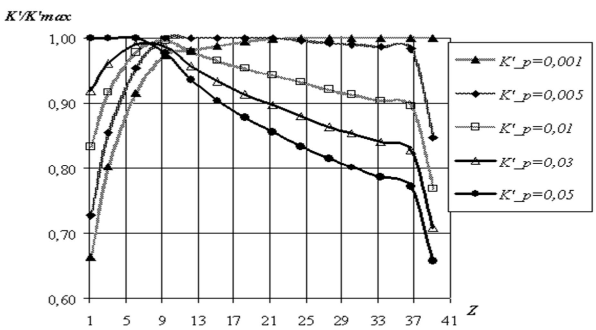

Fig. 2. Dependency graphs $\tilde{K}'/\tilde{K}'_{\max}$ of Z with p = const

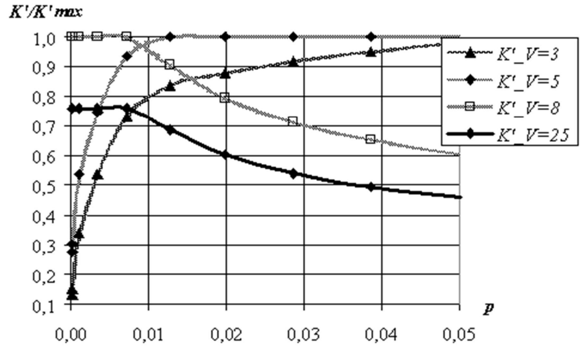

Fig. 3. Dependency graphs $\tilde{K}'/\tilde{K}'_{\max}$ of p with V = const

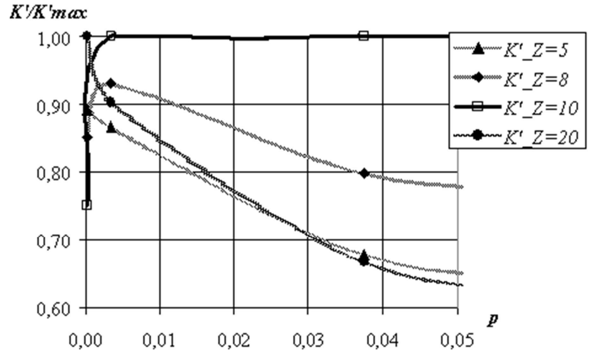

Fig. 4. Dependency graphs $\tilde{K}'/\tilde{K}'_{\max}$ of p with V = const

Conclusion

The primary CUF increase is observed in pure loaded network branches with a small number of channels. Medium and highly loaded branches indicate approximately even CUF change. Therefore, the network data flow provision should be done under following recommendations:

− in case of pure loaded network an even data flow distribution is reasonable;

− if the network load intensity in the direction of communication is medium or high, it is necessary to strive for equality losses in the branches network.

This approach will ensure the high efficiency of the channel utilization, and as a result, mitigates the network equipment requirements to provide QoS demands in various data transfer directions.

1. A.I. Romanov, “Telekommunikatsionnye seti i upravleniye, VPC ” Kiev University, 2003, 247p.

2. A.I. Romanov, M.M Nesterenko, “Modyfikovana metodyka otsinky yakosti obslugovuvannia v telekomunikatsiinykh merezhakh z vstanovlenniam ziednannia,” Scientific Papers of “NTUU “KPI”, vol. 3, 2007, pp. 121-128.

3. B.P. Demidovich, I.A. Maron, E.Z. Shuvalova, “Chislennye metody analiza” Moscow, 1963, 400p.

Received in final form April 27, 2014