NEWS
OF THENATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X **https://doi.org/10.32014/2020.2518-1726.95**
Volume 6, Number 334 (2020), 36 – 44

UDC 004.9: 519.2
IRSTI 27-43-51

M.Yu. Babich[1], M.M. Butaev[1], A.A. Tarasov[1], A.I. Ivanov[2], A.Yu. Malygin[2], K.T. Sauanova[3], Sh.N. Sagyndykova[3]

[1]JSC "Scientific-Production Enterprise “Rubyn”", Penza, Russia;
[2]Education Center "Information Security Systems and Technologies", Penza State University", Penza, Russia;
[3]Almaty University of Power Energy and Telecommunication, Almaty, Kazakhstan.
E-mail: mail @npp-rubin.ru, mal@yandex.ru, sauanovaklara@gmail.com, tomka2001@mail.ru

FUNCTIONS OF DISTRIBUTIONS OF AMOUNTS OF UNIFORALLY DISTRIBUTED RANDOM VALUES OF TIMES OF PROCESSING THE REQUEST OF THE INFOCOMMUNICATION SYSTEM

Abstract. The normal distribution of a random variable is usually used in studies of the probabilistic properties of information systems. Using the normal distribution to approximate the distributions determined over a bounded distorts the physical meaning of the model and the numerical results obtained can only be used as an initial approximation. The purpose of the work is to improve methods for calculating the probability properties of infocommunication systems. The object of study is an analytical method for calculating the request processing time in the system, the subject is the formula for calculating the duration of sequential processing of a request by elements of the system with uniformly distributed independent random processing times. For positive random variables, it is proposed to use finite-interval distribution laws, for example, beta distribution. Density formulas and probability functions for the sums of two, three, and four independent randomly distributed variables are given.

Key words: cumulative distribution function, CDF, probability density function, PDF, probabilistic properties, infocommunication system.

Introduction. In a multicriteria performance evaluation, an additive convolution of its criteria is often used; the weighting coefficients of the criteria are introduced [1]. The meaning of these weights and their values are determined heuristically, based on an expert survey. In contrast to additive convolution, a multiplicative formula is known for converting criteria efficiency functions into a general indicator with the goal of its further optimization [2]. In this case, the output of the complex criteria be more grounded and it is easier to understand. One of the approaches to deriving a multiplicative complex efficiency criterion is based on a probabilistic approach: the probabilities of reaching given criteria values are estimated. The overall probability of achieving the target values of several criteria is calculated as the product of the probabilities of the individual criteria, if criteria do not depend on each other. An increase in the probability of an individual criterion increases the likelihood of a complex criterion, therefore, increases the efficiency. At the same time, for an analytical study using a simple probabilistic multiplicative complex efficiency criterion, an analytical description of the probability distribution functions of the criteria used, which are often difficult to obtain, is necessary.

One of the important indicators of the effectiveness of infocommunication systems is the responsiveness, reactivity, throughput and other properties associated with the evaluation of the system’s performance over time [3]. A probabilistic criterion, for example, the duration of data processing, is defined as the probability of completion of data processing in a given time.

Another problem to be solved at the initial stages of the design of infocommunication systems is to assess the impact of the durations of individual stages of data processing on the total processing time in the system [4, 5]. This problem is solved by the most informative calculation of probability properties durations her data processing steps.

Mathematically, the tasks of calculating the data processing time are usually solved using an acyclic directed stochastic graph. The same apparatus is used to calculate network diagrams using the PERT method [6 - 8]. Computational methods PERT designed for networks of large dimension, so the use of the central limit theorem of probability theory, justifying the use of normal probability distribution of time the law of individual stages and network as a whole [9]. Such an approach, with a small number of data processing stages in the network, causes difficulties in substantiating the use of the normal distribution law defined on the interval ($-\infty$, $+\infty$). In practice a priori known minimum duration of data processing in step, it is caused by physical features of the data processing in the step, and the maximum length is determined by the requirements of the technical tasks. The mean or mode of the duration of the stage is unlikely to occur, therefore, calculation based on them has little reliability. The calculation of only the largest and smallest data processing times (interval method) is not sufficiently informative [10]. Under such conditions, it is advisable to use a probabilistic approach with a uniform density of data processing durations at the stages [11, 12].

With the sequential execution of the stages, the duration of processing requests is reduced to summing the durations of processing at each stage, i.e. to the summation of independent random variables [13 - 15]. The probability distribution function of the sum of independent random variables is calculated through the convolution formula of the distribution functions of the terms.

The topology of the connections of information processing elements in the system can be different, but to evaluate the critical duration of information processing, a sequential chain of elements is allocated that is likely to determine the temporal properties of the system. Estimating the duration of information processing in such a chain of elements, obviously, comes down to summing the processing durations in each of them. For a deeper estimation of temporal properties, it is necessary to know the probability density function (PDF) $f(t)$ and/or the cumulative distribution function (CDF) $F(t)$ of a continuous random variable [14], and the best analytical solution to this problem is an accurate analytical description of the PDF and CDF.

Mathematical model. Let random continuous independent quantities $T \in t_i$, $i \in 1,..., n$ be uniformly distributed on the intervals $t_i \in [0; \beta_i]$, $\beta_i > 0$. To simplify further analytical transformations, random variables are ordered by increasing β_i. The upper bounds of the ordered random variables are indicated by the variables b_i. Ordering condition: $b_i \le b_{i+1}$, $i \in 1,..., n - 1$. For example, the largest values of the random component of the duration of the first stage $\beta_1 = 3$, the second stage $\beta_2 = 5$ and the third stage $\beta_3 = 4$ will receive new notation $b_1 = 3$, $b_2 = 4$, $b_3 = 5$. It is known that, due to the additive properties of the arithmetic operation of summation, changing the sequence of summing random variables does not affect the final result [13].

In contrast to [15], the simplified Heaviside function $H(t)$ [17] was used to simplify analytical dependences:

$$H(t) = \begin{cases} 0, \ t < 0; \\ \text{undefined}, \ t = 0; \\ 1, \ 0 < t. \end{cases} \tag{1}$$

The PDF and CDF of a random variable using (1) are described:

$$f_1(t) = \begin{cases} 0, \ b_1 < t < 0; \\ \dfrac{1}{b_1} H(b_1 - t), \ 0 < t < b_1; \end{cases} \tag{2}$$

$$F_1(t) = \begin{cases} 0, \ t \le 0; \\ \dfrac{t}{b_1}, \ 0 < t < b_1; \\ 1, \ b_1 \le t. \end{cases} \tag{3}$$

Based on these proposals, the formula expressions $F(t)$ and $f(t)$ are obtained for the sums of several independent random variables for a more general case than [13].

The PDF of the sum of two random variables is determined by convolution [15]:

$$f_{12}(t)=\int_0^t f_1(u)\,f_2(t-u)\,du, \tag{4}$$

where $f_1(t)$, $f_2(t)$ are the PDF of the first and second terms; $f_{12}(t)$ is the PDF of the sum of the first and second random variables.

The CDF $F_{12}(t)$ is defined similarly. Using the direct method, $f_{12}(t)$ and $F_{12}(t)$ can be determined by substituting (2), (3) in (4):

$$f_{12}(t)=\begin{cases}0,\ t\le 0;\\ \dfrac{1}{b_1b_2}\left[t-H(t-b_1)(t-b_1)-H(t-b_2)(t-b_2)\right],0<t<b_1+b_2;\\ 0,\ b_1+b_2\le t;\end{cases} \tag{5}$$

$$F_{12}(t)=\begin{cases}0,\ t\le 0;\\ \dfrac{1}{2b_1b_2}\left[t^2-H(t-b_1)(t-b_1)^2-H(t-b_2)(t-b_2)^2\right],0<t<b_1+b_2;\\ 1,\ b_1+b_2\le t.\end{cases} \tag{6}$$

The characteristic function of the sum of two uniformly distributed random variables [17]:

$$\chi_t(p)=\frac{\left(1-e^{-b_1p}\right)\left(1-e^{-b_2p}\right)}{b_1b_2p^2}.$$

The uniformly distributed random variables and their sums are symmetric, therefore, the mean, median and mode are equal, and the asymmetry coefficient is zero. For the sum of two uniformly distributed random variables, the mode is absent [15]:

$$M(T)=Me(T)=\frac{b_1+b_2}{2}.$$

Dissipation properties:

- variance:

$$D(T)=\frac{(b_1^2+b_2^2)}{12};$$

- quantile of the order of 0.2:

$$\sqrt{\frac{2}{5}b_1b_2}\ ;$$

- quantile of the order of 0.8:

$$2\sqrt{\frac{2}{5}b_1b_2}\ ;$$

- 2nd order starting moment :

$$m_2(T)=\frac{b_1^2+b_2^2}{3}+\frac{b_1b_2}{2};$$

- 3rd order starting moment :

$$m_3(T)=\frac{b_1^3+b_2^3}{4}+\frac{b_1b_2(b_1+b_2)}{2};$$

- 4th order starting moment :

$$m_4(T) = \frac{b_1^4 + b_2^4}{5} + \frac{b_1 b_2 \left(b_1^2 + b_2^2\right)}{2} + \frac{2}{3} b_1^2 b_2^2 \ ;$$

- kurtosis coefficient :

$$Ex(T) = \frac{9\left(b_1^4 + b_2^4\right) + 30 b_1^2 b_2^2}{5\left(b_1^2 + b_2^2\right)^2} .$$

Below, without further explanation, formulas for calculating the properties for the sums of three and four uniformly distributed random variables are given. The quantiles for such distributions are usually determined by numerical methods.

The PDF, CDF, characteristic function, mean and variance of the sum of three uniformly distributed random variables

$$f_{123}(t) = \begin{cases} 0, t \le 0; \\ \dfrac{1}{2b_1 b_2 b_3}\Big[t^2 - H(t-b_1)(t-b_1)^2 - H(t-b_2)(t-b_2)^2 - H(t-b_3)(t-b_3)^2 + \\ \quad + H(t-b_1-b_2)(t-b_1-b_2)^2 + H(t-b_1-b_3)(t-b_1-b_3)^2 + \\ \quad + H(t-b_2-b_3)(t-b_2-b_3)^2\Big], 0 < t < b_1 + b_2 + b_3 \ ; \\ 0, b_1 + b_2 + b_3 \le t; \end{cases}$$

$$F_{123}(t) = \begin{cases} 0, t \le 0; \\ \dfrac{1}{6b_1 b_2 b_3}\Big[t^3 - H(t-b_1)(t-b_1)^3 - H(t-b_2)(t-b_2)^3 - H(t-b_3)(t-b_3)^3 + \\ \quad + H(t-b_1-b_2)(t-b_1-b_2)^3 + H(t-b_1-b_3)(t-b_1-b_3)^3 + \\ \quad + H(t-b_2-b_3)(t-b_2-b_3)^3\Big], 0 < t < b_1 + b_2 + b_3 \ ; \\ 1, b_1 + b_2 + b_3 \le t \ ; \end{cases}$$

$$\chi_t(p) = \frac{\left(1 - e^{-b_1 p}\right)\left(1 - e^{-b_2 p}\right)\left(1 - e^{-b_3 p}\right)}{b_1 b_2 b_3 p^3};$$

$$M(T) = Me(T) = Mo(T) = \frac{b_1 + b_2 + b_3}{2};$$

$$D(T) = \frac{b_1^2 + b_2^2 + b_3^2}{12} .$$

A γ-quantile is calculated by numerically solving the equation $F_{123}(t_\gamma) = \gamma$. For example, a 0.2-quantile $t_{0.2} = 1.55675$; 0.8-quantile $t_{0.8} = 2.94325 = 4.5 - 1.55675$ with $b_1 = 1$, $b_2 = 1.5$, $b_3 = 2$. Since $F_{123}(t)$ is a polynomial interval-defined function, before solving the equation, it is necessary to determine the interval of values t_γ in which there will be a solution to the equation. Various methods can be used to determine the interval, for example, using a graph, or by tabulating $F_{123}(t)$ at the boundaries of the intervals, etc.

$$m_2(T) = \frac{b_1^2 + b_2^2 + b_3^2}{3} + \frac{b_1 b_2 + b_1 b_3 + b_2 b_3}{2};$$

$$m_3(T) = \frac{b_1^3 + b_2^3 + b_3^3}{4} + \frac{b_1 b_2 (b_1 + b_2) + b_1 b_3 (b_1 + b_3) + b_2 b_3 (b_2 + b_3)}{2} + \frac{3}{4} b_1 b_2 b_3 ;$$

$$m_4(T) = \frac{b_1^4 + b_2^4 + b_3^4}{5} + \frac{b_1 b_2 (b_1^2 + b_2^2) + b_1 b_3 (b_1^2 + b_3^2) + b_2 b_3 (b_2^2 + b_3^2)}{2} + \\ + \frac{2 (b_1^2 b_2^2 + b_1^2 b_3^2 + b_2^2 b_3^2)}{3} + b_1 b_2 b_3 (b_1 + b_2 + b_3);$$

$$Ex(T) = \frac{9 (b_1^4 + b_2^4 + b_3^4) + 30 (b_1^2 b_2^2 + b_1^2 b_3^2 + b_2^2 b_3^2)}{5 (b_1^2 + b_2^2 + b_3^2)^2}.$$

The PDF, CDF, characteristic function, mean and variance of the sum of four uniformly distributed independent random variables:

$$f_{1234}(t) = \begin{cases} 0, t \le 0; \\ \frac{1}{6 b_1 b_2 b_3 b_4} \Big[t^3 - H(t - b_1)(t - b_1)^3 - H(t - b_2)(t - b_2)^3 - H(t - b_3)(t - b_3)^3 - \\ \quad - H(t - b_4)(t - b_4)^3 + H(t - b_1 - b_2)(t - b_1 - b_2)^3 + H(t - b_1 - b_3)(t - b_1 - b_3)^3 + \\ \quad + H(t - b_1 - b_4)(t - b_1 - b_4)^3 + H(t - b_2 - b_3)(t - b_2 - b_3)^3 + \\ \quad + H(t - b_2 - b_4)(t - b_2 - b_4)^3 + H(t - b_3 - b_4)(t - b_3 - b_4)^3 - \\ \quad - H(t - b_1 - b_2 - b_3)(t - b_1 - b_2 - b_3)^3 - H(t - b_1 - b_2 - b_4)(t - b_1 - b_2 - b_4)^3 - \\ \quad - H(t - b_1 - b_3 - b_4)(t - b_1 - b_3 - b_4)^3 - H(t - b_2 - b_3 - b_4)(t - b_2 - b_3 - b_4)^3 \Big], \\ \quad 0 < t < b_1 + b_2 + b_3 + b_4 ; \\ 0, b_1 + b_2 + b_3 + b_4 \le t ; \end{cases}$$

$$F_{1234}(t) = \begin{cases} 0, t \le 0; \\ \frac{1}{24 b_1 b_2 b_3 b_4} \Big[t^4 - H(t - b_1)(t - b_1)^4 - H(t - b_2)(t - b_2)^4 - H(t - b_3)(t - b_3)^4 - \\ \quad - H(t - b_4)(t - b_4)^4 + H(t - b_1 - b_2)(t - b_1 - b_2)^4 + \\ \quad + H(t - b_1 - b_3)(t - b_1 - b_3)^4 + H(t - b_1 - b_4)(t - b_1 - b_4)^4 + \\ \quad + H(t - b_2 - b_3)(t - b_2 - b_3)^4 + H(t - b_2 - b_4)(t - b_2 - b_4)^4 + \\ \quad + H(t - b_3 - b_4)(t - b_3 - b_4)^4 - H(t - b_1 - b_2 - b_3)(t - b_1 - b_2 - b_3)^4 - \\ \quad - H(t - b_1 - b_2 - b_4)(t - b_1 - b_2 - b_4)^4 - H(t - b_1 - b_3 - b_4)(t - b_1 - b_3 - b_4)^4 - \\ \quad - H(t - b_2 - b_3 - b_4)(t - b_2 - b_3 - b_4)^4 \Big], 0 < t < b_1 + b_2 + b_3 + b_4 ; \\ 1, b_1 + b_2 + b_3 + b_4 \le t ; \end{cases}$$

$$\chi_t(p) = \frac{(1 - e^{-b_1 p})(1 - e^{-b_2 p})(1 - e^{-b_3 p})(1 - e^{-b_4 p})}{b_1 b_2 b_3 b_4 p^4},$$

$$M(T) = Me(T) = Mo(T) = \frac{b_1 + b_2 + b_3 + b_4}{2},$$

$$D(T) = \frac{b_1^2 + b_2^2 + b_3^2 + b_4^2}{12}.$$

For example, a 0.2-quantile $t_{0.2} = 2.56716$ and a 0.8-quantile $t_{0.8} = 4.4328 = 7 - 2.56716$ with $b_1 = 1$, $b_2 = 1.5$, $b_3 = 2$, $b_4 = 2.5$.

$$m_2(T) = \frac{b_1^2 + b_2^2 + b_3^2 + b_4^2}{3} + \frac{b_1 b_2 + b_1 b_3 + b_1 b_4 + b_2 b_3 + b_2 b_4 + b_3 b_4}{2};$$

$$m_3(T) = \frac{b_1^3 + b_2^3 + b_3^3 + b_4^3}{4} + \frac{b_1 b_2 (b_1 + b_2) + b_1 b_3 (b_1 + b_3) + b_1 b_4 (b_1 + b_4) + b_2 b_3 (b_2 + b_3)}{2} +$$
$$+ \frac{b_2 b_4 (b_2 + b_4) + b_3 b_4 (b_3 + b_4)}{2} + \frac{3}{4}\left(b_1 b_2 b_3 + b_1 b_2 b_4 + b_1 b_3 b_4 + b_2 b_3 b_4\right);$$

$$m_4(T) = \frac{b_1^4 + b_2^4 + b_3^4 + b_4^4}{5} + \frac{b_1 b_2 (b_1^2 + b_2^2) + b_1 b_3 (b_1^2 + b_3^2) + b_1 b_4 (b_1^2 + b_4^2) + b_2 b_3 (b_2^2 + b_3^2)}{2} +$$
$$+ \frac{b_2 b_4 (b_2^2 + b_4^2) + b_3 b_4 (b_3^2 + b_4^2)}{2} + \frac{2\left(b_1^2 b_2^2 + b_1^2 b_3^2 + b_1^2 b_4^2 + b_2^2 b_3^2 + b_2^2 b_4^2 + b_3^2 b_4^2\right)}{3} +$$
$$+ b_1 b_2 b_3 (b_1 + b_2 + b_3) + b_1 b_2 b_4 (b_1 + b_2 + b_4) + b_1 b_3 b_4 (b_1 + b_3 + b_4) +$$
$$+ b_2 b_3 b_4 (b_2 + b_3 + b_4) + \frac{3}{2} b_1 b_2 b_3 b_4;$$

$$Ex(T) = \frac{9\left(b_1^4 + b_2^4 + b_3^4 + b_4^4\right) + 30\left(b_1^2 b_2^2 + b_1^2 b_3^2 + b_1^2 b_4^2 + b_2^2 b_3^2 + b_2^2 b_4^2 + b_3^2 b_4^2\right)}{5\left(b_1^2 + b_2^2 + b_3^2 + b_4^2\right)^2}.$$

The proposed approach can be used to derive analytical expressions for the sum of a larger number of random variables, however, the bulkiness of the formulas increases significantly. The implementation of the proposed calculation formulas in computer programs does not cause difficulties. The temporal computational complexity is determined by the number of operations of multiplication and division when calculating the values of the functions in the last sub-interval of determining the functions of the sum of random variables. For example, for $F(t)$ with $n = 2$, the number of operations of multiplication and division is 6, for $n = 3$ – 18, and for $n = 4$ – 50, this means that the number of such operations obeys a quadratic dependence on n. Thus, the computational complexity of the formulas obtained is $O(n^2)$.

To illustrate the application of the obtained formulas, various deciles are calculated. When the probability of completion of information processing is 0.1, the duration of the information process in a chain of four elements having evenly distributed random durations in seconds at intervals [0; 15], [0; 15], [0; 15] and [0; 40] is equal to $t_{0.1} = 24.25$ s.

Risk assessments of untimely data processing are determined through decile calculations [18]. For example, to calculate the duration of processing information in the same chain at a risk level of 10%, the equation $F_{1234}(t_{0.9}) = 1 - 0.1$ is numerically solved, which implies $t_{0.9} = 60.75$ s. To calculate deciles, we used the traditional method of numerical solution of nonlinear equations.

The graphs $f_{12}(t), f_{123}(t), f_{1234}(t)$ are shown in the figure.

Summary. The obtained analytical dependences are applicable for independent random variables of the genus. The implementation of the formulas obtained in the form of computer programs allows the use of more complex distribution functions of random variables for more adequate modeling of data processing in infocommunication systems.

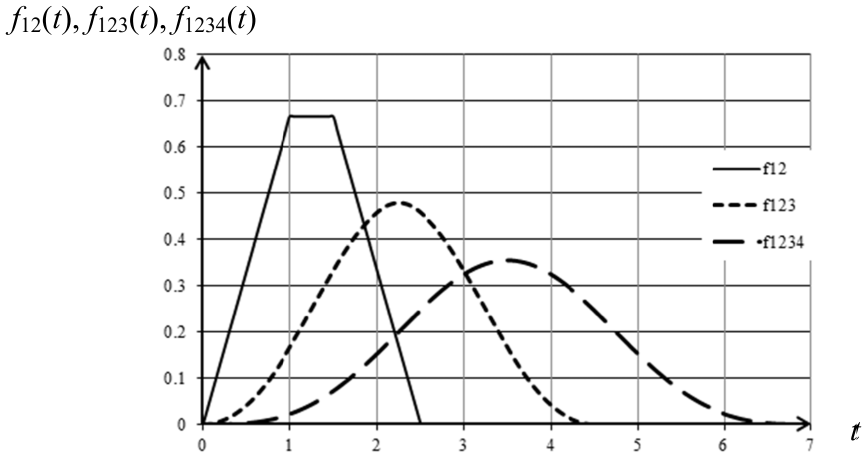

Figure 1- Graphs of the PDF $f_{12}(t)$, $f_{123}(t)$ and $f_{1234}(t)$ for $b_1 = 1$, $b_2 = 1.5$, $b_3 = 2$, $b_4 = 2.5$.

М.Ю. Бабич[1], М.М. Бутаев[1], А.А. Тарасов[1], А.И. Иванов[2], А.Ю. Малыгин[2], К.Т. Сауанова[3], Ш.Н. Сагындыкова[3]

[1]«Рубин» ғылыми-өндірістік кәсіпорны» АҚ, Пенза, Ресей;
[2] Пенза мемлекеттік университеті, Пенза, Ресей;
[3]Алматы энергетика және байланыс университеті, Алматы, Қазақстан

ИНФОКОММУНИКАЦИЯЛЫҚ ЖҮЙЕНІҢ СҰРАНЫСТЫ ӨНДЕУДЕГІ БІРКЕЛКІ ҮЛЕСТІРІМДІ КЕЗДЕЙСОҚ УАҚЫТ ҚОСЫНДЫСЫНЫҢ ТАРАЛУ ФУНКЦИЯЛАРЫ

Аннотация. Инфокоммуникациялық жүйелерді жобалаудың бастапқы кезеңдерінде шешілген маңызды міндеттердің бірі – деректерді өңдеудің жеке кезең ұзақтығының жүйеде жалпы өңдеу уақытына әсерін бағалау. Бұл мәселе деректерді өңдеу кезеңі ұзақтығының ықтималды сипаттамаларын есептеу арқылы шешіледі.

Математикалық тұрғыдан өңдеу уақытын есептеу мәселесі әдетте ациклдік бағытталған стохастикалық график арқылы шешімін табады. Аталған әдістер PERT әдісі арқылы желілік сұлбаларды есептеу үшін қолданылады. PERT есептеу әдістері ауқымды желілерге арналған, сондықтан бөлек кезеңдер мен тұтас барлық желінің орындалу уақыты үшін ықтималдық үлестірім заңдылығын қолдануды негіздейтін ықтималдық теориясының орталық шекті теоремасы қолданылады. Желідегі мәліметтерді өңдеу кезеңдері аз болған жағдайда бұл тәсіл (-∞, + ∞) аралығында анықталған қалыпты тарату заңын қолдануда қиындықтар туғызады. Тәжірибеде кезеңдегі мәліметтерді өңдеудің минималды ұзақтығы априорлы белгілі, ол мәліметтерді өңдеудің физикалық сипаттамаларына байланысты, ал максималды ұзақтығы техникалық тапсырма талаптары негізінде анықталады. Кезең ұзақтығының орташа мәні немесе мода ықтималдығы өте аз, сондықтан оған негізделген есептеуге деген сенім де шамалы. Ең ұзақ және қысқа өңдеу уақытын есептеу (интервал әдісі) ақпараты жеткіліксіз. Мұндай жағдайда кезеңдер бойынша деректерді өңдеу уақытының біркелкі тығыздық ықтималдық әдісін қолданған жөн. Кезең бойынша дәйекті түрде орындалғанда сұранысты өңдеу ұзақтығы әр кезеңде өңдеу уақытының қосындысы – тәуелсіз кездейсоқ шама қосындысы есептеледі. Тәуелсіз кездейсоқ шама қосындысының ықтималдылықты үлестіру функциясы мүшелердің таралу функциялары бойынша конволюция формуласы арқылы есептеледі.

Жүйедегі ақпаратты өңдеу элементтерінің байланыс топологиясы әртүрлі болуы мүмкін, бірақ ақпаратты өңдеудің шектік ұзақтығын бағалау үшін жүйенің уақытша сипаттамасын анықтайтын тізбекті түрдегі элементтер тізбегін ерекшелейді. Осындай элементтер тізбегінде ақпаратты өңдеу ұзақтығын бағалау, әрине, олардың әрқайсысындағы өңдеу уақытының қосындысына алынады. Уақытша сипаттамаларды тереңірек бағалау үшін үздіксіз кездейсоқ шаманың ықтималдық тығыздығы функциясын f(t) және / немесе F(t) үлестірім функциясын білу қажет және бұл мәселенің ең жақсы шешімі – таралу мен тығыздық функцияларының дәл аналитикалық сипаттамасы.

Жұмыстың мақсаты – инфокоммуникациялық жүйелердің ықтималдық сипаттамаларын есептеу әдістерін жетілдіру. Зерттеу нысаны – жүйеде сұранысты өңдеу уақытын есептеудің аналитикалық әдісі, тақырыбы – біркелкі бөлінген тәуелсіз кездейсоқ өңдеу уақыты бар жүйе элементтері арқылы сұранысты ретімен өңдеу ұзақтығын есептеу формуласы. Екі, үш және төрт тәуелсіз біркелкі үлестірілген кездейсоқ шама қосындысы үшін тығыздық формулалары мен ықтималдықтың функциялары келтірілді.

Түйін сөздер: ықтималдық тығыздығы, үлестіру функциясы, инфокоммуникациялық жүйелердің ықтимал сипаттамасы.

М.Ю. Бабич[1], М.М. Бутаев[1], А.А. Тарасов[1], А.И. Иванов[2], А.Ю. Малыгин[2], К.Т. Сауанова[3], Ш.Н. Сагындыкова[3]

[1]АО «Научно-производственное предприятие «Рубин», Пенза, Россия;
[2]Пензенский Государственный Университет, Пенза, Россия;
[3]Алматинский Университет Энергетики и Связи, Алматы, Казахстан

ФУНКЦИИ РАСПРЕДЕЛЕНИЙ СУММ РАВНОМЕРНО РАСПРЕДЕЛЁННЫХ СЛУЧАЙНЫХ ЗНАЧЕНИЙ ВРЕМЕН ОБРАБОТКИ ЗАПРОСА ИНФОКОММУНИКАЦИОННОЙ СИСТЕМОЙ

Аннотация. Одной из важных задач, решаемых на начальных этапах проектирования инфокоммуникационных систем является оценка влияния продолжительностей отдельных этапов обработки данных на общую продолжительность обработки в системе. Эта задача наиболее информативно решается через расчёт вероятностных характеристик длительностей этапов обработки данных. Математически задачи расчёта времени обработки данных обычно решаются с помощью ациклического направленного стохастического графа. Этот же аппарат используется для расчёта сетевых графиков методом PERT. Расчётные методы PERT созданы для сетей большой размерности, поэтому использована центральная предельная теорема теории вероятности, обосновывающая применение нормального закона распределения вероятности времени выполнения отдельных этапов и всей сети в целом. Такой подход при небольшом числе этапов обработки данных в сети вызывает сложности в обосновании использования нормального закона распределения, который определён на интервале (-∞, +∞). На практике априорно известны минимальная продолжительность обработки данных на этапе, она обусловлена физическими особенностями обработки данных на этапе, а максимальная продолжительность определяется требованиями технического задания. Среднее значение или мода продолжительности этапа имеют место с малой вероятностью, поэтому расчёт на их основе имеет небольшую достоверность. Расчёт только наибольшего и наименьшего времени обработки данных (интервальный метод) недостаточно информативен. В таких условиях целесообразно использовать вероятностный подход с равномерной плотностью длительностей обработки данных на этапах. При последовательном выполнении этапов продолжительность обработки запросов сводится к суммированию продолжительностей обработки на каждом из этапов, т.е. к суммированию независимых случайных величин. Функция распределения вероятности суммы независимых случайных величин рассчитывается через формулу свёртки функций распределения слагаемых.

Топология соединений элементов обработки информации в системе может быть различной, но для оценки критической продолжительности обработки информации выделяют последовательную цепочку элементов, которая, вероятно, будет определять временные характеристики системы. Оценка продолжительности обработки информации в такой цепочке элементов, очевидно, сводится к суммированию продолжительностей обработки в каждом из них. Для более глубокой оценки временных характеристик необходимо знать функцию плотности вероятности f(t) и/или функцию распределения F(t) непрерывной случайной величины, и лучшим решением этой задачи является точное аналитическое описание функций распределения и плотности.

Цель работы – совершенствование методов расчёта вероятностных характеристик инфокоммуникационных систем. Объектом исследования является аналитический метод расчёта времени обработки запроса в системе, предметом – формулы расчёта продолжительности последовательной обработки запроса элементами системы с равномерно распределёнными независимыми случайными временами обработки. Приведены формулы плотности и функции вероятности для сумм двух, трёх и четырёх независимых равномерно распределённых случайных величин.

Ключевые слова: плотность вероятности, функция распределения, вероятностные характеристики инфокоммуникационных систем.

Information about the authors:

Babich M. Doctor of Technical Sciences, associate professor, JSC ”Scientific-Production Enterprise “Rubyn”", Penza, Russia, mail@npp-rubin.ru. https://orcid.org/0000-0003-2969-1055;

Butaev M.M. Doctor of Technical Sciences, professor, JSC ”Scientific-Production Enterprise “Rubyn”", Penza, Russia, mail@npp-rubin.ru. https://orcid.org/0000-0001-6421-0048;

Tarasov A.A. Candidate of Technical Sciences, JSC ”Scientific-Production Enterprise “Rubyn” ", Penza, Russia, mail@npp-rubin.ru. https://orcid.org/00000-0001-7028-836X;

Ivanov A.I. Doctor of Technical Science, associate professor, Penza State University, Penza, Russia, bio.ivan.penza@mail.ru. https://orcid.org/0000-0003-3475-2182;

Malygin A.Yu., Doctor of Technical Sciences, professor, Penza State University, Penza, Russia, mal@yandex.ru. https://orcid.org/0000-0002-1410-7269;

Sauanova K.T. Candidate of Technical Sciences, AUPET, Almaty, Kazakhstan. e-mail:sauanovaklara@gmail.com. http://orcid.org/0000-0001-8484-5446;

Sagyndykova Sh.N Candidate of Technical Sciences, AUPET, Almaty, Kazakhstan, tomka2001@mail.ru. http://orcid.org/0000-0001-5909-584X

REFERENCES

[1] Ralph L. Keeney, Howard Raiffa (1993) Decisions with Multiple Objectives:Preferences and Value. Cambridge University press. ISBN: 0-521-43883-7 (in Eng.).

[2] Green J.M, Johnson B.W. (2002) Towards a Theory of Measures of Effectiveness.Dubley Knox Library. Monterey California. https://www.nps.edu/library.

[3] Kleinrock L. (1964) Queuing systems: vol. 1 - Theory. Wiley Interscience, New York. ISBN: 978-0471491101.

[4] Sovetov B.Ya., Yakovlev S.A.(2013) Modeling systems [Modelirovanie sistem]. Moskva:YUrajt. ISBN: 978-5-9916-26989 (in Russ.)

[5] Roy N., Dubey A., A. Gokhale, Dowdy L. A (2011) Capacity Planning Process for Performance Assurance of Component Based Distributed Systems. ICPE'11 - Second Joint WOSP/SIPEW International Conference on Performance Engineering, Karlsruhe, Germany, 39(3):259-270. DOI: 10.1145/1958746.1958784.

[6] Golenko-Ginzburg D.I. (2010) Stochastic Network Models for Development Planning and Management: Monograph. Voronezh : "Scientific Book ". ISBN 978-5-98222-646-4

[7] Weng- Ming Chu, Koan-Yuh Chang, Chien -Yu Lu, Chang-Hung Hsu, Chien -Hung Liu, Yung-Chia Hsiao. (2014) The New Approach to A the Determine the Critical Stochastic has the Path in the Activity the Network. Mathematical Problems in Engineering 2014(3):1123-1128. International Symposium on Computer, Consumer and Control (IS3C). DOI: 10.1109 / IS3C.2014.292 2014.292.

[8] Leemis LM, Duggan M.J, Drew J.H, Mallozzi J.A, Connell K.W. (2006) Algorithms to Calculate the Distribution of the Longest Path Length of a Stochastic Activity Network with Continuous Activity Durations. Networks. 48 (3): 143-165. DOI: 10.1002/net.20125.

[9] HOU Zhen-ting, ZHANG Xuan, KONG Xiang- xing.(2010) A new analytical algorithm for computing probability distribution of project completion time. Journal of Central South University of Technology: 1006-1010. DOI: 10.1007 / s11771-010-0591-4.

[10] Dubois D., Fargie, H., Fortin J. (2005) Computational methods for determining the latest starting times and floats of tasks in interval-valued activity networks Journal of Intelligent Manufacturing 16(4):407-421. DOI: 10.1007/s10845-005-1654-5.

[11] Abdelkader Y.H, Al-Ohali M. (2013) Estimating the Completion time of Stochastic Activity networks with uniform distributed activity Times Archives Des Sciences Journal Vol 66(4):115-134. https://www.researchgate.net/publication/236135085_Estimating_the_Completion_time_of_Stochastic_Activity_networks_with_uniform_distributed_activity_Times.

[12] Sireesha V., Ravi Shankar N. (2011) Analysis of time estimates on right skewed distribution of activity times in PERT // Journal of Statistics and Mathematics. V. 2, Issue 1:15-22. http://www.bioinfo.in/contents.php?id=85

[13] Feller V. (1971) Introduction to probability theory and its applications. V. 2. ISBN: 978-0-471-25709-7.

[14] Vadzinsky R. N. (2001) Handbook of probability distributions. St. Petersburg Science. ISBN: 5-02-024919-X.

[15] Kendall M. Stuart A. (2006) Advanced Theory of Statistics. Distribution Theory. Journal of the American Statistical Association. DOI: 10.2307 / 27639788.

[16] Butaev M.M., Tarasov A.A. (2018) An analytical estimate of the processing time of a request by an information system. Radio electronics issues.V 12. : 69-73.(in Russ.).

[17] Korn G., Korn T. (2010) Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review.Dover Publications, Inc, Mineola, New York. ISBN:0-486-41147-8.

[18] Udoumoh E.F., Ebong D.W. (2017) A Review of Activity Time Distributions in Risk Analysis American Journal of Operations Research 07(06):356-371. DOI: 10.4236/ajor.2017.76027.