Visiblestructureshighlighting model analysis aimed at object image detectionproblem

V.V. Mokshin[1], I.R.Saifudinov[1], P.I. Tutubalin[1], L.M. Sharnin[1]

[1]Kazan National Research Technical Universitynamed after A. N. Tupolev - KAI,10 Karl Marx Str., Kazan, Russia, 420111

Abstract. The research considers an approach to solving the problem of reducing the data processed in mobile platformsoriented video-analytic systems.Different models of human visual attentionhas been analyzed and also classified according to the image segmentation.The results obtained are presented in the form of the method used for isolating the borders of the most significant object in the image. They are based on the optimization the length and curvature values of the object borders. This approach allows to filter significant structures in the image and process them with the help of image segmentation techniques.The method was evaluated according to usage the accuracy criterion along with other methods in delineation of boundaries: threshold value, morphological processing and watershed. Software and hardware system for registering dump trucks has been improved. It gives the possibility to automatize the process of registration dump trucks, training them during the construction of roads.

Keywords:pattern recognition, structural visibility, image map.

1. Introduction

Automated video surveillance systems (video analytics) used in many spheres of human activity, likewise state institutionsand manufacturing enterprisesare extremely popular nowadays. The systems usage allows tocreate options forevents automaticalarm, efficiencyin employees labour improvement by means of direct control of the performing activity.Herewith, because of the large number of tasks solved in video analytics, as well as their multi-criteria,autonomy tendency, and indistinct nature of tasks, there is a certain necessity to find the effective approach for highlighting image visible structures that allows to work on the basis of mobile platforms.The attempt for analysis of various tasks has beenmade in the research. [1], [3], [6-8].

The method of highlighting of visible structures in an image based on visibility measuring of length and curvature of the curve that is similar to the concept suggested by Lowe [4]has been taken into consideration. The research deals with the measure of visibility according to the criterion of false positive detection in comparison to the method of delimiting contour of object [2].

2. Visibility network construction

Orientation elements are the basic computational elements of the network [9]. Each element $p_i$is connected with a processor that normally can perform certain calculations based on the conditions and those with $k$denotation performed by the neighbour processor.This defines a single network containing kn^2 processing blocks with the local communication. In the current implementation, k is equal to 48, which provides a reasonable angular resolution.Let's appeal to the associatedorientation sequence of the elements$p_i, \dots, p_{i+N}$, , where each element is a linear segment or interval, like lengthof a curve$N$. (curves can be continuous or with any number of intervals). The optimization task is formulated to maximizethevalue of $\Phi_N$above overall length of a curve$N$, starting with p_i.

$$\max_{(p_{i+1},\dots,p_{i+N})\in\delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N})$$

where$\delta^N(p_i)$ is a number of all possible length of a curve$N$, starting with p_i.

For a certain of measuresclass$\Phi(\cdot)$, the calculation of Φ_N can be obtained by simple local computations made repeatedly.To illustrate, let’s take the first curves three elements long only. In this case:

$$\max_{(p_{i+1},p_{i+2})\in\delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2})$$

wherep_i is definedby p_{i+1}(one of p_i's $k$of neighbours) and p_{i+2} (p_{i+1} neighbour) for a given elementsothatthe rate$\Phi_2(p_i, p_{i+1}, p_{i+2})$ will be in its maximum. Simple approach (brute-force method)in the analysis of k^2 value in different curveswill be required anew.Suppose, however, that $\Phi_2$corresponds to the condition of:

$$\max_{\delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2}) = \max_{p_{i+1}} \Phi_1(p_i, \max_{p_{i+2}} \Phi_1(p_{i+1}, p_{i+2}))$$

In this case, the maximization of the rateΦ_2 can be achieved by application of $\Phi_1$used repeatedly over shorter curves. The general approach can be formulated the same way, e.i:

$$\max_{\delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N}) = \max_{p_{i+1}\in\delta(p_i)} \Phi_1(p_i, \max_{\delta^{N-1}(p_{i+1})} \Phi_{N-1}(p_{i+1}, \dots, p_{i+N})) \quad (1)$$

where$\delta(p_i)$ is equal to$\delta^1(p_i)$. Thus, the searching area required for each lengthof a curve N is being reduced, from p_i tokN, instead of k^Nwhich is essential for a brute-force method approach.The concept(1) is related to the optimality principle, underlining for all multistage decision-making processes. This is a special case in dynamic programming in particular. It refers to the family of functions that follows concept (1) ofextensible functions.

There are two factors that are essential for visibility measure.The first one is related to the length of a curve, and the second factor is related to its shape. The length of a curve is determined by the number of its elements that have a factual curve(rather than an interval) passing through these elements. They are called active elements. Whereas elements that are associated with intervals are referred to as virtual elementswhere local visibility $\sigma_i$corresponds to p_i. If p_i is the active element, then σ_i has positive value, which is equal to 1 and 0 for the virtual elementσ_i. A measure associated with the length of a curve $p_i, \dots, p_{i+N}$ is determined by the equation:

$$\sum_{j=i}^{i+N} \sigma_j \quad (2)$$

The measure rate above(2) presents the sum of the local values of the visibility of active elements along the curve.

Then, the attenuation function associated with the curve $p_i, \dots, p_j$ is defined as follows:

$$\rho_{i,j} = \prod_{k=i+1}^{j} \rho_k$$

where$\rho_{i,i} = 1$. The measure in (2) is modified by the attenuation coefficients is:

$$\sum_{j=i}^{i+N} \rho_{i,j}\sigma_j \quad (3)$$

The measure rate in equation (3) is weighted contribution of the local visibility values σ_j along a curve, that are in reverse dependency to a number of virtual elements along $p_i, \dots, p_j$. To measure the shape of a curve, measurethatin reverse dependency to the total curvature of a curve is used. The total curvature of γ is defined as$\int_\gamma \left(\frac{d\theta}{ds}\right)^2 ds$, where $\theta(s)$is a slope along a curve and$\frac{d\theta}{ds}$at the point P is known as the local curvature at this point(the reciprocal value of the radius of curvature R).It is necessary to use the total curvature to obtain a measure that is limitedandinthe position of reverse dependency to the total curvature. The next measure is relevant to the following:

$$\exp^{-\int_{\gamma}\left(\frac{d\theta}{ds}\right)^2 ds} \tag{4}$$

In order to obtain a discrete approximation to the measure in (4), we denote$a_k$to indicate orientation difference between the k-th element and its successor andΔsas length of the orientation element. A discrete approximation to full curvature of a measure along $p_i, \dots p_j$, will be:

$$C_{i,j} = \prod_{k=i}^{j-1} f_{k,k+1}$$

where

$$f_{k,k+1} = \exp^{-\frac{2a_k \operatorname{tg}\frac{a_k}{2}}{\Delta s}} \tag{5}$$

$C_{i,j}$ is the weight of each value of local visibility σ_j along a curve. A measure that shows a high rate for long curves with low overall curvature is now defined as:

$$\sum_{j=1}^{i+N} C_{i,j}\, \rho_{i,j}\, \sigma_j \tag{6}$$

The measure in equation (6) is weighted contribution of the local visibility values σ_j along a curve. The curves that will receive a high measure on (6) are long curves, more straight with the least number of intervals.

3. Analysis of the visibility network: discussion and results.

To analyse visibility network we compared the percentage of false positive detections. Test sampleslocatedaround the perimeter of circle at equal intervals consist of short, oriented segmentsin the field of segments with random position and orientationas it shown in figure 1.

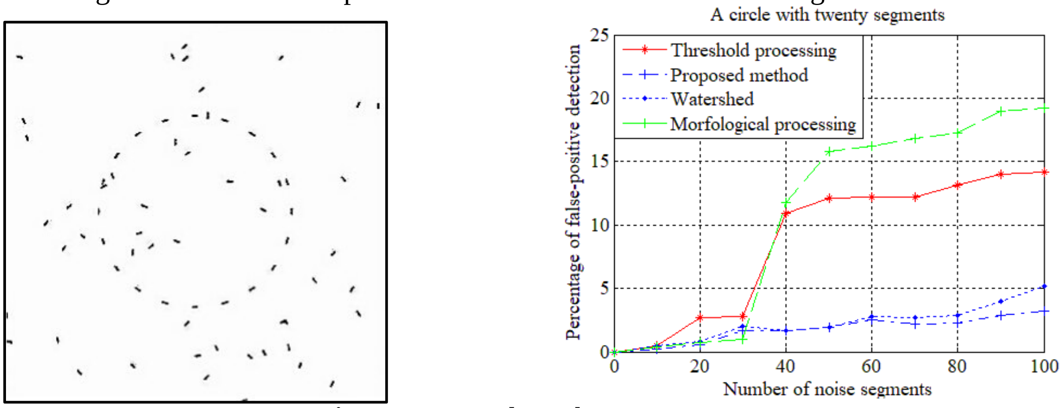

Figure 1. A circle with twenty segments.

The methods that are used for the calculation visibility forms and noise segments are: the threshold value [10], the watershed [11,13], the morphology [12], and the proposed approach.Segments were sorted in ascending order according to their visibilityof the most (ϕ_1) and less (ϕ_n) noticeable segments. For given $m$form segments, false-positive are defined as noise segments, which are assigned a visibility greater than ϕ_{m+1}.A false positive estimate for each method was calculated for samples consisting of different numbers of shape and noise segments rate. A false positive rate for each combination (for example, 20 shape segments and 70 noise segments) was estimated by averaging false positive value by means of more than ten attempts with different noise samples.The picture in the right side of Figure 2 is the graph of false positive rate percentage for a circle with twenty segments.

Each method can be accomplished well enough (less than 10% of false positives) at a low noise level (40 noise segments or less). The results of methods applied begin to diverge at higher noise levels. It should be noted that threshold processing is superior to morphological processingwhile watershedisrelative to the approach suggested, although the latter is more expensive to calculate. And, finally, at lower signal-to-noise ratios watershed and the proposed approach have significantly lower false-positive estimates.The next comparison was identical to the first one, but shape segments are formed by an unlimited sinusoid (Figure 2).

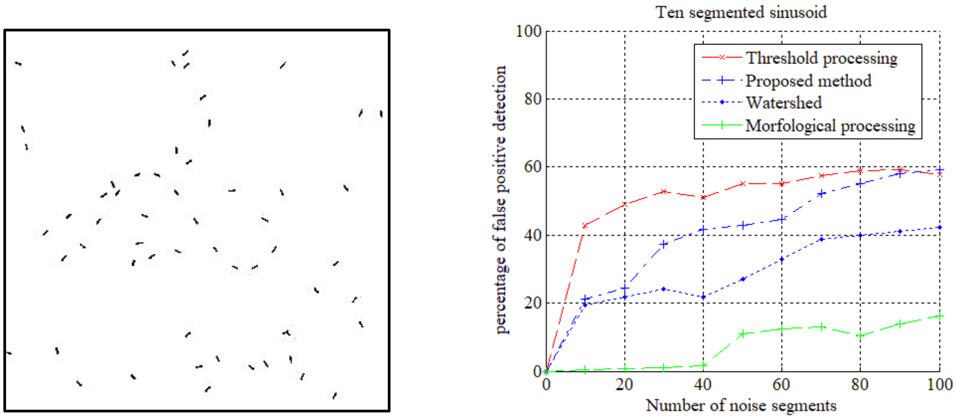

Figure 2.A circle with a ten segment sinusoid.

In the right half of Figure 2 is a graph of the percentage of false positives relative to the number of noise segments for a ten segment sinusoidal curve.Comparatively low indicators of the proposed method compared with other methods can be attributed to its apparent dependence on closure.However, it still outperforms the threshold processing for higher signal-to-noise ratios and has an error rate comparable to threshold processing (i.e., within 5%) at lower signal-to-noise ratios.

In the third comparison a field consisting of correlated noises (i.e., a dipole) was used (Figure 3).The dipole consists of two collinear segments separated by an interval equal to the distance between neighboring segments of the circle.Since the two segments forming the dipole are collinear, the degree of closeness between the segments forming the dipole is greater than between the adjacent segments of the circle.Therefore, it is impossible to distinguish noise segments from the shape segments using only a local measurement. From the graph it can be seen that the methods of threshold processing and morphological processing have almost a 100% false positive level, whereas the watershed and the proposed method are much better able to cope with this task

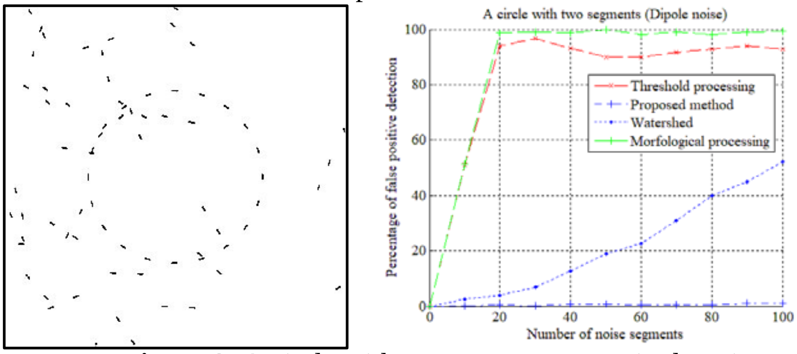

Figure 3. A circle with twenty segments. Dipole noise.

In the fourth comparison (Figure 4), ten segments are used. This is a complex picture, because the sampling frequency is so small, so that there is only one segment at 36 degrees of the circle. Most of methods do not work well even at relatively high signal-to-noise ratios. For noise level 80, threshold processing and the morphological processing method are performed at 90% false positive level. The watershed method is performed a bit better, with a false-positive level of 70%.In contrast, the false-positive level for the proposed method is less than 5%.

Let's consider practical use examples of the specified methods in the decision of the object segmentation problem on the image in the video-analytical system focused on mobile platforms for the account number of transportations dump trucks. The image segmentation results are shown below.Segmentation for threshold processing is shown in Figure 5.

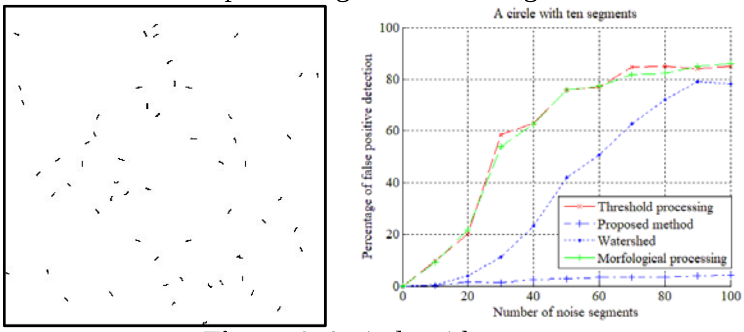

Figure 4. A circle with ten segments.

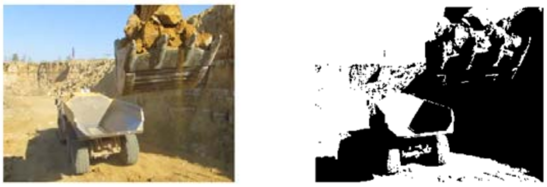

Figure 5. Segmentation. Threshold value.

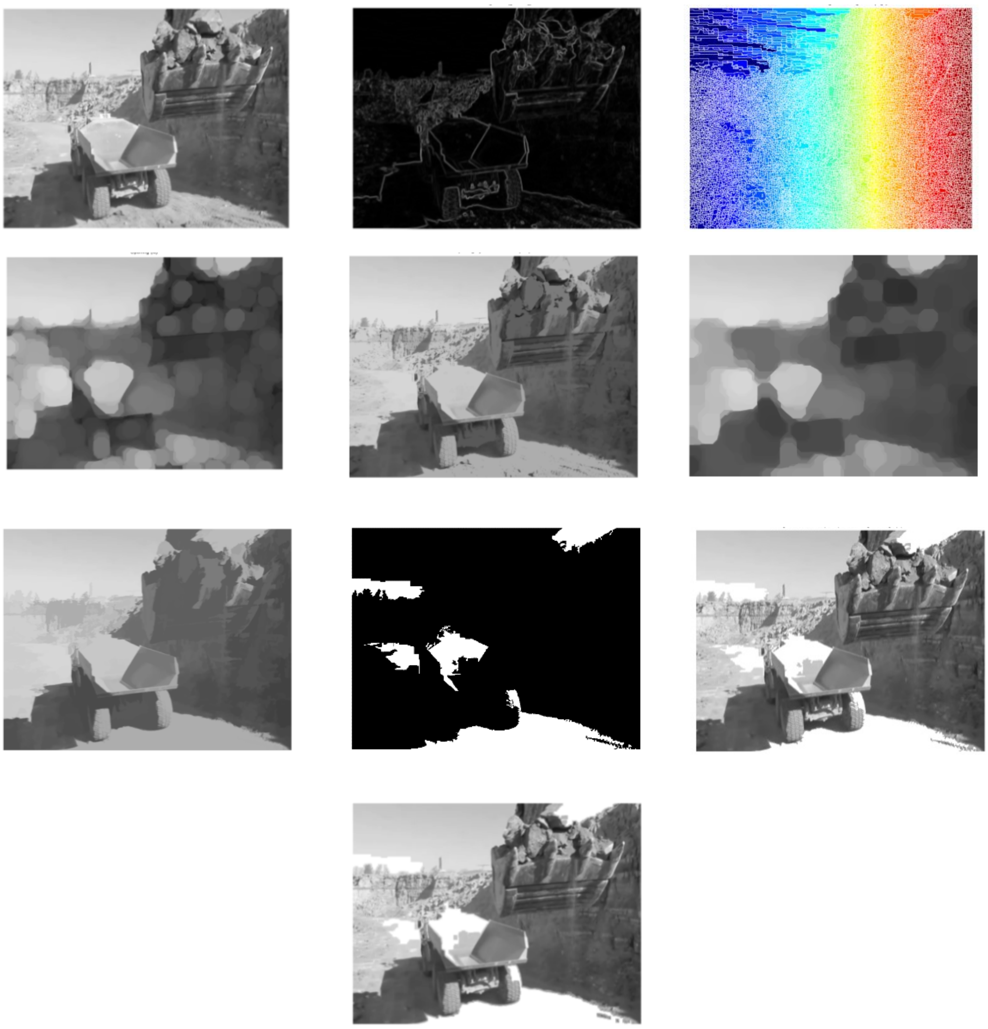

Figure 6. Stages of watershed segmentation. From left to right, from top to bottom.

Segmentation for the watershed is shown in Figure 6.

The results of morphological processing segmentation are shown in Figure 7.

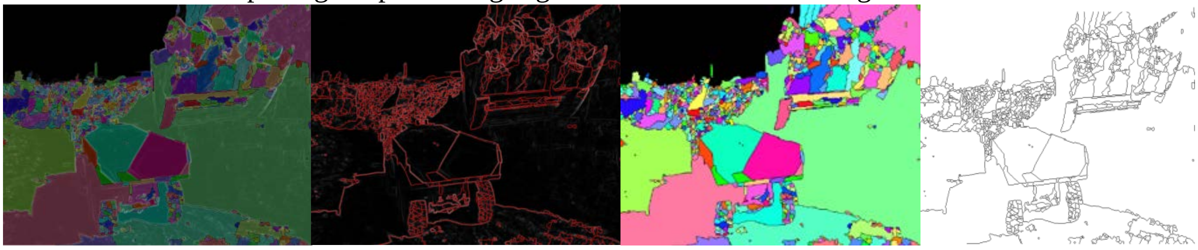

Figure 7. Stages of morphological processing segmentation. From left to right.

The results of the proposed methodsegmentation for isolating significant structures are shown in Figure 8.

Figure 8. Left. Preprocessing an image using the Sobel operator. The middle. Map of the importance of the image. On right. The selected area of the image.

4. Conclusions

In addition to comparing false positive detection, Table 1 shows the time required to run the algorithms on the Android 5.0 operating system. with a resolution of 640x480 pixels.It can be seen from the table that the proposed method for identifying significant structures shows the fastest and least resource-consuming results.Thus, the proposed model for significant structuresidentifying is an effective tool for primary segmentation of images in video-analytical systems oriented at mobile platforms.As results have shown, the method of isolating notable structures showed the most accurate results.The method of selecting notable structures is based on the measure of visibility calculation that includes applies the principle of dynamic programming [5, 15].

Table 1. Comparison of the time spent on segmentation and resource intensity.

Method	Calculation time (seconds)	RAM Consumption (MB)
Threshold processing	1.36	70
Watershed	2.03	105
Morphological treatment	1.78	85
Allocation of significant structures	1.09	45

Also it would be used queueing system to analyse such kind models [16, 17].Two internal properties of the network of significance, such as extensibility and geometric convergence, allow us to optimize the measure of significance and effectively restore optimal curves (in polynomial time).

At the same time, they restrict the range of possible functions that can be used as a measure of significance, so the method has some limitations with scale invariance, fusion of curves, and grouping in the presence of compounds.In addition, overcoming overlapping problems will require asymptotically increasing the complexity of the method, since the discretization is closely related with using dynamic programming to effectively optimize the chosen measure.

5. References

[1] Moksin, V.V. Vehicle recognition based on heuristic data and machine learning / V.V.Moksin, I.R.Saifudinov, A.P.Kirpichnikov, L.M. Sharnin// Bull. KazanTechn. Univ. –2016. –Vol. 19(5).– P. 130-137. (inRussian)
[2] Saifudinov, I.R. Grouping contours of objects of structural images based on the network of element visibility / I.R.Saifudinov, V.V.Mokshin, A.P. Kirpichnikov//Bull. Kazan Techn. Univ.– 2017. –Vol. 20(9). –P. 120-123. (in Russian)
[3] Treisman, A. Perceptual Grouping and Attention in Visual Search for Features and for Objects / A.Treisman// Journal of Experimental Psychology: Human Perception and Performance. – 1982. – Vol. 8(2). – P. 194-214.
[4] Lowe, D.G. Perceptual Organization and Visual Recognition /D.G. Lowe // Kluwer Academic Publishers, Boston, Mass., 1985.The book is an extended version of the PhD Thesis with same title, Computer Science Dept., Stanford University, 1984.
[5] Kormen, T.Dynamic programming / T.Kormen, Ch. Leiserson, R. Rivest, K. Shtaine// Algorithms: construction and analysis = IntroductiontoAlgorithmsedited by Krasik. Moscow: Williams, 2005. – 1296 pp. (inRussia)
[6] Gorilik, A.L. Recognition methods/ A.L. Gorilik, V.A. Skripkin // Higher School Publishing. – 1989.–232 pp.(in Russia)
[7] Vapnik, V.N. Theory of pattern recognition / V.N.Vapnik, A.I. Chervonenkis. – Moscow: Science, 1974. –416 pp. (inRussia)
[8] Nesteruk, V.F. Questions of the theory of perception of subject images and a quantitative assessment of their contrast / V.F.Nesteruk, V.A.Sokolova// Optoelectronic industry. –1980. –Vol. 5. –P.11-13. (inRussia)
[9] Shashua, A.Structural saliency: The detection of globally salient structures using a locally connected network / A. Shashua, S.Ullman // 2nd Intl. Conf. on Computer Vision (ICCV ’88), Clearwater, FL, 1988.
[10] Laurent,H. Figure-Ground Discrimination: a Combinatorial Optimization Approach / H.Laurent, H.Radu // IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers. – 1993. – Vol. 15(9). –P.899-914.
[11] Sarkar, S. Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors / S. Sarkar, K.L. Boyer // Computer Vision and Pattern Recognition,1996.
[12] Lance, R. Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Salience / R.W. Lance, W. David // Neural Computation. – 1997. – Vol. 9(4). –P.837-858.
[13] Shakowat, Md. Morphological based technique for image segmentation / Md.Shakowat, Z.Sarker, W.H. Tan, R.Logeswaran// International Journal of Information Technology. –2007. –Vol. 14(1).– P. 55-80.
[14] Sahoo, P.K. A survey of thresholding techniques / P.K. Sahoo // Computer Vision, Graphics and Image Processing. –1988. – Vol. 41. –P. 233-260.
[15] Rutkowski, W.S. Shape completion / W.S.Rutkowski// Computer Vision, Graphics and Image Processing. – 1979. – Vol. 9. – P. 89-101.
[16] Yakimov, I. The comparison of structured modeling and simulation modeling of queueing systems / I.Yakimov, A.Kirpichnikov, V.Mokshin, Z.Yakhina, R.Gainullin// Communications in Computer and Information Science (CCIS). –2017. – Vol. 800. DOI: 10.1007/978-3-319-68069-9_21.
[17] Tutubalin, P.I. The Evaluation of the cryptographic strength of asymmetric encryption algorithms / P.I. Tutubalin, V.V. Mokshin // Second Russia and Pacific Conference on Computer Technology and Applications (RPC) IEEE. – 2017. –P.180-183. DOI: 10.1109/RPC.2017.8168094.