Scholars Journal of Engineering and Technology (SJET) **ISSN 2321-435X (Online)**
Sch. J. Eng. Tech., 2014; 2(3C):491-497 **ISSN 2347-9523 (Print)**
©Scholars Academic and Scientific Publisher
(An International Publisher for Academic and Scientific Resources)
www.saspublisher.com

Research Article

Automated Controllers Functioning Criteria in Content Distribution Systems

Vitalii M. Tkachov [1,2,*], Vadym E. Savanevych[1]
[1] Department of Electronic Computers, Faculty of Computer Engineering and Control, Kharkiv National University of Radio Electronics, Lenin ave. 14 , Kharkiv, Ukraine, 61166
[2] Department of Space Radio Physics, Institute of Radio Astronomy of the National Academy of Sciences of Ukraine, Chervonopraporna St. 4, Kharkiv, Ukraine, 61002

***Corresponding author**
Vitalii M. Tkachov
Email: vitalii.tkachov@kture.kharkov.ua

Abstract: This paper describes content distribution systems (CDS) used in telecommunication datacenters. First or last, the CDS becomes large, and its detailed engineering can not be performed during reasonable time span. Therefore, its optimization is out of question. In this case, formulating a model for a huge system functioning, not the system optimization, is required. This paper describes the problem of automated controllers functioning in telecommunication up-diffused datacenters. We propose the functional criteria of the practicability CDS behavior agreed with a specification of earlier informational system criterion. It belongs to a class of additive functional-cost criterion. The mentioned problem of CDS functioning can be solved using the Lagrange's method of undetermined coefficients. This allows to develop the system with minimal average costs during the extended period of its operational time under conforming to the possibility of tasks solving. At that the behavior model parameters values are equivalent to the corresponding Lagrange coefficients. The presented approach can be practically implemented in the suballocated service systems or the systems working according to the "observer-subscriber" principle.
Keywords: Functioning criteria, Content distribution systems, Lagrange coefficients, Atemporal criterion.

INTRODUCTION

All other things being equal, a complexity of any functional system is determined by diversity of its input stimulus. Content distribution systems (CDS) used in the field of telecommunications consist of all possible input stimuli forming a request space (RS) that contains all possible combinations of users' requests during the long interval of system operation. Therefore, on the one hand, a CDS complexity is determined by the content diversity and on the other hand – by the entropy of its request space or by the average cost of acquisition, storage and distribution for the users in the general case [1,2].

The subclass of small CDS with rather simple RS can be distinguished from the CDS class. Their operation is quite monotonous. These systems have the same request type which identically penetrates through the same aggregates (subnetworks) and results in the same operating of CDS elements. The content hub in a small local network or the banking systems of information sharing by SMS can be classified as systems mentioned above. All of them are designed in classical manner. A CDS operational model is formalized; a criterion is chosen; an optimization problem is formulated and solved. For example, there is a problem of minimization of costs associated with CDS creation and management under conforming to consumer service quality.

The CDS of a small city have more complicated RS structure. There is also the possibility in principle of design of the corresponding CDS of minimal complexity in this case. Unfortunately, that possibility is only prospective in many cases. First of all, it is quite difficult to consider all possible cases of request sequences and their runtime environment since some moment. But even if it is possible, the design time of these systems will be increased considerably.

First or last, the CDS becomes large, and its detailed engineering can not be performed during reasonable time span. Besides, the estimation of optimal response for another request of such a system under condition of the certain state of its subsystems, is also impossible during the reasonable time. Thereby, there are extremal CDS that potentially provide desired service quality under the minimal costs, on the one hand, but these CDS require for too long or even infinite control time on the other hand. For example, it is not practical to synthesize the optimal content arrangement according to the certain structure of telecommunication data center.

There are the functional systems assumed detail engineering, but they are practically compromised. There are functionally similar systems which are practically realistic, but can not be designed in a proper way at the same time.

There is only one possibility of above mentioned conflict resolution. It is necessary to express the system as the complex of an executive system and mutually independent aggregates. The system itself should be provided with behavioral pattern. The executive system should operate upon the CDS functional (behavior) criterion. Provided that the executive and the aggregates may be pure abstract objects without any real implementation. The aggregate structure should be simple and available for detail engineering (optimization). Effectively, the system complexity caused by the RS inconvenience (diversity of outer requests) is transferred from the system structure complexity to its operating complexity. The biunique correspondence between structure and behavior is vanished as for these systems. These are the systems that are considered to be complex [3].

One more reason can be advanced for such an approach. According to the Von Neumann's research, a system element structure should be fixed and should not become more involved. This practical conclusion is based upon the theoretical possibility of existence of some universal automaton: the system of high-level complexity which is able to perform functions of any other system, having all necessary instructions. For example, such an invariance is typical for computer operation. That is to say, the necessity for further loss of simplicity of these systems and their aggregates vanishes at the certain evolution stage of functional systems. For this purpose it is enough to set increasingly detailed and complex instructions.

The objective of this manuscript is the implementation of CDS functional criterion and the proposal development for the selection procedure of criterion parameters.

PUBLICATION ANALYSIS

In point of fact, the mentioned above proposal provided CDS with an adaptive behavior. The practicability measure of system behavior is required for the quantitative assay of the operation of the behavioral type functional system operation

Initially, the criterion of system operation cost minimization and the criterion of entropy maximum were considered as these practicability measures. In the former case the criterion illustrated rather the system structure than its behavior. In the latter case there were some noticeable shortcomings: the reliability of entropy estimation depends on the effective goal-setting and model construction (there are no reliable criteria for the performance evaluation of these models and there is a strong necessity to apply heuristic approaches); a lot of conditional chance data are required for estimation of a complex system entropy; but there are no methods of their theoretical estimation and designation.

Up-to-date optimization methods, the methods of "operating quality" increasing and the functional models of informational systems are determined by a great number of various theoretical tendencies.

Some researchers examine the object domain of data storage and processing as chaotic system. The authors base their statements for the reason that such modern systems are highly dependent of original conditions and even insignificant change of original model results in unexpected effects.

The methods of multicriteria interpretation of these systems are also commonly used. However, such an interpretation is effective only for optimizing of limited complexity systems.

The "operating quality" (grade of its behavior practicability) of the chance models is determined by the probability distribution type $P(X,U)$ of complex event in the fact that request stream $x_j \in X$ is delivered to the system input and the response stream $u_i \in U$ appears on the system output in a certain period of time, where X and U are the multitudes of requests and responses respectively.

That model correlates with the data integration method. In that very case, the data processing system is technically determined by the group of three $< G, S, M >$, where G is a global output scheme, S is a scheme of sources and M is a relation between these schemes. The current correspondence determines the mechanisms of input request service. The multitude of such mechanisms determines the diversity of models of demand service in the data management systems.

The function $f[P(X,U)]$ with the following two properties is used for the context of the chance models as a measure or estimate of the practicability system behavior:

1) having been taken on its minimal value, the function $f[P(X,U)]$ is identical to null under the complete statistic independence of the requests and responses of system $p(x_j, u_i) = p(x_j)p(u_i)$.

2) the function $f[P(X,U)]$ takes its maximal value due to the complete relation between the request

$p(u_i/x_j) = p(x_j, u_i)/p(x_j) = 1$ and response.

3) the function $f[P(X,U)]$ is additive.

The following function meets all the characteristics.

$$f[P(X,U)] = log\frac{p(x_j, u_i)}{p(x_j)p(u_i)}, \quad (1)$$

which is a mutual random information.

Hereafter, having considered any stimulus and result of system operation to be of equal priority, average over them the function (1)

$$C = \sum_{j=1}^{J}\sum_{i=1}^{I} p(x_j, u_i) log\frac{p(x_j, u_i)}{p(x_j)p(u_i)} = I_{XU} \quad (2)$$

where *J*, *I* – a number of statuses of requests and responses of CDS respectively.

Provided that the practicability measure of system behavior (useful effect) becomes equal to average mutual information of the system multitudes of requests and responses. The costs are not included in that measure and not calculated while the estimation of system useful effect.

While the CDS design, the authors suggest to use the increment ratio of average information amount (in the quality of state space entropy difference before and after examination) to the expenditures connected with that examination, as a criteria of examination efficiency [4]. That criteria is well correlated with the functional-cost analysis.

It may be illustrated that prescribed value of average mutual information (AMI) between the requests and responses of CDS do not produce any unequivocal requirement for the consumer service quality. At that, there are the infinite set of the systems of the same AMI, even under the certain spaces of requests and responses. Thereby, the AMI may be considered as an effect measure, but it can not be regarded as behavioral measure in order to achieve the goal. The transition from the random to the average mutual information often described only the necessary criterion of goal achieving [5].

The adoption of subtest population of the functional system practicability is not a new approach. For example, the quality of speech transmission in the packet based networks is defined by the such parameters as a delay, a jitter and a percentage loss of packs.

A PRESENTATION OF TEMPORAL FUNCTIONAL CRITERIA OF LARGE TELECOMMUNICATION SYSTEM

The assumption mentioned above illustrates that different states of streams of demands and packs for CDS have different importance. The next step to the goal achieving requirement specification is a vector of average particular mutual informations between every CDS request and response with elements:

$$I_{x_jU} = \sum_{i=1}^{I} p(u_i/x_j) log\frac{p(u_i/x_j)}{p(u_i)}. \quad (3)$$

The vector with the elements calculated by using the formula (3), exactly describes the necessary criterion of the goal achieving; that vector is a useful measure of TCS (CDS specifically) purposeful behavior.

Sometimes, the ends do not justify the means. This is often more preferable to reformulate the goal than pay too high a price for its achieving. That is why the combination of different-type costs, e.g. energy (data transmission) or computational (data compaction) costs, should be taken into account. Then we can turn from the subtest population to a scalar criteria by setting a system goal weight coefficients $B = \{\beta_1, ..., \beta_I\}$ and weight coefficients which determine the system resource shortage $\Gamma = \{\gamma_1, ..., \gamma_V\}$:

$$K_{AT} = \sum_{j=1}^{J}\beta_j I_{x_jU} - \sum_{v=1}^{V}\gamma_v z_v, \quad (4)$$

where V is a number of CDS aggregates.

The functional criteria of the practicability system behavior (4) is a specification of earlier informational system criterion. It belongs to a class of additive functional-cost criterion.

The real system analysis allows to consider that weight coefficients β_j, γ_v depends on the system tasks, input system stimulus and on the state of system aids and resources. All these components are the time functions, especially during the large system engineering. Thereby, an atemporal criteria of practical behavior (4) should be exchanged for the temporal (time-dependent) criteria:

$$K_T = \sum_{j=1}^{J}\beta_j(t) I_{x_jU} - \sum_{v=1}^{V_{зап}}\gamma_v(t) z_v. \quad (5)$$

The minimization (5) is rather difficult. This minimization becomes more difficult anyway because of the lack of prospective goals of system and its state in the sequence of time intervals. Therefore, it is suggested to use a conception of lose-range interaction forming the basis of behavior of many human being objects and technical system operation [6].

According to that conception, the optimization to be provided only for current step action. The last statement allows to replace the particular mutual informations I_{x_jU} and costs z_v with their increments at steps ΔI_{x_jU} and Δz_v respectively. At that, the temporal criteria of rational behavior (5) with account of the lose-range interaction conception will be as follows:

$$K_{ТБ} = \sum_{j=1}^{J} \beta_j(t)\Delta I_{x_jU} - \sum_{v=1}^{Vзза} \gamma_v(t)\Delta z_v \; . \qquad (6)$$

The model of purposeful behavior (5), (6) is quite general and widely used. The actions of functional system, according to the maximization of corresponding criterion, would be so far various as it is required for their goal-achieving.

A FORMAL CONNECTION OF THE INTRODUCED CRITERIA WITH A GENERAL PROBLEM DEFINITION OF OPTIMAL ACTIONS OF A CONSUMER REQUEST TELECOMMUNICATION DATA CENTER

One more reason for using the functional criteria of large CDS is its conformity to a general solution of optimal action problem as for the customer demand service in the distribution data center.

Q customers with the price plan $\Omega_Q = \{\theta_1,....,\theta_j,...\theta_Q\}$ are in the operating zone of the distribution data center. The certain combination of requests is delivered to the CDS by the time t_0. Requests of the same customer or of the different customers are independent of one another. It is considered that the customer do not receive a content without request. If customer orders a i-th content, he will obtain that i-th content or no content at all (the mistakes of a content indication are not examined). Repeated requests are not examined too.

The systems of data storage (DSS) of the distribution data center have bandwidth-limited communication channels. If the "costs" means the costs for content transmission from the CDS to a customer, than we consider only a partial bandwidth capacity of its communication channel, which is used for the demand performing.

Anyhow, during h-th time interval, the distribution data center is able to transmit a total content capacity V_h, which should not exceed some capacity C_h:

$$V_h = \sum_{j=1}^{Q} V_{hj} \le C_h \, , \qquad (7)$$

where V_{hj} is a partial bandwidth capacity of communication channel associated with j-th customer's service during the h-th time interval.

While there are simultaneous limitations for the bandwidth capacity of single CDS, the limitation (7) should be complemented with following N limitations:

$$V_{hn} = \sum_{j=1}^{Q} V_{hnj} \le C_{hn} \quad \text{for} \quad n = \overline{1, N} \, , \qquad (8)$$

where C_{hn} is a bandwidth capacity of communication channel of n-th CDS during h-th time interval; V_{hn}, V_{hnj} is data transmit volume for the service channel of j-th customer of n-th CDS during h-th time interval.

The costs for service of customer's demand are the partial bandwidth capacity of data center globally or it's CDS, associated with certain customer's service or request service. There are following options as the distribution data center responses:

- to send a content from the main data center;
- to send a content from the data center, which is nearest to the customer;
- to send a content to the data center, which is next to that data center and then send it to the customer;
- to delay the request processing and send notification to the client (to queue up);
- to refuse the request.

Anyway, the costs, associated with certain v-th method of request processing from certain j-th customer Z_{vj} are supposed to be prescribed.

There are virtually no bit-errors of content transmission and time delivery requirements are performed with unitary probability at the present time. As it stands, the service quality of j-th customer is determined by the content volume he has obtained. At that, customer's service requirements for the distribution data center by the time t_1 can be presented as vectors with independent requirements for the customer's data volume $W_{t1} = \{W_{11},...,W_{1J}\}$. One can say that there is a problem of conforming of these requirements g_1.

Thus, the optimization problem under the requirements only for the total bandwidth capacity of the system or for the bandwidth capacity of single CDS (7), will be as follows:

$$Z = \sum_{h=1}^{H} z_h \rightarrow \min \quad \text{while} \quad V_j = W_j, \quad V_h \le C_h$$

for $j = \overline{1,Q}$, $h = \overline{1,H}$. (9)

Than it is supposed that the volume of average mutual information between the stimulus and reaction in the earlier criterion of practicability system behavior in the object domain according to the current data center theory and practice, is defined by the data volume transmitted to a customer ($I_j = V_j$).

The problem (9) can be solved using the Lagrange's method of undetermined coefficients [9] by the maximization of formula

$$\sum_{j=1}^{J} \beta_j I_j - \sum_{h=1}^{H} \gamma_h \sum_{j=1}^{J} z_{hj} \rightarrow \max , \qquad (10)$$

where γ_h, β_j are Lagrangian coefficients determined by the conditions (7), (8) and informational requirements respectively.

Recording form of the optimization problem (10) is identical to a recording form of maximization of functional informational criteria (5).

In practice, instead of the task of customer service by the distribution data center by the time t_1 is more often solved the task of continuous customer service.

In addition, there are different tasks for CDS $G = \{g_1, g_2, ..., g_h, ...\}$, according to the different requirements $W = \{W_1, W_2, ..., W_t, ...\}$, that are formulated starting from the time t_1 (the start of system operation t_0) at the points in time $T_{зад} = \{t_1, t_2, ..., t_h, ...\}$. At that, the task g_i cannot be solved during the time interval $\Delta t_i = t_i - t_{i-1}$ and requires a significant amount of time. Thereby, we should solve the task g_i or accumulate the information for its solution, which is required only by the time t_i, during the simultaneous solution of tasks g_{i-1}, g_{i-2} etc.

The current CDS tasks may be redefined subject to the goals of mega system (user community), the system state, the conditions of system operation, the solution quality of previous tasks and costs for these solutions. The strategy of redefinition may be given or the current tasks are determined by the mega system each time. If it is impossible to solve the task, it should be redefined. As a rule, if there is an excess of the system resource, the task is redefined too. In all circumstances, the stream of tasks itself G and the corresponding matrix W depends on many factors which are not always can be counted and predicted. G and W can be considered as random.

Therefore, it is necessary to develop the system with minimal average costs during the extended period of its operational time Z_Σ under conforming to the possibility of solving of task random sequence $\mathrm{G} = \{\mathrm{g}_1, \mathrm{g}_2, \ldots, \mathrm{g}_н, \ldots\}$ with requirements $\mathrm{W} = \{\mathrm{W}_1, \mathrm{W}_2, \ldots, \mathrm{W}_н, \ldots\}$:

$$Z_\Sigma \rightarrow \min \quad \text{while} \quad I_{tj} \ge W_{tj} \quad \text{for} \quad j = \overline{1,Q_t},$$

$t = \overline{1,H}$ (11)

if there are preset bandwidth limitations.

SIMPLIFICATION OF CDS COSTS

Main computational costs while solving (10) are connected with the definition of values of Lagrange's undetermined coefficients. Even the definition of corresponding Lagrange's coefficients is not a required problem solution in some applications. For example, the requests are formed in consecutive order during the analyzed time interval. They should be processed over that very interval. Every request should be processed by its own certain optimal method established by the system. The content requested by j-th customer can be distributed to him by one of the CDS or can not be distributed at all. It is impossible to wait for an optimal CDS method of action for every request performing. Because, if we wait for the interval end, there will no time for the processing of corresponding requests.

Thereby, there are tasks, having the characteristic of crucial impossibility (inexpediency) of exact optimization solving for the appropriate time in the branch of telecommunication system operation. The problem of minimization of CDS costs should be transformed into the problem of simplification of CDS costs at the same time. The simplification of CDS costs is a reducing of these costs to the accepted level (generally, this level is not as low as practicable) along with the processing of customer's requests of preset quality.

Rationalization of CDS costs during their designing is based on the following suggestions. It is suggested to work with requests as with data increments. Abstraction from the place of specific requests in the original RS is done at that. It is suggested to consider the requests independent formally. The original RS is replaced by a request space that is an intersection of the subspaces of possible values of each request separately. The latter enables designing aggregates of single request processing. These aggregates (sending j-th content from n-th DSS, "delaying" request execution, request servicing failure etc.) enable precise designing. It is suggested to regard the system as consisting from the above mentioned single request processing aggregates and a request execution manager. The latter distributes requests among the aggregates. It is suggested to associate costs not with the RS initial points, but with the request – request servicing aggregate pairs. It makes intuitive sense that the undetermined Lagrange coefficients depend on the task given to the CDS, on the CDS state, on the current request stream characteristics and the external conditions. Due to this fact it is suggested to replace the Lagrange coefficients actual values (calculated on basis of the corresponding optimization tasks limitations) with the values of empirical functions from the above mentioned parameters. Since the latter are the functions of time, the undetermined coefficients are the same. If the coefficients are known, the process of the request manager working rule calculation appears to be trivial.

Application of the last suggestion definitely resolves the optimization tasks (10) to the form that corresponds with information criterion of a purposeful system functioning (5).

Ignoring the mutual dependence of the messages enables using the short-range interaction principle, thus enabling using the corresponding increment criterion (6):

$$K_{m\delta} = \sum_{j=1}^{J} \beta_j(t)\Delta I_{xjU} - \sum_{v=1}^{V} \gamma_v(t)\Delta z_v \rightarrow \max \quad (12)$$

where ΔI_{xjU}, Δz_v is the increment of corresponding information and costs during the processing of another request by v aggregate.

Since CDS is huge, its optimization is out of question. However, it is essential that the system worked with maximum expediency. Therefore formulating a model for a huge system functioning, not the system optimization, is required. The system functioning model is sought in the category that satisfies the information criterion (12) by choosing functions for undetermined Lagrange coefficients:

$$\beta_j = \beta_j(t, W(G), S, C);$$
$$\gamma_v = \gamma_v(t, W(G), S, C) \text{ for } j = \overline{1,J},\ v = \overline{1,V}, \quad (13)$$

-

where S, C are the system state and the external conditions correspondingly.

The choice of functions (13) is performed in the predefined class (e.g. N-th degree polynomial or a piece lineal function) by the parameters selection through methods of learning with tutor.

The problem of CDS costs rationalization is set in the following way. Develop system aggregates for all the types of servicing single request using the object domain methods. Find the request manager purposeful operation model parameters by finding the functions (13), which ensure the lowest system costs of all the possible functions of the selected class:

$$Z_{\Omega\beta,\Omega\gamma} \rightarrow \min$$
$$I_{tj} \geq W_{tj},\ z_h \leq z_{h\max},$$

with $\beta_j(t, W(G), S, C) \in \Omega_\beta$, $\gamma_v(t, W(G), S, C) \in \Omega_\gamma$, where Ω_β, Ω_γ are the predefined function classes.

For solving the operation model parameters choice task (13) actual samples can be used (though in limited volumes) on the one side, and, on the other side, the experience of the experts (decision-making persons) concerning results of the request processing system operation in different conditions can be used.

EXAMPLE

All the requests entering the CDS get to the CDS manager input buffer (CDS IB). CDS bears the following costs when performing any of the above mentioned operations: $Z_n, Z_{ou}, Z_{ob}, Z_{rej}$ accordingly. Certain average amount of information can be assigned to each request.

The own resources shortage system estimation Γ is a matrix $H \times Q_{num}$ in size, where H is the number of time intervals which must be used for client request streams servicing and which is considered when planning the request servicing currently, Q_{num} is the number of tariff plans. When solving the task of request servicing for specific time matrix Γ is a column matrix consisting of Q_{num} elements. For each client request of the predefined k-th tariff plan the resources shortage system estimation γ_k is equal. The γ_k value, corresponding to the k-th tariff plan, depends only on the number of requests present at the CDS manager input buffer N_B. At that the $\gamma_k = fk(N_B)$ dependence can be found only in the first or second degree polynomial class.

The weight coefficients of the goal B system totality is a column matrix with ($\hat{Q}+1$) elements, j-th element of which β_i corresponds to the j-th client. The goal column matrix elements depend only on the I_j and W_i values. It is considered during operation that the $\beta_j(I_j, W_j)$ function is piece lineal (fig. 1).

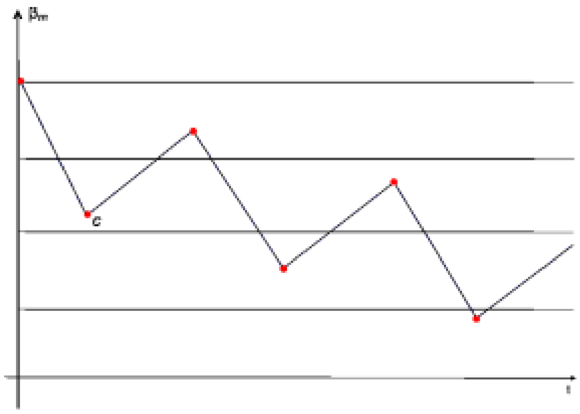

Fig-1: Diagram of CDS client serving politics formation

To define the function (13) for β_i we need to estimate 2 parameters using any methods of learning with tutor [7].

The function $\beta_j(I_j, W_j)$ parameters are constant for all the requests from clients with common tariff plan. After defining all the parameters of functions $\gamma_k = fk(N_B)$ and $\beta_j(I_j, W_j)$ the CDS message manager works in a following way. The β_i, γ_k functions values, corresponding to the amount of information in the content, which corresponds to the request, and its servicing in one of the system aggregates are defined for each request.

The presented approach can be practically implemented in the suballocated service systems or the systems working according to the "observer-subscriber" principle. The core of such a system is the complex server organization, which includes services of storing and processing the requests coming from users. Systems of such type often cannot be optimized by all the parameters simultaneously. Because of that the approach to rationalization and modernization that is suggested here is perspective. Embedding the described method to actual CDS is the purpose for future researches of the authors of this article.

CONCLUSIONS

The atemporal criterion of the functioning system (4) has been specified by replacing average mutual information with weighted sum of average particular mutual informations between every request and the system response space. Additional temporal behavior criteria (5) - (6) have been introduced. Comparative analysis of standard form of client requests problem solution, written in the form of Lagrange undetermined coefficients method, and temporal system behavior criteria shows their identity. At that the behavior model parameters values are equivalent to the corresponding Lagrange coefficients. It is suggested to determine the latter not through the optimization tasks solution, but through empirical functions from tasks given to telecommunication CDS, from CDS state, current request characteristics and external conditions. Using any methods of learning with the tutor is sufficient for defining the parameters of the latter. It is reasonable to concentrate future researches on these methods [8].

REFERENCES

1. Henri Lauer; Radio Engineering Principles. Hard Press, 2012: 340
2. Savanevych VE; Statement of the problem of synthesis of algorithms for the minimum complexity. Information processing systems: Collected Works. Kharkiv, Ukraine, 2002; 4 (20): 67– 69.
3. Laszlo Fuchs; Partially Ordered Algebraic Systems. Dover Publications, 2011: 229 p.
4. Viatcheslav B; Melas Functional Approach to Optimal Experimental Design. Springer, 2006: 333
5. Viorel Barbu; Analysis and Optimization of Differential Systems. Springer, 2003: 442
6. Solomon Kullback; Information Theory and Statistics Courier. Dover Publications, 2012: 416
7. Thomas M. Cover, Joy A. Thomas; Elements of Information Theory. John Wiley & Sons, 2012: 776
8. The CoLiTec project will offer their ideas to NASA. Available from http://www.neoastrosoft.com/the-colitec-project-will-offer-their-ideas-to-nasa_en/