UDC 004.93.14 https://doi.org/10.15407/vidbir2022.50.078

HYBRID SIMULATION MODELS FOR COMPLEX DECISION-MAKING PROBLEMS UNDER CONDITIONS OF PARTIAL UNCERTAINTY

V. O. Filatov, A. L. Yerokhin, O. V. Zolotukhin, M. S. Kudryavtseva

Kharkiv National University of Radio Electronics, Kharkiv

E-mail: andriy.yerokhin@nure.ua, oleg.zolotukhin@nure.ua

Specific features of application of hybrid simulation and control models in information systems and system support for decision-making in solving practical problems under conditions of uncertainty, vagueness, inaccuracy, stochasticity of processes of subject areas are considered. To obtain reliable data, it is necessary to use poorly formalized operational and long-term data on the state of the object of control, expert knowledge information on the application of mathematical programming methods with stochastic or fuzzy constraints, as well as many cause-and-effecr relations between processes that may be presented in the form of production rules: “condition–action”. Based on research and analysis of complex decision-making problems using hybrid simulation-control models in conditions of partial uncertainty, an estimate of their complexity in terms of practical implementations, which did not exceed the quadratic dependence on the number of operations is obtained. The peculiarities of their use in real developments are determined, which allowed us to increase the reliability of decisions in information systems, to reduce development time to 12% in the conditions of fuzzy, stochastic character of researched processes of real objects. The results that confirm their effective use in solving practical problems: an example of solving situational analysis using hybrid simulation-control models in the information-analytical decision support system, are presented.

Keywords: *hybrid simulation control model, decision-making process with fuzzy algorithmic constraints, Petri net.*

ГІБРИДНІ ІМІТАЦІЙНІ МОДЕЛІ ДЛЯ ЗАДАЧ ПРИЙНЯТТЯ РІШЕНЬ В УМОВАХ ЧАСТКОВОЇ НЕВИЗНАЧЕНОСТІ

В. О. Філатов, А. Л. Єрохін, О. В. Золотухін, М. С. Кудрявцева

Харківський національний університет радіоелектроніки, Харків

Проаналізовано особливості застосування гібридних імітаційно-керуючих моделей в інформаційних системах та системах підтримки прийняття рішень для розв’язування практичних задач за умов невизначеності, нечіткості, неточності та стохастичності процесів предметних галузей. Для отримання достовірних даних необхідні слабо формалізовані оперативні та довгострокові результати про стан об’єкта управління, знання експертів, інформація про застосування методів математичного програмування під час стохастичних або нечітких обмежень, а також слід урахувати множинні причинно-наслідкові зв’язки між процесами, які можна подати у вигляді продукційних правил: “умова–дія”. На основі дослідження та аналізу складних задач прийняття рішень з використанням гібридних імітаційно-керуючих моделей в умовах часткової невизначеності оцінено їх складність в умовах практичних реалізацій. Визначено особливості використання операцій у реальних розробках, що дало можливість підвищити достовірність рішень в інформаційних системах, скоротити час розробок до 12% в умовах нечіткого стохастичного характеру досліджуваних процесів реальних об’єктів. Наведено результати, що підтверджують їх ефективне застосування під час вирішення практичних завдань.

Ключові слова: *гібридна імітаційно-керуюча модель, процес прийняття рішень з нечіткими алгоритмічними обмеженнями, мережа Петрі.*

© V. O. Filatov, A. L. Yerokhin, O. V. Zolotukhin, M. S. Kudryavtseva, 2022

Introduction. The solution of practical problems in large-scale systems, which, first of all, should include decision-making problems, is often carried out under conditions of uncertainty, fuzziness, inaccuracy, and stochasticity of processes in subject areas. To obtain reliable data, it becomes necessary to use weakly formalized operational and long-term data on the state of the research object, expert knowledge, mathematical programming methods with stochastic or fuzzy constraints, as well as to take into account the set of cause-and-effect relations between processes that can be represented in coordinates "condition–action" [1].

These processes include those that are implemented to provide many nomenclature industries with material resources in small and medium-sized businesses. These processes are laborious, subject to the influence of a subjective factor, risk, are characterized by the presence of competitors with advanced achievements, informational, resource and other advantages. The use of existing decision-making tools in such conditions is ineffective, their possible extensions have been little studied. The solution of such problems determines the relevance of the work.

The purpose of the research is to improve the efficiency of decision-making processes when processing data in information systems of various classes based on the use of hybrid simulation and control models that ensure the functioning of the system in conditions of fuzzy, stochastic and inaccurate data and knowledge.

Statement of the problem. Let some set of decision-making processes be given $\{\Delta_\omega, \omega \in \Omega\}$, as well as hybrid simulation and control models [1, 2], methods of modeling and management of decision-making processes [3, 4] in the conditions of fuzziness, inaccuracy and stochasticity of the processes under study.

It is necessary to propose and justify:

– methods of modeling and management of decision-making processes in automated information systems for decision-making support using hybrid simulation and control models;

– assessment of the complexity of the implementation of decision-making processes based on existing approaches.

Solutions should be oriented towards modern information technologies.

Applications of hybrid simulation control models in information systems. The design and operation of automated decision support systems operating under conditions of severe resource constraints cause difficulties associated with the need to ensure the specified functional and operational properties of the system. Existing diagnostic and testing tools, as a rule, provide only monitoring of the state of the system and, to a lesser extent, control of the execution of tasks by specific applications. The functioning of the system based on serial software and hardware is largely determined by the quality of applications. The situation is complicated by the distributed nature of data processing and the significant influence of the subjective factor at the design stage.

Representation of the system model in the form of a targeted interaction of a network of frames allows the implementation of complex system tasks. However, the use of frames has known limitations in their computer implementation. In this regard, it becomes necessary to develop and use an effective apparatus for modeling and analyzing the behavioral properties of the system in the process of data processing. Considering the significant costs for the design and operation of such systems, high requirements for the quality of the solutions obtained, the problem of modeling, complex analysis and quality assurance of the behavioral properties of automated systems in specific applications is quite relevant, and currently does not have acceptable solutions.

Another well-known formalism of the system, situational analysis, which allows one to generate and study various scenarios for the operation of control systems, simulate various (including conflict and undesirable) situations, study and evaluate the consequences of decisions on the behavior of a controlled system, are Petri nets. Petri nets allow simulating the parallel and distributed operation of the algorithm. At the same time, the use of Petri nets in modeling control algorithms makes it possible to use the theory of Petri nets to analyze their semantic (behavioral) properties.

In particular, many parallel control algorithms and information processing processes are naturally modeled by graph-schemes. The graph-scheme of the algorithm is a directed graph: $G = \langle V, E \rangle$, where V is a set of peaks, E is many arcs.

In works [1–4], some features of the use of hybrid simulation-control models for solving various classes of applied problems are considered. Modeling and management of decision-making processes [3, 4] involves the implementation of the following functions:

– analysis and filtering of input data;

– formation and filling of databases;

– satisfaction of relevant requests and applications;

– formation and filling of knowledge bases;

– execution of inference algorithms;

– performing computational actions by means of mathematical programming;

– formation of the structure and state space of the E-network $\mathrm{E}^{(m)}$ [4] as dynamic objects;

– formation of conditions for performing network transitions;

– formation of a set of allowed connections on the network and exclusion from consideration of a set of forbidden transitions and connections;

– determination on the model of a set of alternatives that meet the existing criteria on a set of constraints of the subject area;

– formation of recommendations for decision-making.

Proceeding from the peculiarities of using hybrid simulation-control models in practical applications, we will consider some aspects of the proposed approach.

Formation and filling of the database. To work with objects and data characteristic of a given subject area, it is advisable to use standard database management systems (DBMS). This is due to the fact that an enterprise usually has some kind of DBMS that is used to collect and process data related to its core business. It may contain information about suppliers, the range of goods and their characteristics, weather conditions, etc. This data can be used to solve the problem of making decisions using tools for modeling production processes and decision-making processes.

Formation and filling of the knowledge base. A number of modern enterprises usually have some versions of decision-making systems, expert systems. In this case, it is advisable to adapt existing solutions to apply the considered approach. The knowledge base is filled with expert knowledge using the knowledge acquisition subsystem in a formalized language, taking into account the agreements adopted in the system. Usually, the knowledge base is filled out through the input and editing of the parameters of computational positions, and the definition of membership functions for linguistic variables.

A production or rule-based model allows knowledge to be represented in the form of sentences such as “if (condition), then (action)”. A “condition” (antecedent) means a certain sample sentence by which a search is carried out in a knowledge base, and an “action” (consequent) means actions performed upon a successful search outcome (they can be intermediate, acting further as conditions, terminal or target shutting down the system) [5].

An important stage is the definition and specification of the type and parameters of membership functions in fuzzy rules. For this purpose, it is advisable to use analytical representations of the type [6]. For fuzzy inference, algorithms proposed by Mamdani are applicable [7]. The result of the inference procedures execution determines the marking of the computational positions of the E-network, as well as the value of algorithmic constraints when solving mathematical programming problems.

Implementation of computational solutions by means of mathematical programming. According to the method proposed in [4] for solving the problem of optimizing decision-making processes with fuzzy algorithmic constraints, we will formulate the goals and objectives of implementing the decision-making processes:

– we determine the initial data, including those that are presented on a set of linguistic variables;

– we form the goal function for the task, which is defined in the work as a linear programming task;

– we form constraints of two types: for the classical approach and fuzzy algorithmic;

– for fuzzy algorithmic constraints, we form a knowledge base based on rules, linguistic variables and, thus, the knowledge base is configured;

– taking into account the complexity, nature and characteristics of the problem, the classical method of mathematical linear programming is determined and substantiated; the linear programming problem is solved by the classical method without taking into account algorithmic fuzzy constraints.

If, taking into account the fuzzy algorithmic constraints, the results are satisfactory, then we mark the computational positions of the E-network. Otherwise, we carry out the reverse fuzzy logical conclusion of Mamdani and specify the constraints on the objective function. The result of the subsequent solution gives us new conditions for marking the computational positions of the E-network.

Formation of the structure in the state space of the network $\mathrm{E}^{(\mathrm{m})}$ is determined on the model of alternatives from a set of possible solutions that satisfy the formulated criteria on a set of constraints of the subject area and can be implemented in the form of a set of targeted actions.

When developing the method [3], the concepts of dynamic objects [3, 8] were introduced and used. Dynamic objects define the vertices, state space and connections of the E-network, formats and attributes are defined for them, which can be modified during the modeling process or, if necessary, taking into account the peculiarities of the subject area. For some computing positions, stochastic streaming networks are additionally introduced, which complement dynamic objects of the “computing position” type $M(RcN - C_k)$. This allows us, in fact, to consider the E-network at the level of interacting dynamic objects of a given subject area.

Features of the use of hybrid simulation and control models in decision-making problems. In practical implementations of systems operating in real time, it is important to estimate the required computing and time resources for solving applied problems, taking into account the complexity of solving problems depending on the dimension of the hybrid simulation-control model.

In this regard, we will consider the problem of implementing decision-making processes based on the proposed hybrid models, which include:

– data at the input of the system; $\{R_i\},\ i \in M$ – a lot of data about objects;

– many rules if / then (D) – deterministic rules, if / then (Pr) – probabilistic rules, if / then (F) – fuzzy rules;

– means of mathematical programming with fuzzy algorithmic constraints;

– network models with a managed structure;

– a decision-maker.

The complexity of identifying data at the input of the system and a set of data about the objects of research $\{R_i\},\ i \in M$ is determined by means of the applied DBMS [9], used by software and hardware. The complexity of decision-making processes by means of mathematical programming is determined by existing software implementations.

In this regard, we will consider the assessment of the complexity of fuzzy inference procedures and procedures for performing the E-network.

Assessment of the complexity of fuzzy inference procedures based on production rules. In some special cases, questions of complexity of research for network models were presented in a number of works, in particular, in [10]. Note that inference based on fuzzy production rules is used both in inference procedures and in problems of forming fuzzy algorithmic constraints of decision-making procedures by means of linear programming of the model [4].

Thus, in [4], a special case was considered under the assumption that there were a limited number of rules that define algorithmic constraints. This somewhat reduces the universality of the results obtained and requires additional research.

Let there be many fuzzy rules like if / then ... As you know, a direct fuzzy inference can be represented in the form [7]

$$\begin{array}{l} \text{if } x \text{ is } \mu_1(x) \text{ then } y \text{ is } \mu_2(y) \\ \quad x \text{ is } \mu_1'(x) \\ \hline \quad y \text{ is } \mu_2'(y) \end{array}, \tag{1}$$

where $\mu_1(x)$, $\mu_2(y)$ are the corresponding membership functions, $\mu_1'(x)$ is the known value of the membership function, $\mu_2'(y)$ is the required value of the membership function.

The calculation of the values of the membership functions for (1) are calculated as folows:

$$\mu_2'(y) = \vee_x [\mu_1'(x) \wedge \mu(x, y)], \tag{2}$$

where $\mu(x, y)$ is the membership function of a known relation $\tilde{R}(x, y)$ ...

Statement 1. If some set of fuzzy rules is given W (1), then the upper bound for the complexity of solution (2) for $|W| = 1$ is equal to

$$C = k_1 \rho^2, \tag{3}$$

where k_1 is some coefficient determined by the features of the algorithm implementations, ρ is the number of discretes when displaying some membership function $\mu(x)$.

Indeed, solution (2) includes a set of minimax operations for the vector $\mu_1'(x)$ dimensions ρ and the membership function of the relation $\mu(x, y)$, which is represented by a matrix of dimensions ρ^2 ... Then the number of operations for comparing values and finding the resulting vector $\mu_2'(y)$ is ρ^2 ... Coefficient k_1 is scalable.

Statement 2. If some set of fuzzy rules is given W (4.10), then the upper bound for the complexity of solution (2) for $|W| > 1$ is equal to

$$C = k_2 \rho^3 |W|, \tag{4}$$

where k_2 is some coefficient, determined by the peculiarities of the implementation of the algorithm, ρ is the number of discretes when displaying some membership function $\mu(x)$...

Indeed, solution (2) includes a set of minimax operations for each pair of matrices of dimensions ρ^2 membership functions of relations $\mu(x,y)$, $\mu(y,z)$ some neighboring rules. Then the number of operations for comparing values and finding the resulting ratio matrix $\mu(x,z)$ for these rules is ρ^3... Given that the rules in this case $|W|$, then the validity of (4) is obvious. Coefficient k_2 is scalable.

Consequence. Then, taking into account the results of statements 1, 2, the value of the upper bound for the complexity of the algorithm is proportional to

$$C' = k_1\rho^2 + k_2\rho^3 |W|, \tag{5}$$

Comment. The complexity of defuzzification procedures [9] was not taken into account in (5), which is a fairly well-studied procedure and does not require additional research.

For (5), the problem arises of choosing the sampling frequency ω_0, i.e. select value ρ, as well as identifying stationary points for which it is necessary to set the value of the membership function. As follows from [6], analytical dependences of membership functions can be represented by a superposition of linear functions (straight lines, triangles, trapezoids, etc.), curves of the second and higher (Gaussians, sigmoids) orders. For the first case, it is enough to display characteristic points when constructing relations [9]: mutual intersection and intersection with coordinate axes. For the second case, it is necessary to display all the characteristic sections of the function, including the stationary points of finding the extrema and inflection points.

Let some membership function be given $\mu(x)$... Then, at least some of the stationary points can be determined by finding the extrema of the function $\mu(x)$ and inflection points by calculating derivatives of the first or higher order and examining their roots. The validity of the position is obvious if we take into account that, excluding the trivial case, $\mu(x)=1$, membership functions [6, 7] usually have a clearly defined extremum (maximum), and, for example, Gaussians are characterized by the existence of inflection points.

The very same sampling rate ω_0 is largely determined empirically, and as follows from the above, for some special cases it can be determined from the appropriateness of the number of discrete $\rho \le 10$.

Assessment of the complexity of implementation of procedures for performing modeling processes on the E-network. Consider a network for which $|P|=m$, $|T|=n$... Then, to assess the complexity of the network execution, we must take into account that for objects of the "position" and "transition" types, it is necessary to perform a certain number of operations, which are proportional to the number of nonzero components of the incidence function.

Statement 3. If there is an E-network, then the number of actions during the execution of the network is proportional to the number of nonzero components of the incidence function $F = I \cup O$. E-network and can be represented as

$$C'' = k_0 mn, \tag{6}$$

where k_0 is determined by the sparseness of the incidence function, i.e. fractions of nonzero components of the function F...

The validity of the statement directly follows from the conditions for performing transitions, marking positions during the operation of the E-network.

Consequence. Taking into account (5) and (6), the required complexity

$$C_{\Sigma} = C' + C'', \tag{7}$$

or

$$C_{\Sigma} = k_0 mn + k_1 \rho^2 + k_2 \rho^3 \, |W|, \tag{8}$$

The theoretical results are confirmed on test examples in the implementation of decision-making procedures for complex production processes, which allows the effective software implementation. In this case, membership functions of both the first and the second types were used. The time spent on the implementation of the proposed version of the algorithm does not exceed theoretical dependences, which is confirmed on real objects. This makes it possible to use it in automated real-time systems.

Example. Consider the tasks most frequently encountered in decision support systems.

Objective 1. Resource type A_1 is provided with information support and stored in database files, for example .dbf, $B_{11}, B_{12}, ...B_{1n}, n \in N$, where N is the number of databases. The search condition for the considered fragment of the decision-making system is formulated as follows: determine a given amount of a resource of type $A_1 = A_1^*$, on the set $\{B_{11}, B_{12}, ...B_{1n}\}, n \in N$ ... The search for a solution is carried out in two types of actions – at the first step SEARCH in the database B_1 if the condition is not met, the DATABASES JOIN is performed $B_1 + B_2$ and the search is repeated [11, 12].

The procedure for searching and merging databases can be performed in terms of structured SQL query language using the SELECT and INSERT options.

For example:

Select sum ([A1]), from b1.dbf;

Insert into b1… Select b2… .From b2.

We represent a fragment of the problem considered above in the form of a modified predicate Petri net S (Fig. 1). Many positions $\{p_0, p_1\}$ interpret, respectively, the input and output conditions for the execution of the action, which is represented on the network by the transition t_0, many positions $\{p_1, p_2\}$ interpret, respectively, the input and output conditions for the execution of the action t_1... Initial marking vector $M_0 = (1, 0, 0)$ defines the initial state space of the network.

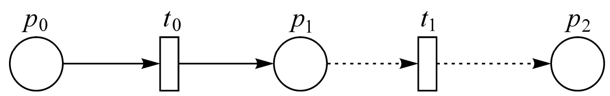

Fig. 1. Petri net *S*.

Description of the main constraints are shown in Fig. 2, where $V_{p_k{}^I}$ is the input condition for the transition $t_1 \in T$, $k \in K$, $V_{p_k{}^o}$ is the exit condition of the transition t_1. If the precondition $V_{p_k{}^I}$ performing the transition t_1 is not satisfied: $V_{p_k{}^I} \neq V_{t_k}$, then for this transition $t_1 \in T$, $k \in K$ a slave network S' is generated where position

p_1 is a metaposition. Fig. 2 shows an extended network $S \cup S'$... Initial marking vector $M_0 = (1,0,0,0,0)$ defines the initial state space of the network.

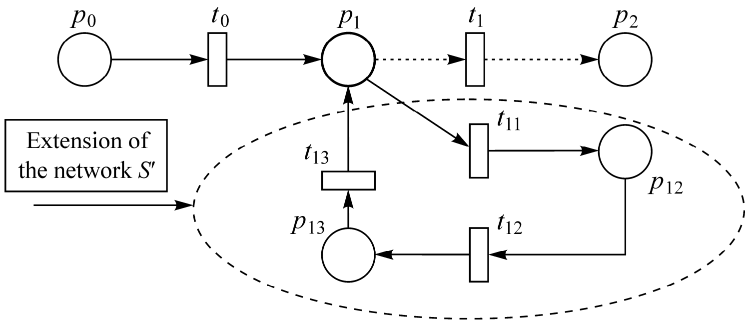

Fig. 2. Petri net S with extension S'.

In order to solve the problem, a network S'... is formed. Position marking p_{13} and performing the transition t_{13} marks metaposition p_1 which triggers the execution of the network S... If the desired solution is not feasible by means $S \cup S'$, then the search for solutions is done from the set $\{B_{11}, B_{12}, ...B_{1n}\}$, $n \in N$ by combining databases and expanding the network S'... Consider a modification of Problem 1 and the decision support procedure.

Objective 2. Resource type A_1 is provided with information support and is stored in files of two databases, for example .dbf: B_{11}, B_{12}, $...B_{1n_1}$, $n_1 \in N_1$, $N_1 \subset N$; B_{21}, B_{22}, $...B_{2n_2}$, $n_2 \in N_2$, $N_2 \subset N$, where N is the number of databases. The search condition for the considered fragment of the decision-making system, by analogy with the condition of problem 1, is formulated as follows: determine the given amount of resource of the type $A_1 = A_1^*$ on sets $\{B_{11}, B_{12}, ...B_{1n_1}\}$, $n_1 \in N_1$ and $\{B_{21}, B_{22}, ... B_{2n_2}\}$, $n_2 \in N_2$.

The search for a solution is carried out for two database systems by two types of actions – at the first step SEARCH in the database B_1 if the condition is not met, the DATABASES JOIN is performed $B_1 + B_2$ and the search is repeated. The procedure for searching and merging databases can be performed in terms of structured SQL query language using the SELECT and INSERT options. The desired solution is determined by the optimal choice from the existing design solutions that satisfy the condition of providing a resource of the type $A_1 = A_1^*$ and some additional restrictions.

CONCLUSIONS

The features of solving practical problems in large-scale systems, which, first of all, should include decision-making problems, often carried out in conditions of uncertainty, fuzziness, inaccuracy, and stochasticity of processes in subject areas are discussed. To obtain reliable data, it becomes necessary to use weakly formalized operational and long-term data on the state of the control object, the knowledge of experts, the mathematical programming methods with stochastic or fuzzy constraints, as well as to take into account the set of cause-and-effect relations between processes, which can be presented in the form of production processes. rules: "condition–action".

Based on the study and analysis of complex decision-making problems using hybrid simulation and control models under conditions of partial uncertainty, an estimate of their complexity in the conditions of practical implementations was obtained, which did not exceed the quadratic dependence on the number of operations performed. The features of their use in real developments were determined, which made it possible to increase the reliability of decisions made in information systems, to reduce the development time to 12% in the conditions of a fuzzy, stochastic nature of the investigated processes of real objects. An example of solving the problem of situational analysis using hybrid simulation-control models in an information-analytical decision support system is given.

1. Kucherenko, V.E. Models for the analysis of data processing processes in automated systems *Radioelectronics and Informatics*, **2004**, 3, 103–107.
2. Filatov, V.A.; Kucherenko, V.E. Odata processing models based on the hierarchy of high-level networks. *Radioelektronika. Informatics. Management*, **2004**, 1, 95–100.
3. Kucherenko, V.E. Methods for analyzing the processes of modeling and processing of data in folding systems. *Artificial intelligence,* **2004**, 4, 126–134. (in Ukrainian)
4. Filatov, V.A.; Kucherenko, V.E. On methods of optimization of decision-making processes in automated control systems under algorithmic constraints. *Applied Radio Electronics*, **2005**, 4, 298–303.
5. Knowledge base of intelligent systems. Gavrilova, T.A.; Khoroshevsky, V.F. Eds. Peter, 2000.
6. Kucherenko, V.E. Models and methods of data presentation in information and analytical systems. In Education, science, production and management in the 21st century: Sat. Proceedings of the International Scientific Conf. in 4 volumes: Krakhta, V.B. Ed. Stary Oskol, Russian Federation, 2004, pp. 307–310.
7. Connolly, T.; Begg, K. Database Systems: A Practical Approach to Design, Implementation, and Management, Pearson, 2014.
8. Kucherenko, V.E. Nutrition of automation of processes to take decisions in algorithmic intercourse. In 9th International Youth Forum "Radioelectronics and young people in the 21st century": Kharkiv, KHNURE, 2005, p. 437.
9. Tsoukalas, L.H.; Uhrig, R.E. Fuzzy and Neural Approaches in Engineering. John Wiley & Sons. Inc., 1997.
10. Kaufmann, A. Introduction to the Theory of Fuzzy Subsets. Academic Pr., 1975.
11. Filatov, V.; Yerokhin, A.; Zolotukhin, O.; Kudryavtseva, M. Personalized Adaptation of Learning Environments. In 8th International Conf. on Advanced Optoelectronics and Lasers CAOL2019. September 06–08, 2019, Sozopol, Bulgaria. https://doi.org/10.1109/CAOL46282.2019.9019525
12. Filatov, V.; Semenets, V.; Zolotukhin, O. Synthesis of semantic model of subject area at integration of relational databases. In International Conf. on Advanced Optoelectronics and Lasers. Sozopol, Bulgaria, September 06–08, 2019, pp. 598–601. https://doi.org/10.1109/CAOL46282.2019.9019532

Received 18.01.2022