ISSN 1727–6209

TERNOPIL NATIONAL ECONOMIC UNIVERSITY
RESEARCH INSTITUTE OF INTELLIGENT COMPUTER SYSTEMS
IN COOPERATION WITH V.M. GLUSHKOV INSTITUTE FOR
CYBERNETICS, NATIONAL ACADEMY OF SCIENCES, UKRAINE

ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
НАУКОВО-ДОСЛІДНИЙ ІНСТИТУТ ІНТЕЛЕКТУАЛЬНИХ
КОМП'ЮТЕРНИХ СИСТЕМ
У СПІВПРАЦІ З ІНСТИТУТОМ КІБЕРНЕТИКИ ІМ. В.М. ГЛУШКОВА,
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

International Journal of

Computing

Published since 2002

September 2012, Volume 11, Issue 3

Міжнародний науково-технічний журнал

Комп'ютинг

Видається з 2002 року

Вересень 2012, Том 11, Випуск 3

Постановою президії ВАК України № 1-05/3 від 14 квітня 2010 року науково-технічний журнал "Комп'ютинг" віднесено до переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата технічних наук

Тернопіль
ТНЕУ
2012

International Journal of Computing

September 2012, Vol. 11, Issue 3

**Research Institute of Intelligent Computer Systems,
Ternopil National Economic University**

**Registered by the Ministry of Justice of Ukraine.
Registration certificate of printed mass media – KB #17050-5820PR, 15.07.2010.**

It's published under resolution of the TNEU Scientific Council, protocol # 8, May 30, 2012

Editorial Board

Address of the Editorial Board

Research Institute of Intelligent Computer Systems
Ternopil National Economic University
3, Peremoga Square
Ternopil, 46020, Ukraine

Phone: +380 (352) 47-5050 ext. 12234
Fax: +380 (352) 47-5053 (24 hours)
computing@computingonline.net
www.computingonline.net

Міжнародний журнал “Комп’ютинг”

Вересень 2012, том 11, випуск 3

Науково-дослідний інститут інтелектуальних комп’ютерних систем
Тернопільський національний економічний університет

Зареєстрований Міністерством юстиції України. Свідоцтво про державну реєстрацію друкованого засобу масової інформації – серія КВ №17050-5820ПР від 15.07.2010.

Друкується за постановою вченої ради ТНЕУ, протокол № 8 від 30 травня 2012 року

Редакційна колегія

Адреса редакції :

НДІ інтелектуальних комп’ютерних систем
Тернопільський національний економічний університет
площа Перемоги, 3
Тернопіль, 46020, Україна

Тел.: 0 (352) 47-50-50 внутр. 12234
Факс: 0 (352) 47-50-53
computing@computingonline.net
www.computingonline.net

CONTENTS

A. Kumar, V. Goyal
EARCT – Environment for Automated Rank Based Continuous Agent Testing 180

A. Mykhailiuk, O. Mykhailiuk, O. Pylypchuk, V. Tarasenko
A Creation of the Linguistic Ontology Based on a Structured Electronic Encyclopedic Resource 191

M. Polyakova, V. Krylov, N. Volkova
The Method of Wavelets Construction by Transformation of a Graph of Power Function for Edge Detection 203

D.V. Patil, R.S. Bichkar
Integrated Effect of Data Cleaning and Sampling on Decision Tree Learning of Large Data Sets 215

A. Kabysh, V. Golovko
General Model for Organizing Interactions in Multi-Agent Systems 224

A. Voronych, Y. Nykolaychuk, V. Hladyuk
Formation and Processing of Entropy-Manipulated Correcting Signal Codes in Wireless Sensor Networks 234

S. Antoschuk, D. Maevsky
Structural Synthesis of Dynamic Information Systems 245

S.P.Khandait, R.C.Thool, P.D.Khandait
ANFIS and Neural Network Based Facial Expression Recognition Using Curvelet Features 255

B.Y. Volochiy, L.D. Ozirkovskyy, I.W. Kulyk
The Method of Building of Behavior Model of Non-Markov Complex Systems as a Graph of States and Transitions 262

Y. Semchyshyn, I. Kulpa, I. Kolosovskyi, O. Hrechnikov, P. Hayda
Developing Semantic Network Management System 272

V.G. Deibuk, I.M. Yuriychuk, R.I. Yuriychuk
Spin Model of Full Summator on Peres Gates 282

J. Drozd, A. Drozd, J. Sulima
Natural Resources and Their Use for Checkability Increasing the Digital Components of Safety-Critical Systems 293

V.A. Luzhetsky, Y.V. Baryshev
The Generalized Construction of pseudonondeterministic hashing 302

Abstracts 309

Information for Papers Submission to Journal 318

CONTENTS / ЗМІСТ / СОДЕРЖАНИЕ

A. Kumar, V. Goyal
EARCT – Environment for Automated Rank Based Continuous Agent Testing 180

А. Михайлюк, О. Михайлюк, О. Пилипчук, В. Тарасенко
Формування лінгвістичної онтології на базі структурованого електронного енциклопедичного ресурсу 191

М. Полякова, В. Крылов, Н. Волкова
Метод построения улучшенных вейвлетов путем преобразования графика степенной функции для задачи контурной сегментации изображений 203

D.V. Patil, R.S. Bichkar
Integrated Effect of Data Cleaning and Sampling on Decision Tree Learning of Large Data Sets 215

А. Кабыш, В. Головко
Обобщенная модель организации взаимодействий в мультиагентной системе 224

А. Воронич, Я. Николайчук, В. Гладюк
Формування та опрацювання ентропійно-маніпульованих сигнальних коректуючих кодів у безпровідних сенсорних мережах 234

С. Антощук, Д. Маєвський
Структурний синтез динамічних інформаційних систем 245

S.P.Khandait, R.C.Thool, P.D.Khandait
ANFIS and Neural Network Based Facial Expression Recognition Using Curvelet Features 255

Б. Волочій, Л. Озірковський, І. Кулик
Метод побудови моделей поведінки складних систем немарковського типу у вигляді графа станів і переходів 262

Ю. Семчишин, І. Кульпа, І. Колосовський, О. Гречніков, П. Гайда
Розробка системи керування семантичними мережами 272

В.Г. Дейбук, І.М. Юрійчук, Р.І. Юрійчук
Спінова модель повного однорозрядного суматора на елементах Переса 282

Ю.В. Дрозд, А.В. Дрозд, Ю.Ю. Сулима
Естественные ресурсы и их использование для повышения контролепригодности цифровых компонентов систем критического применения 293

V.A. Luzhetsky, Y.V. Baryshev
The Generalized Construction of pseudonondeterministic hashing 302

Abstracts / Резюме 309

Information for Papers Submission to Journal / Інформація для оформлення статей до журналу / Информация для оформления статей в журнал 318

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

EARCT- ENVIRONMENT FOR AUTOMATED RANK BASED CONTINUOUS AGENT TESTING

Ashok Kumar [1], Vinay Goyal [2]

[1] Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India
[2] Panipat Institute of Engineering & Technology, Samalkha, Panipat, India
vinaykuk@gmail.com

Abstract: *Testing MAS (Multi-agent System) is a challenging task because these systems are distributed, complex and autonomous in nature. Agents exist in an open environment having their own locus of control and they require context awareness. So due to these agent's characteristics, testing MAS system using existing testing techniques becomes a very tedious job. Agents also pose problems regarding message communication and semantic interoperability, as well as synchronization with other agents existing in the environment. All these features are known to be hard not only to design and to code, but also to test. In this paper we will propose a unique environment EARCT to test MAS keeping in mind the essential software engineering paradigms such as effort consumed, errors revealed etc.*

Keywords: *Agents Testing Autonomous Ranking.*

1. INTRODUCTION

Current research and development on agent oriented technology mainly put emphasis on designing architecture, formalizing protocols, designing frameworks etc.Very limited research work has been carried out on testing multi agent system [1,2,3]. The agent-oriented paradigm is considered a natural extension to the object-oriented (OO) paradigm, but agents are different from objects in many ways [4,5]. Although there are well-defined OO testing techniques, agent-oriented development has neither a standard development process nor a standard testing technique. Since, in MAS (Multi agent system) there are several agents existing in an environment [6,7] as distributed components, which are proactive and autonomous, so it is possible that same inputs can provide different outputs on different execution. The problems posed by the agents and Multi agent system testing are well recognized [8], and some of the more significant ones are discussed in next section.

2. BACKGROUND

As noted above, this section identifies and describes some of the major problems posed by Multi-Agent Systems.

Autonomous and social nature: Agents are autonomous in nature and due to their social capabilities- they cooperate with other agents present in the environment. MAS testing tools must have a comprehensive view over all distributed agents in addition to local knowledge about individual agents, in order to check whether the whole system operate accordingly to the specifications or not. Moreover, it may be possible that a single agent ran successfully and correctly as a stand-alone entity but incorrectly in a community or *vice versa*.

Complexity: As agents are autonomous and are run concurrently in a distributed environment, it becomes very difficult to determine the exact boundary of a test case. Distributed and concurrent environments often pose a challenge before testing team.

Agent Communications: Agents communicate with each other via message passing and not by method invocation as in object oriented technology, so existing object oriented technology testing techniques are not applicable for agent based testing.

Non-Deterministic nature: Agent's nature is non-deterministic in nature because it is not possible to determine in prior all possible interactions of an agent during its execution.

Irreproducibility effect: Due to an agent's pro-active and autonomous nature, we cannot guarantee that two executions of the systems will lead to the same output state, even if the same inputs are used because agents can modify their knowledge between any two executions. As a result, looking for a specific error can be difficult if it cannot be replicated [9].

Amount of data: MAS can constitute numerous agents, as each agent processes its own data. To test such a huge amount of data is itself a very challenging task for the testing team. Also agents operate asynchronously and in a parallel manner which is also challenging for a testing team.

Agent Ranking: There is no way a testing team can assess out the importance or rank of a particular agent. It might be possible that a particular agent is acting as a core element of the system, and another as an outside scaled agent with less knowledge about its internal state and behavior. Agents with top ranking (most important agents) must be given extra attention and resources during testing and third party agents or agents with lower ranking may be tested with limited resources. This will result in more effective resource utilization during testing phase and will improve proper operation of the core functionalities for the whole system.

Agent Inter-Dependency: As agents are social in nature and are autonomous also, they can interact with other agents present in the environment to fulfill their goal. Because of this, there exist many execution paths which can be followed by an agent. It is difficult to test each and every path for correctness, completeness and consistency.

Black Box MAS: MAS can also be considered as «black-box»; that is, they may provide very little or some time no observational primitives to the outside world, resulting in limited access to the internal agents' state, their expected behavior and knowledge. This kind of MAS could be quite difficult to test, in that the test result (PASS or FAIL) may be hard to assess.

3. RELATED WORK

There is a very brief literature available on the testing of Agent. In fact, in recent times, few automated techniques are proposed to test agents and test strategies soon. Also work agent testing was done at various levels of testing such as unit-level (agent level), the level of integration (MAS) and system level. In fact, there is very little that AOSE explicitly define the test phase and test mechanisms. Zhang [10] introduced a framework for Model Based Testing using design patterns of the Prometheus agent development methodology. This framework focuses on testing agent plans (units) and mechanisms to generate test cases and appropriate to determine the order in which units are to be tested. Ekinci [11] stated that the goals of agents are the smallest testable units and MAS proposed to test these units through the test objectives. Each test objective is conceptually divided into three sub-goals: Installation (System Preparation), the objective of the test (perform actions on the objective), and purpose statement (check the satisfaction of goals). The first and last objective is preparing pre-conditions and post-conditions checked by testing the objective under test, respectively. In addition, they introduce a testing tool, called as SEAUnit that provides the infrastructure to support the proposed approach. Agile PASSI [12] proposes a framework to support trials of single agents. They develop a test suite specifically for the verification agent. Test plans are prepared before the coding phase according to the specifications and the tool is also capable of generating agents AgentFactory can also generate driver and stub to accelerate the testing of a specific agent. Lam and Barber [13] proposed a semi-automated process for understanding the behavior of software agents. The approach mimics what a human user (can be a tester) in the understanding of the software: building and refining knowledge base of agent behavior, and use it to verify and explain the behavior of agents at runtime. Nunez [14] introduced a formal framework for specifying the behavior of autonomous e-commerce agents. Desired behaviors of the agents being tested are presented using a new formalism, called state machine utility that embodies the users preferences in their states. Two test methods have been proposed to check if an implementation of a specified agent behaves as expected (i.e. compliance testing). In their approach to asset tests, they used for each test agent test (special agent) who makes the formal specification of the agent to facilitate reaching a specific state. The trace of the operational agent is then compared to the specification in order to detect defects. Moreover, the authors also proposed using the passive test in which the agents being tested, were only observed, not stimulated as in the active test. Invalid traces, if any, are then identified through formal specifications of agents. Coelho [15] proposed a framework for unit testing of MAS based on the use of simulated agents, which simulate real agents to communicate with the test agents were implemented manually, each corresponding to an agent role. Sharing the inspiration of JUnit [16] with Coelho [15], Tiryaki [17] proposed a test-driven development approach that supported MAS iterative and incremental construction MAS. A testing framework called SUnit, which was built above and JUnit Seagent [18], was developed to support the process. The framework allows writing tests for the agents' behavior and interactions between agents. Gomez-Sanz [19] introduces advance in testing and debugging methodology. In fact INGENIAS [5], the meta-model INGENIAS was extended with concepts for defining tests to integrate the reporting of the test, i.e. testing and test packets. Work has also provided facilities for access to mental states of individual agents to check them at runtime. Houhamdi [20] introduced an approach to derive test

suite for testing agent that is goal-oriented requirements analysis artifact that the basic elements for developing test cases. The proposed process has been illustrated with respect to the Tropos development process. It provides systematic guidance for generating test suites, the detailed design agent. These test suites, on the one hand, can be used to refine the analysis of objectives and to detect problems early in the development process. On the other hand, they are subsequently executed to test the objectives from which they were established. Agile [21] defines a test phase based testing framework JUnit [22]. To use this tool, designed for testing OO, MAS test in context, they need to implement a platform agent sequential, strictly used for testing, which simulates asynchronous message passing. The ACLAnalyser [23] tool runs on the JADE [24] platform, it intercepts all the messages exchanged between agents and stores them in a relational database. This approach exploits the clustering techniques to construct graphs of interaction of agents that support the detection of failed communication between agents that are expected to interact, configurations execution asymmetric and the data exchanged between agents. Padgham [25] uses design artifacts (e.g., interaction protocols and agent design specification) to provide automatic identification of the source of errors detected during execution. A debug agent is added to the central MAS to monitor the conversations of agents. It receives a carbon copy of every communication between agents, in a specific conversation. The interaction protocol specifications to the call are taken and analyzed to detect erroneous conditions automatically. Rodrigues [26] proposed to exploit the social conventions, norms, rules, that prescribe the authorizations, obligations, and / or ban agents in MAS open to an integration test. Information available in the specifications of these agreements give rise to a number of types of assertions, such as time to live the role, cardinality, and so on. During the test run of a special agent, called agent will report to observe a events and messages to generate analysis results thereafter. Nguyen [27] propose to use the ontology (s) extracted from MAS under test and a set of OCL constraints, which act as a test oracle. Having as an input a representation of the ontology (s) used, the idea is to build an agent capable of delivering messages whose content is inspired by these ontologies. The resulting behavior is believed to be correct by using the input set of OCL constraints: if the message satisfy the constraints, the message is correct, this procedure is supported by ECAT, a software tool. Houhamdi and Athamena [20] has introduced a new approach to goal-oriented software testing integration. They propose an approach to derive a test suite for integration testing, that takes goal-oriented artifact needs analysis to derive test cases. They discussed how to derive test suites for testing the integration of architectural design and detailed system objectives. These test suites can be used to observe the emergent properties, resulting from potential agents and make sure that a group of agents and contextual resources function properly together. This approach defines a structured test and overall integration and junction sequence of processes for engineering software agents by providing a systematic way to derive test cases from the analysis objective. Houhamdi and Athamena [28] introduced an approach to derive a test suite to test the system that is goal-oriented. Requirement analysis artifact that are the basic elements for developing test cases. The proposed process has been illustrated with respect to the Tropos development process, provides a systematic guidance for generating test suites for modeling artifacts, produced with the development process. They discussed how to derive test suites to test the system and delay the requirement to architectural design. These test suites, on the one hand, can be used to refine the analysis of objectives and to detect problems early in the development process. On the other hand, they are subsequently executed to test the objectives from which they were established.

4. AUTOMATED CONTINOUS TESTING

As observed from the issues highlighted above in the section 2, manual testing of MAS is a troublesome job. Testing MAS can be effectively done by automating the testing job and the process should be done in a continuous fashion. The continuous streaming of testing process is required due to uncertain and complex nature of MAS along with huge data to be tested. The proposed automated continuous testing framework will extensively test the system covering maximum units. The proposed framework complements the manual test case with each other rather than replacing the manual test cases. Due to continuous automated testing, various conditions which are responsible for errors can be reveled which are otherwise hard to reproduce manually. The other issue is to figure out the importance/rank of the module under consideration. The higher rank modules must be tested exhaustively while on the other hand, modules with lower rank of importance do not need exhaustive testing. This differentiation is done to make optimal use of available resources along with making the objective to produce an error free system. We will introduce two different strategies of testing, named as Rank Oriented Random Testing and Rank Oriented Exhaustive Testing. Depending upon the rank of a particular module, the appropriate testing strategies can be applied. The idea of introducing

rank based testing is to save time, cost and other resources incurred in testing process especially during testing of a multi agent system. The basic necessity of the testing process is to figure out the maximum number of errors in the system. The other consideration is of the agent interdependency. As agents are interacting with other agents in a loosely coupled environment, so establishing a valid inter-dependency relationship between the chains of interacting agents to fulfill the goal is really required. This can be achieved naturally because agents interact with each other by message passing. The valid states of the caller and cal lee agents in MAS can be checked during testing process in order to test the agent dependency. As the nature of MAS can change over time, may be between two successive executions, due to their learning capabilities, a single test case execution is not useful to reveal all the errors. The test time and number of test suites can be increased by using autonomous property of the agents. Automated continuous agents can continuously test the MAS units without the need of any human intervention and can proceed on their own without human attention. Continuous testing of MAS requires that the tester agent has the capability to develop existing test suites and to generate new test suites, with an aim of exercising and stressing the application as much as possible. The final goal is to reveal yet unknown faults.

5. EARCT (ENVIRONMENT FOR AUTOMATED RANK BASED CONTINOUS TESTING)

We propose a framework for automated rank based continuous testing named EARCT (Environment for Automated Rank based Continuous Testing). We propose three main components: The Autonomous tester agent, the observer agent, the ranking agents and two testing strategies: the rank based random test case generation and rank based progressive mutant test case generation.

The Autonomous Tester Agent: The dedicated autonomous tester agent will continuously test the MAS by means of message passing. It will continuously interact with the agents under test (AUT) by sending message to other agents, simulating the behavior of the caller or cal lee agents analogous to writing stubs and drivers in top down and bottom up testing respectively. The process will be autonomous and will execute in background, without any human intervention and continuous fashion in order to achieve the basic goal of revealing maximum faults in the system. The test suites of tester agents will contain the dummy messages to be sent to AUT. These messages can be extracted from the goal diagram using TROPOS.

The Observer Agent: The observer agent works like a watch dog over the autonomous tester agent and AUT. It observes the communication pattern between both of them. The observer agent will have knowledge about the pre and post state conditions of the AUT, error conditions, crash situations and even deadlock conditions. In case of any of the above mentioned case occurred the responsibility lies with the observer agent to inform about the problem(s) to the testing team. The testing team can then figure out the faults in the system. In figure 1 shown, the observer agent will be working as a master for the local observer agents. This is very much required because MAS works in heterogeneous environment. It is always the case that an agent in one environment will interact with the other agents in some other environment, and it is not necessary that the two environments has to be the same. To avoid side effects, the role of these local observer agents becomes very crucial. These local bodies report to the central observer agent who provides a global view about the agent inter-relationship and the environment. This global view in turn helps the ranking agent to evaluate the AUTs behavior after the inclusion of mutants into them. Thus the role of observing agent is to keep track of the interactions between AUTs and their pre and post conditions along with providing the execution scenario to the ranking agent. This covers the ‘black box’ problem discussed in the previous section.

The Ranking Agent: One of the issues related with the continuous automated testing is that how many test suites must be executed on AUT in order to get a satisfactory condition about functionality of the components or units. Applying an exhaustive testing technique on those functional units which are having low importance or we can say low ranking is not a good idea. In the same way, the AUT having high rankings must be tested fully using exhaustive testing technique discussed in the subsequent sections. For low rank AUT, random testing technique can be applied, which will save time and other efforts of the testing phase. The other issue is related with the agent inter-dependency. As agents interact with each other seamlessly with other agents in the environment, the chain of agent interdependency sometimes grows profusely. It gets difficult to test the long and complex chains of the communicating agents. The role of ranking agent is to provide pathway to the tester and observer agent to limit the number of test cases of AUT based upon their rankings and to limit the length of the chain of interacting agents. Higher the ranking, more rigorous testing will be followed on AUT and maximum inter-dependent chain of agents will be tested and *vice-versa.*

The main aim of EARCT is to enhance the efficiency and quality of one of the most important

phase of software engineering- The testing phase. Agents are very complex in nature and can pose extreme problems before the testing team. Testing based on ranks can help in figuring out maximum errors by using optimal resources in any MAS.

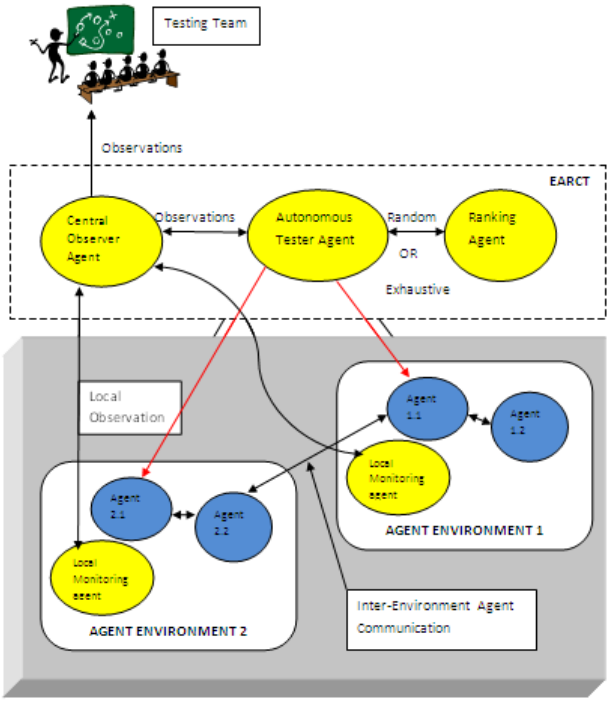


Fig. 1 – EARCT (Environment for Automated Rank based Continuous Testing)

6. TESTING STRAEGIES

As discussed earlier, the exhaustive testing is not possible for all the agents in the environment. The selective or random testing is done for low rank agents and exhaustive testing is done for the high rank agents. The two techniques are discussed in the following section:-

Rank Oriented Random Testing: The tester agent is capable of sending random generated data through random test data generation mechanisms [30, 31] to the AUT, using some communication protocol. The tester agent has to select one of the communication protocols available from the domain specification, to include meaningful and domain specific data into the communication messages. A model of the domain data, coming from the business domain of the MAS under test, must be also supplied. Various types of communication protocols are described and practiced such as UML based sequence diagram, complex cooperation protocol, activity diagrams, collaboration diagrams and the most widely accepted, the FIPA Interaction Protocol [29] in JADE [24]. In addition to the protocol, the format for the message passing between the tester agent and the AUT must also be supplied. One such type is FIPA ACLMessage [29]. Various ACLs (Agent Communication Language) are proposed by many researchers but the two most discussed and practiced ACLs are KQML (Knowledge Query and Manipulation language) and FIPA-ACL (Foundation for intelligent physical agent ACL). Both of them rely on speech act theory developed by Searle in 1960 and enhanced by Winograd and Flores in the 1970s. Speech act theory is derived from the linguistic analysis of human communication, based on the idea that with language the speaker not only makes statements, but also performs actions. But out of these two ACLs, FIPA-ACL is a standardization consortium. The JADE platform (Java Agent Development Environment) provides basis for FIPA-ACL. We will be using FIPA Interaction Protocol for agent communication and FIPA-ACL as the communication language. Due to their wide acceptability, the random test data generation technique then has to be selected by the tester agent. The random testing technique will be initiated by the rank agent based upon the ranking of the AUT. The testing team can decide the minimum (Min-R) and maximum rank (Max-R) number to initiate the testing strategy. Moreover, as discussed above, the number of AUTs collaborating together in a chain can form a huge series. The ranking of agents will be used to limit the chain for testing purpose. Higher the rank, more AUTs in the chain will be tested. The testing team can decide the minimum and maximum rank numbers and these values can be encoded in ACL message as parameters. The communication protocol can be extended to accommodate the ranks of the AUTs. In order to define the rigor of testing effort and to limit testing effort in the chain of interacting agents, the overall model of the rank based agent random testing will serve the following purpose:

- The model will prescribe the range and the structure of the data that are produced randomly, either in terms of generation rules or in the (simpler) form of sets of admissible data that are sampled randomly.
- Long and meaningful data and interaction sequence using random sampling is very hard to generate. The ranking mechanism of agents will limit the data and number of interactions between AUTs, making rank based random testing a cheap and efficient testing technique to reveal faults in the agent based system. Experimental details will be presented in the subsequent sections.

Randomly generated messages generated by the tester agent are then sent to AUT and response is observed by the observer agent. The response can be a successful state transition, an exception, a deadlock condition or a crash etc. Whenever the observer agent observes a divergence from the agent's expected behavior, the exception(s) is reported back to the testing team. It is the

responsibility of the observer agent to keep track of the actions and reactions occurring between the tester agent and the AUTs. The idea is not to fully automate the testing process but to support the manual testing in order to improve overall testing experience by reducing testing effort.

Rank Oriented Exhaustive Testing: Ranking Agent help the autonomous tester agent to decide on which AUT, random testing technique has to be applied. In the case where it has been decided by the Ranking and tester agent on the basis of ranking that random testing is not sufficient for the AUT, another testing strategy has to be applied. For those AUTs, having higher order of ranking, exhaustive testing is useful. It is evident that longer the sequence generated for the series of interacting agents, the likelihood of revealing faults is maximized, required for higher degree ranking holder AUTs. Sophisticated techniques as compared to random testing are required. One such proposed technique is rank oriented exhaustive testing. We are proposing the combination of progressive testing and mutation testing with the name- agent oriented rank based exhaustive testing. Simple exhaustive testing for a chain of agents is not possible due to amount of data to be tested. So for higher ranking agents, a combination of progressive and mutation testing [32, 33] can be done. Mutation testing is a kind of software testing, which involves modifying agent's source code in small ways. Mutation operators will be applied to original source code to inject the artificial and known defects. A mutant can be a modified branch condition, a wrong variable name or a modified method invocation process etc. A test suite will be considered as defective which does not detect and reject the mutated code. On the contrary, if the test case is able to detect the artificially seeded defect in the program, the mutant is considered to be 'killed'. The adequacy of the test case is measured as the ratio of total killed mutants over total number of mutants generated. The purpose is to help the tester develop effective tests or locate weaknesses in the test data used for the program or in sections of the code that are seldom or never accessed during execution. This is required for higher ranking AUTs or chain of AUTs to make them more reliable and robust. The next step is to combine the mutation testing with the concept of progressive testing and ranks of the AUTs. For moderate or higher ranking AUTs, this kind of exhaustive may be applied. The decision it again left with the testing team depending upon the project management and software engineering requirements. In exhaustive testing, we assume that if we want to have a best test suit, it can be developed gradually in a progressive way [34, 35]. The idea is to evolve test suites by applying mutation operators to the test cases themselves by the tester agent. The ranking agent will provide a pathway to the tester agent in the form of a heuristic value, which is the defined as the shortfall of the test case to achieve the testing goal. Lower the value (i.e. test case is an effective one), it is more likely to happen that ranking agent will choose the particular test case for progression into another stringent version. It is assumed that test suites, which can kill maximum mutants will have higher probability to reveal faults. We have designed a simulator for the same as shown in figure 2.

Phase1:	**Mutants and Test Case Generation:** 1.1 ← For test on AUTs in MAS, generate mutants {M_1, M_2... M_n} 1.2 ← Get Rank values {R_1, R2... R_n} from the ranking agent for each AUT 1.3 ← Apply all/maximum mutants to higher rank AUTs and one or more to lower rank AUTs. 1.4 ← Randomly generate test cases {TC_1, TC_2... TC_n}
Phase 2:	**Execution Phase and heuristic value evaluation:** 2.1 ← Autonomous tester agent to execute each test case on all mutants generated in the phase 1. 2.2 ← Ranking agent will compute the heuristic value of each test case: HV[TC_i]=MK_1/n, where HV stands for heuristic value, MK stands for mutants killed and n stands for total number of mutants generated.
Phase 3:	**Test case progression:** 3.1 ← Select test cases with higher order of heuristic value HV[TC_i]. 3.2 ← Apply another set of generated mutants on the test case selected by the ranking agent. 3.3 Add the new mutated test cases to the existing set of test cases.
Phase 4:	**Check Status:** 4.1 ← If maximum numbers of generations are achieved then STOP and EXIT. else 4.2 ← Check improvements in the heuristics value for all the progressive iterations. 4.3 ← If there is no improvement, go to step 1.1 else 4.4 ← Report test results to testing team.

Fig. 2 – Simulator design for Rank Oriented Exhaustive Testing

7. EXPERIMENTAL DETAILS

We have simulated the EARCT environment using a hypothetical case study derived from a hospitality sector. We have also used TAOM4E (Tool for Agent Oriented Visual Modeling for the Eclipse platform) for goal oriented modeling, code generation and testing for goal-directed system. As TAOM4E follows TROPOS methodology as its basis, we will also use the same methodology to implement our case study. TAOM4E supports TROPOS's early and late requirement modeling requirements, architectural design and also provides support for automated agent oriented implementation and testing. TAOM4E is developed by the Software Engineering unit at Fondazione Bruno Kessler (FBK), Trento. The current version

0.6.3 is downloadable under GPL license from the tool homepage http://selab.fbk.eu/taom. The t2x code generation tool provided by TAOM4E can transform problem domain to solution domain by mapping goal models to a goal-directed implementation on the Jadex BDI agent platform. Explicitly preserving goal models at run-time and providing the proper middleware for navigating this model and acting according to it. Agent code can be generated from the graphical interface, and the implemented prototypes are executable directly from the Eclipse user interface. The recent addition in TAOM4E framework is Goal Oriented Testing tool to support testing and validation along the process phases. The EARCT framework will directly derive test cases from the goal models of TROPOS and uses them to implement agents.

8. RESULT AND CONCUSION

We have done manual testing as well as testing using EARCT framework on the case study. In manual testing, we applied two manual testing techniques viz. random testing which includes branch coverage and mutation testing. The choice is obvious, we will perform the same type of testing using EARCT framework by rank oriented random testing and rank oriented exhaustive testing strategies. The results will be compared then. The results are compared on the basis of three parameters: 1) Number of errors revealed and error type, where error type could be classified as fatal, Moderate or Low based upon the severity impact of these errors on the software, if these errors would have remained hidden during testing. 2) Effort utilized; this is measured as the time taken in minutes to figure out the errors(s) in a particular module/agent. 3) As discussed, the agents can form a complex chain reaction to fulfill the social goal. We are considering one more parameter i.e. number of linked branches/modules/AUTs tested to test the unit in question. Longer the tested chain, more the team will have confidence in the testing process.

The desirable software engineering scenario is to have maximum number of errors revealed by testing maximum units in the chain and utilizing minimum effort in terms of time spent on testing.

Further, as already discussed, using agent based testing we are also implementing the concept of ranks for the agents. The ranking is done in ascending order i.e. agents having ranking «1» will be the most important module and module with ranking «2» is the next unit in the importance list. The ranks will be fetched directly from the requirement phase. Again the idea is to put maximum effort on the important modules and pay somewhat less emphasis on the less important module. This is necessary to as to optimize the quality-effort ratio. As agents based systems are very complex in nature, it is not feasible to test each and every agent exhaustively, rather few important one can be tested fully and rest partially to save time, resources and efforts.

In the table 1 shown below, we have manually tested four modules from the case study using Random Branch Testing. Table 1 shows the time taken (in minutes) to test the module, number of errors revealed, error description, error type and number of linked branches tested. In table 2, the table is having one more column in the end- the rank of the module. Again table 3 and 4 shows the parameters obtained using manual mutation testing and agent based exhaustive testing respectively.

Table 1. Random Branch Testing (Manual)

Random Branch Testing (Manual)					
Module	**Effort (Time taken in Mins) (Manual Mode)**	**No. of errors Revealed (Manual Mode)**	**Error Description**	**Error Type**	**Branch level covered (Manual Mode)**
1	20	3	Calculation Error	FATAL	3
2	34	2	Boundary-check related error	MODERATE	2
3	17	4	Boundary-check related error	MODERATE	5
4	23	5	Compatibility and intersystem defect	MODERATE	3

Table 2. Agent Based Continuous Rank Oriented Random Testing

EARCT-Agent Based Continuous Rank Oriented Random Testing						
Module	**Effort (Time taken in Mins.) (Agent Based)**	**No. of errors Revealed (Agent Based)**	**Error Description**	**Error Type**	**Number of AUTs Covered**	**Rank**
1	11	4	Calculation Error, Control Flow error	Fatal	9	**2**
2	6	3	Boundary check missing	Medium	6	**3**
3	13	7	Unhandled condition, Calculation logic error, Control Flow defects	Fatal	13	**1**
4	10	4	Performance bug	Low	6	**4**

Table 3. Mutation Testing (Manual)

Manual Mutation Testing					
Module	**Effort (Time taken in Mins) (Manual Mode)**	**No. of errors Revealed (Manual Mode)**	**Error Description**	**Error Type**	**Modules Covered (Manual Mode)**
1	20	2	Keyword constraint violation	FATAL	3
2	34	2	Update Anomaly	FATAL	2
3	17	4	Return Empty	LOW	5
4	23	5	Un-expected Result	MODERATE	3

Table 4. Agent Based Continuous Exhaustive Testing

EARCT-Agent Based Continuous Exhaustive Testing						
Module	**Effort (Time taken in Mins) (Agent Based)**	**No. of errors Revealed (Agent Based)**	**Error Description**	**Error Type**	**Number of AUTs Covered**	**Rank**
1	11	4	Calculation Error, Control Flow error	Fatal	9	**2**
2	6	3	Boundary check missing	Medium	6	**3**
3	13	7	Unhandled condition, Calculation logic error, Control Flow defects	Fatal	13	**1**
4	10	4	Performance bug	Low	6	**4**

As shown in figures 3, 4, 5, 6, 7, 8 the comparative results are highly noticeable. The time taken by the automated agent based testing is substantially low as compared to the manual one for each of the module. The evaluated error count and linked units/agents covered are also on the higher end as compared to the manual one. One noticeable observation is that in case of module 4, the performance delivered by the agent based testing mechanism is not that good in terms of evaluated error count because of its low ranking. The exhaustive strategy is not applied on this module due to its low ranking, resulting in less error discovery and less linked units covered. This is one compromise the testing team has to make while taking decisions about the subsequent testing strategy and other software engineering paradigms like correctness, completeness, consistency and of course time and quality. If the module is having higher ranking the resources utilized will be on the higher end and vice-versa.

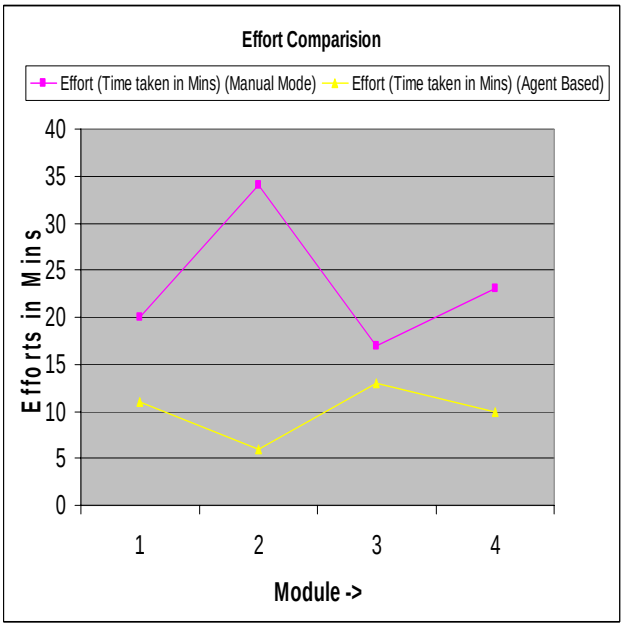


Fig. 3 – Effort Comparison- Manual Random Branch Testing Vs Agent Based Continuous Rank Oriented Random Testing

Fig. 4 – Evaluated Error Count Comparison- Manual Random Branch Testing Vs Agent Based Continuous Rank Oriented Random Testing

In this paper, we have shown that current manual techniques for testing are not good enough to answer the issues mentioned in the earlier sections. The problems like autonomy, social nature, complexity, agents' inter-dependency and non-deterministic nature can be answered using continuous testing only. Further to optimize various software

engineering paradigms like number of errors revealed in unit time and effort spent on the chain of interacting or inter-dependent agents etc, the agent ranking method is useful to help the testing team to take decisions about the subsequent steps.

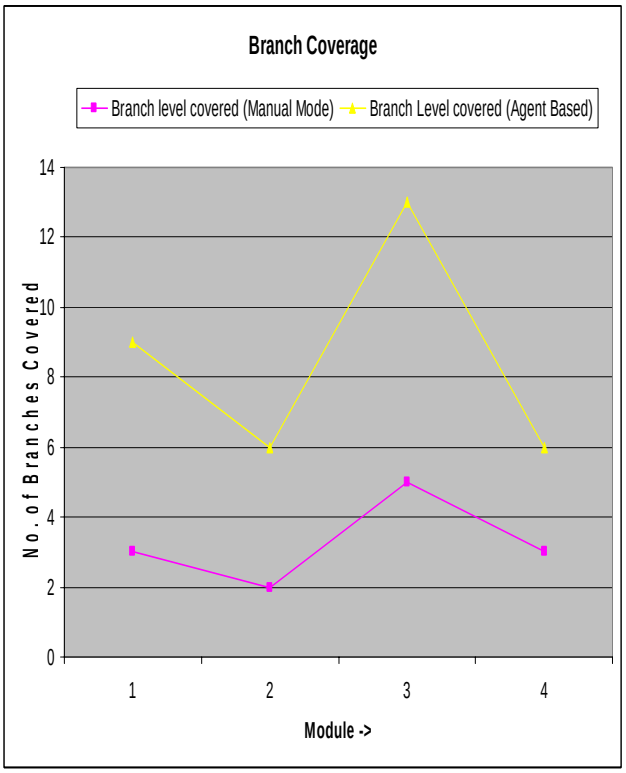


Fig. 5 – Branch/AUTs Coverage Comparison- Manual Random Branch Testing Vs Agent Based Continuous Rank Oriented Random Testing

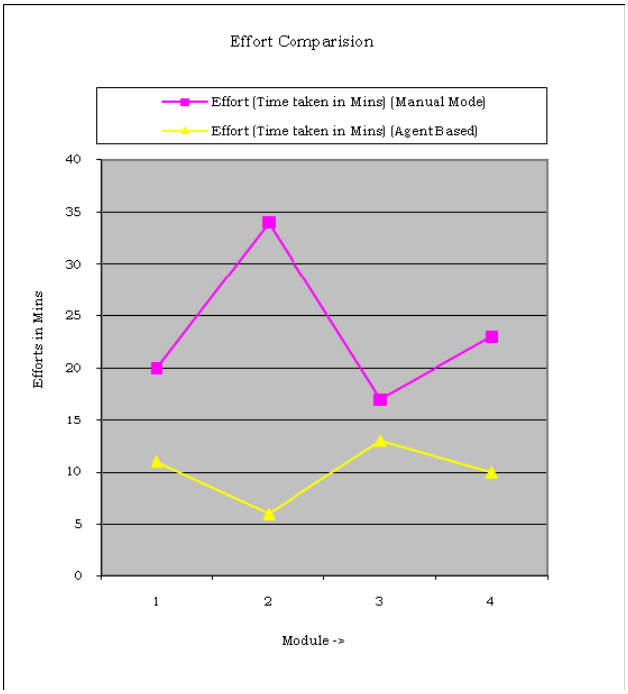


Fig. 6 – Effort Comparison- Manual Mutation Testing Vs Agent Based Continuous Rank Oriented Random Testing

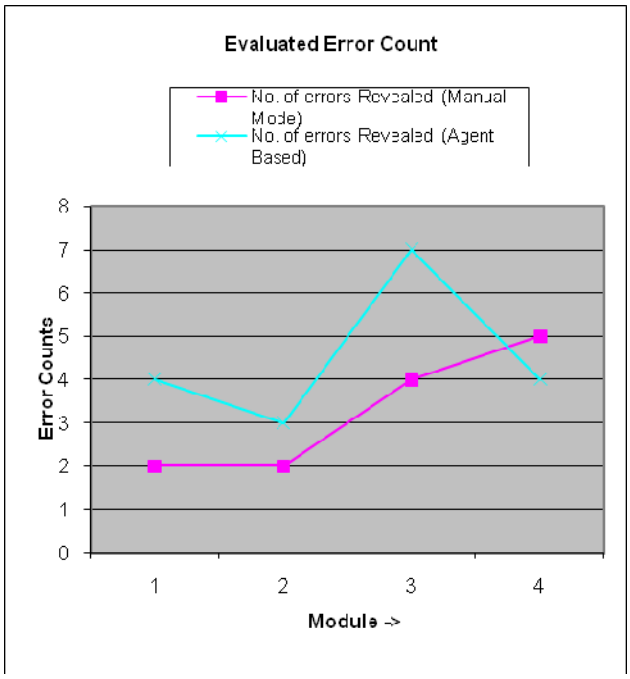


Fig. 7 – Evaluated Error Count Comparison- Manual Mutation Testing Vs Agent Based Continuous Rank Oriented Random Testing

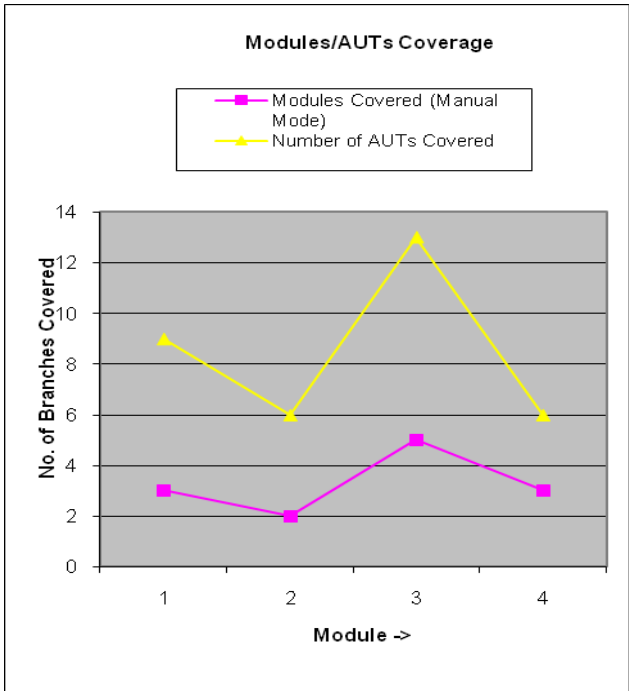


Fig. 8 – Modules/AUTs Coverage Comparison- Manual Mutation Testing Vs Agent Based Continuous Rank Oriented Random Testing

8. REFERENCES

[1] Tiryaki A.M., Oztuna S., Dikenelli O., and Erdur Sunit R., A unit testing framework for test driven development of multi-agent systems, In *7th International Workshop on Agent Oriented Software Engineering*, 2006.

[2] Roberta Coelho A.S., Kulesza U and Lucena

C., Unit testing in multi-agent systems using mock agents and aspects, *International Workshop on Software Engineering for Large-scale Multi-Agent Systems*, May 2006.

[3] Blaya J. A. B., Hernansaez J. M., and Gomez Skarmeta A. F., Towards and approach for debugging multi-agent systems through the analysis of agent messages" Comput. Syst. Sci. Eng., (20) 4 (2005).

[4] Cossentino M., *From requirements to code with PASSI methodology*, In Vijayan Sugumaran (Ed.), Intelligent Information Technologies: Concepts, Methodologies, Tools, and Applications, USA, 2008.

[5] Pavon J., Gomez-Sanz J., and Fuentes-Fernandez R., *The INGENIAS methodology and tools*, In Agent Oriented Methodologies (eds. Henderson-Sellers and Giorgini), Idea group, (2005), pp. 236-276.

[6] Huget M., and Demazeau Y., Evaluating multi agent systems: a record/replay approach, *Intelligent Agent Technology, IAT 2004, Proceedings IEEE/WIC/ACM International Conference*, (2004), pp. 536-539.

[7] Jennings N.R., An agent-based approach for building complex software systems, *Communications of the ACM*, (44) 4 (2001), pp. 35-41.

[8] Rouff C., *A Test Agent for Testing Agents and Their Communities*, IEEE, 2002.

[9] Sommerville I., *Software Engineering*, 9th edition, Addison Wesley, 2011.

[10] Zhang Z., Thangarajah J., and Padgham L., Automated unit testing for agent systems, *2nd International Working Conference on Evaluation of Novel Approaches to Software Engineering, ENASE'07*, Spain, (2007), pp. 10-18.

[11] Ekinci E., Tiryaki M., Cetin O., and Dikenelli O., Goal-oriented agent testing revisited, *Proceeding of the 9th International Workshop on Agent-Oriented Software Engineering*, (2008), pp. 85-96.

[12] Caire G., Cossentino M., Negri A., and Poggi A., Multi-agent systems implementation and testing, *Proceedings of the 7th European Meeting on Cybernetics and Systems Research – EMCSR2004*, Vienna, Austrian Society for Cybernetic Studies, (2004), pp. 14-16.

[13] Lam D., and Barber K., Debugging agent behavior in an implemented agent system, *2nd International Workshop, ProMAS, Springer, Berlin*, (2005), pp. 104-125.

[14] Nunez M., Rodriguez I., and Rubio F., Specification and testing of autonomous agents in e-commerce systems, *Software Testing, Verification and Reliability*, (15) 4 (2005), pp. 211-233.

[15] Coelho R., Kulesza U., Staa A., and Lucena C., Unit testing in multi-agent systems using mock agents and aspects, *Proceedings of the international workshop on Software engineering for large-scale multi-agent systems*, ACM Press, New York, (2006), pp. 83-90.

[16] Gamma E., and Beck K., JUnit: a regression testing framework, http://www.junit.org, 2000.

[17] Tiryaki A.M., Oztuna S., Dikenelli O., and Erdur Sunit R., A unit testing framework for test driven development of multi-agent systems, *In 7th International Workshop on Agent Oriented Software Engineering*, 2006.

[18] Dikenelli O., Erdur R., and Gumus O., Seagent: a platform for developing semantic web based multi agent systems, *AAMAS'05 Proceedings of the fourth International Joint Conference on Autonomous agents and multi-agent systems*, ACM Press, New York, (2005), pp. 1271-1272.

[19] Gomez-Sanz J., Botia J., Serrano E., and Pavon J., Testing and debugging of MAS interactions with INGENIAS, *Agent-Oriented Software Engineering IX*, Springer, Berlin, (2009), pp. 199-212.

[20] Houhamdi Z., Test suite generation process for agent testing, *Indian Journal of Computer Science and Engineering*, (2) 2 (2011).

[21] Knublauch H., Extreme programming of multi-agent systems, *International Joint Conference on Autonomous Agent and Multi-Agent Systems*, Bologna. ACM Press, (2002), pp. 704-711.

[22] Gamma E., and Beck K., JUnit: a regression testing framework, http://www.junit.org, 2000.

[23] Botia J., Lopez-Acosta A., and Skarmeta G., ACLAnalyser: a tool for debugging multi-agent systems, *Proceeding of the 16th European Conference on Artificial Intelligence*, IOS Press, (2004), pp. 967-968.

[24] TILAB, Java agent development framework, http://jade.tilab.com/. Accessed on 17th May 2011.

[25] Padgham L., Winikoff M., and Poutakidis D., Adding debugging support to the Prometheus methodology, *Engineering Applications of Artificial Intelligence*, (18) 2 (2005), pp. 173-190.

[26] Rodrigues L., Carvalho G., Barros P., and Lucena C., Towards an integration test architecture for open MAS, *1st Workshop on Software Engineering for Agent-Oriented Systems/SBES*, (2005), pp. 60-66.

[27] Nguyen C., Perini A., and Tonella P., Goal-oriented testing for MAS, *Agent-Oriented Software Engineering VIII, Lecture Notes in*

Computer Science, (4951) (2008), pp. 58-72.

[28] Houhamdi Z., and Athamena B., Structured system test suite generation process for multi-agent system, *International Journal on Computer Science and Engineering*, (3) 4 (2011), pp.1681-1688.

[29] Foundations for Intelligent Physical Agents, FIPA-specifications. http://www.fipa.org/ specifications. Accessed on 29th Nov, 2011.

[30] Mills H. D., Dyer M. D., and Linger R. C., Cleanroom software engineering, *IEEE Software*, (4) 5 (1987), pp. 19-25.

[31] Fosse P. Thevenod and Waeselynck H., Statemate: applied to statistical software testing, *In Proc. of the Int. Symposium on Software Testing and Analysis (ISSTA)*, (June 1993), pp. 78-81.

[32] DeMillo R. A., Lipton R. J., and Sayward F. G., Hints on test data selection: help for the practicing programmer, *IEEE Computer*, (11) 4 (1978), pp. 34-41.

[33] Hamlet R. G., Testing programs with the aid of a compiler, *IEEE Transactions on Software Engineering*, (3) 4 (1977), pp. 279-290.

[34] Wegener J., *Stochastic algorithms: foundations and applications*, *In Evolutionary Testing Techniques, Springer Berlin, Heidelberg*, Chapter 9, 2005, pp. 82-94.

[35] McMinn P., and Holcombe M., The state problem for evolutionary testing, *Proceedings of the International Conference on Genetic and Evolutionary Computation*, Springer, Berlin, (2003), pp. 2488-2498.

Dr. Ashok Kumar *is working as Professor and Chairman at Department of Computer Science and Applications, Kurukshetra University, Kurukshetra (India). He is PhD in Computer Science with vast teaching and research experience of more than 30 years.*

He had supervised more than 35 PhD scholars in different computing areas including operation research, software engineering, testing and designing etc.

Mr. Vinay Goyal is working as Asst. professor and Head of Department (MCA) at Panipat Institute of Engineering and Technology, Panipat (India). He had published 6 research papers in International Journals of repute on the topic Agent Oriented Software Engineering.

Besides this he had conducted an international conference with the title ICACCT 2008 at APIIT SD INDIA, Panipat (India) in Nov 2008.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ФОРМУВАННЯ ЛІНГВІСТИЧНОЇ ОНТОЛОГІЇ НА БАЗІ СТРУКТУРОВАНОГО ЕЛЕКТРОННОГО ЕНЦИКЛОПЕДИЧНОГО РЕСУРСУ

Антон Михайлюк [1)], Олена Михайлюк [2)], Олексій Пилипчук [2)], Володимир Тарасенко [2)]

1) Київський університет імені Бориса Грінченка,Тимошенко, 13-Б, Київ, 04212, Україна, may-62@ukr.net
2) НТУУ «КПІ», кафедра СПСКС, просп. Перемоги, 37, Київ, 03056, Україна.
mes@scs.ntu-kpi.kiev.ua, ilexcorp@ukr.net, vtarasen@scs.ntu-kpi.kiev.ua

Резюме: *Вирішення задачі інтелектуалізації інформаційно-пошукової діяльності вимагає застосування допоміжних спеціалізованих лінгвістичних ресурсів. Одним із таких ресурсів може бути лінгвістична онтологія предметної галузі. Стаття розглядає підхід до організації програмних засобів для автоматизованого формування онтологічної бази знань шляхом конвертації структурованого енциклопедичного ресурсу у відповідні об'єкти онтології. Розглядаються процедури створення понятєвої бази онтології, ієрархії понять та мережі асоціативних зв'язків. Проводиться дослідження якісного та кількісного складу сформованої експериментальної онтології на основі українського сегменту Вікіпедії.*

Ключові слова: *лінгвістична онтологія, семантичні відношення, структурована енциклопедія.*

A CREATION OF THE LINGUISTIC ONTOLOGY BASED ON A STRUCTURED ELECTRONIC ENCYCLOPEDIC RESOURCE

Anton Mykhailiuk [1)], Olena Mykhailiuk [2)], Oleksiy Pylypchuk [2)], Volodymyr Tarasenko [2)]

[1)] Borys Grinchenko Kyiv University, 13-b Tymoshenko str., Kyiv, 04212, Ukraine, may-62@ukr.net
[2)] NTUU «KPI», SP SCS, 37 Peremogy avenue, Kyiv, 03062, Ukraine
mes@scs.ntu-kpi.kiev.ua, ilexcorp@ukr.net, vtarasen@scs.ntu-kpi.kiev.ua

Abstract: *Solving the problem of intelligent information retrieval requires the use of additional specialized linguistic resources. One of such type of resources is a linguistic ontology of a subject field. The paper considers an approach to develop software for an automatic creation of an ontological knowledge base by the converting structured encyclopedic resource into appropriate ontology's objects. The procedures of creation of the ontology entities base, the hierarchy of entities and the network of associative links are considered. Qualitative and quantitative analysis of the produced experimental ontology based on Ukrainian part of Wikipedia is investigated.*

Keywords: *linguistic ontology, semantic relations, structured encyclopedia.*

ВСТУП

Інтелектуалізація процедур аналізу контенту та інформаційно-пошукової діяльності вимагає розробки спеціалізованих лінгвістичних ресурсів, які б могли бути використані для підвищення ефективності інформаційно-аналітичних програмних засобів. Це дало б суттєвий поштовх розвитку таких лінгвістичних ресурсів, як словники синонімів, комп'ютерні тезауруси [1], семантичні мережі [2] тощо. Зокрема, останнім часом значно посилився інтерес до застосування онтологій під час роботи з текстомісткими об'єктами, що не випадково, оскільки онтологія за своєю природою має відображати структуру людських знань про оточуючий світ. Це дозволяє використовувати її для привнесення змістовної компоненти в процес обробки інформаційних об'єктів, їх аналізу, а

також в процедури інформаційного пошуку по цих об'єктах. Поняття онтології дуже широко трактується як в загально філософському сенсі, так і безпосередньо в рамках інформаційних технологій. Це пов'язано, в першу чергу, з тим, що онтології можуть різнитись за рівнем подання, за призначенням, за сферою застосування та методикою формування тощо. Оскільки однією з найпоширеніших галузей, де застосування онтології вбачається доцільним та перспективним, є сфера інформаційно-пошукової діяльності (в першу чергу це пошук [3, 4] або інформаційний моніторинг [5] текстових об'єктів в глобальній або локальній інформаційній мережі), тому онтологію доцільно розглядати з позиції лінгвістики, оскільки об'єкти мають текстову природу. Лінгвістична онтологія – це спеціальна база знань, що описує поняття зовнішнього світу та відношення між ними. Разом з тим це особливий клас онтологій, де поняття формуються на основі мовних одиниць, що відносяться до певної предметної галузі [1]. Можна виділити наступні важливі риси для структурно-логічних властивостей подібних онтологій:

- кожне поняття предметної галузі подається в онтології за допомогою синсета – сукупності близьких синонімів, що мають подібне значення;
- кожний синсет має певний зміст, що подається за допомогою унікального тлумачення;
- різні поняття можуть мати однакове мовне подання, в такому разі можна говорити про багатозначні терміни, а смислове розрізнення таких понять можливе за допомогою тлумачень;
- всі синсети об'єднуються в єдину ієрархічну таксономію понять – від абстрактних понять до конкретних;
- синсети онтології зв'язані між собою за допомогою семантичних відношень, найпоширеніші з яких – це відношення асоціації та гіпонім-гіперонім [1].

Крім того, використання онтологій в автоматичному режимі для синтетичних мов, до яких належить і українська мова, вимагає доповнення онтології нормальними формами слів для понять онтології.

Формування онтології програмними засобами передбачає аналіз певного текстового масиву даних, виокремлення понять та ідентифікацію зв'язків між ними у відповідності до заданих семантичних відношень. На жаль, така процедура вимагає надто складного апарату семантичного аналізу, який втім не гарантує якісної реалізації онтології через неоднозначність інтерпретації природомовних текстів. Вирішенням цього питання, на нашу думку, може бути залучення колекції документів енциклопедичного (словникового) характеру, що мають чітку структуру та створюються експертами. Саме такою структурованою колекцією для нас вбачається вільна електронна енциклопедія – Вікіпедія [6]. Її статті являють собою опис того чи іншого поняття, при цьому описова частина має цілком чітку структуру і, крім того, містить посилання на інші статті. Це робить можливим автоматизацію обробки таких документів та формування онтології.

Отже мета даної статті полягає в розробці способів організації програмних засобів, які б дозволили в автоматичному або напівавтоматичному режимі створити лінгвістичну онтологічну базу знань на основі статей певного мовного сегменту (напр., українського сегменту) Вікіпедії для застосування в інформаційно-пошуковій діяльності, зокрема в процедурах квазісемантичного пошуку [7].

1. СТРУКТУРА ВХІДНИХ ДАНИХ

Робота програмних засобів формування онтології має використовувати особливості структурної організації матеріалів Вікіпедії. Створення статей відбувається з використанням спеціального формату MediaWiki [8] (Рис.1).

```
<page>
  <title>Кортеж</title>
  <id>19488</id>
  <revision>
    <timestamp>2010-01-21T18:45:56Z</timestamp>
    <contributor>
      <username>Igor Yalovecky</username>
    </contributor>
    <text xml:space="preserve">'''Кортéж''' або "n"-ка  – в [[математика|математиці]] впорядкована та [[скінченна множина|скінченна]] сукупність елементів …
==Дивіться також==
* [[Декартів добуток]]
* [[Формальна мова]]
[[Категорія:Теорія множин]]
[[Категорія:Реляційна модель даних]]
[[en:Tuple]] [[ru:Кортеж]]
    <text>
  </revision>
</page>
```

Рис.1 – Фрагмент XML-документа статті Вікіпедії

Цей формат дозволяє однозначно описувати структурні елементи статті (заголовки, розділи, посилання, мета-елементи тощо), що в свою чергу дає змогу ефективно в автоматичному режимі їх обробляти. Весь архів колекції певного мовного сегменту статей Вікіпедії зберігається у вигляді XML-документа і доступний для завантаження. XML-розмітка дає змогу виділяти необхідні вузли для подальшої обробки. На рис.1 наведений приклад вузла, що описує одну із статей із українського сегмента. Даний фрагмент ілюструє структуру XML-вузла для статті, що описує поняття «Кортеж». Як видно із цього фрагмента, даний вузол містить не лише вихідний код статті, але й деяку допоміжну метаінформацію. Власне для аналізу контенту інтерес становлять два поля – це поле <title> (заголовок статті) та поле <text> (безпосередньо текст статті).

Усю множину статей Вікіпедії за різними критеріями можна віднести до декількох груп, що зведені в табл. 1.

Аналіз структури Вікіпедії дозволяє зробити висновок про те, що статті Вікіпедії разом із взаємними внутрішніми посиланнями створюють певний прототип онтологічної бази знань [6]. Основну роль тут відіграють тлумачні статті та статті категорій, а група статей багатозначних понять виконує допоміжну роль при обробці тлумачних статей по кожному зі значень. Крім того, незавершені статті можуть бути використанні для ідентифікації зв'язку між іншими статтями, якщо ті будуть містити на неї посилання.

Таблиця 1. Види статей Вікіпедії

Вид статей	Опис
Тлумачні статті	Статті описують певне поняття, подію або явище. Відповідно до назви є основним джерелом інформаційного наповнення енциклопедії та відповідно служать головним джерелом для створення онтології.
Статті, що описують багатозначне поняття	Цей вид статей призначений для зберігання списку усіх наявних на поточний момент у відповідному сегменті Вікіпедії значень деякого терміну. Такі статті містять посилання на відповідну тлумачну статтю, якщо така є, і короткий унікальний опис по кожному з семантичних значень терміна. Основний індикатор таких статей – це наявність службової мітки *{{disambig}}* на початку текстової частини.
Статті категорій	Статті, що описують поняття-категорію із загальної ієрархії категорій Вікіпедії. В тлумачних статтях одне або декілька таких понять можуть вказуватись як батьківські категорії.
Статті, що описують файли	Файлові статті описують файлові об'єкти Вікіпедії (наприклад, зображення) та містять специфічну для відповідних об'єктів інформацію – посилання, розмір, тип тощо.
Незавершені статті	В тлумачних статтях часто зустрічаються посилання на статті, які з тих чи інших причин ще не написані. Вони не мають описової частини, але мають заголовок.
Службові статті	Такі статті не мають безпосереднього інформаційного навантаження, спрямованого безпосередньо на користувача, і використовуються в процесі розробки Вікіпедії. Зокрема це можуть бути шаблони статей, довідки, зауваження, обговорення і т. ін.

2. СТРУКТУРА ВИХІДНИХ ДАНИХ

Виявлені особливості структури Вікіпедії дозволяють розглядати програмне забезпечення для формування лінгвістичної онтології як засіб своєрідної конвертації елементів Вікіпедії у об'єкти онтології. Оскільки останні будуть активно використовуватись в інформаційно-пошуковій діяльності, питання зберігання таких об'єктів з можливістю ефективного та оперативного доступу до них може бути вирішено шляхом застосування системи баз даних. Відповідно до логічної організації онтології, на рис. 2 запропонована структурна схема реляційної бази даних, яка містить основні компоненти онтології та зв'язки між ними. Результатом формування інформаційного наповнення онтології будуть заповнені відповідними значеннями таблиці бази даних. Розглянемо детальніше всі реляційні відношення в наведеній базі даних.

Synset – центральне реляційне відношення, яке відображає синсет онтології. Воно містить наступні атрибути: унікальний ідентифікатор (id), символьне подання синсета українською (ua), російською (ru) та англійською (en) мовами, унікальне смислове тлумачення синсету (descr).

Поля ru та en введені для встановлення потенційного зв'язку україномовної онтології з аналогічними онтологіями, створеними на основі російського та англійського сегментів Вікіпедії. В залежності від потреб, множину мов можна як розширювати, додаючи відповідні поля до складу реляційного відношення синсету, так і зовсім відмовитись від додаткових мовних атрибутів, залишивши лише назву синсету основною мовою.

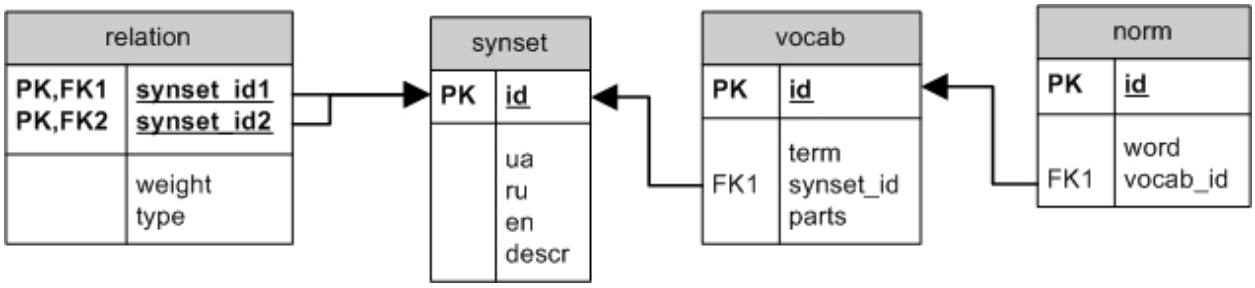


Рис. 2 – Структура бази даних онтології

Vocab – допоміжне реляційне відношення, призначене для зберігання усіх синонімічних входів синсету. Воно складається з унікального ідентифікатора (id), терміну – символьного подання синонімічного входу (term), ідентифікатора батьківського синсету (synset_id) та поля, що містить кількість слів присутніх в терміні (parts).

Norm – допоміжне реляційне відношення для зберігання нормальних форм слів, що входять до складу символьного подання терміну з Vocab. Складається з унікального ідентифікатора (id), символьного подання нормальної форми слова з терміну (word) та ідентифікатора батьківського терміну з Vocab (vocab_id).

Relation – ключове реляційне відношення, що відображає наявні зв'язки між поняттями онтології. Поля synset_id1 та synset_id2 містять ідентифікатори синсетів, між якими присутній семантичний зв'язок. Поле weight характеризує цей зв'язок певним значенням ваги, чим більше це значення, тим сильнішим вважається зв'язок між синсетами. Нарешті, поле type зберігає тип семантичного відношення, так в контексті квазісемантичного пошуку розглядаються два типи відношень – асоціація та гіпонім-гіперонім.

3. АЛГОРИТМІЧНА ОРГАНІЗАЦІЯ ФОРМУВАННЯ ОНТОЛОГІЇ

Процес формування онтології можна розділити на декілька етапів: підготовчий етап, етап створення бази синсетів, етапи побудови ієрархічних та асоціативних зв'язків, а також заключного етапу корекції. Наведений на діаграмі (рис.3) клас mediawiki_parser забезпечує виконання всіх вказаних кроків по створенню онтології. Оскільки вихідні дані статей Вікіпедії подаються в форматі XML-документа, необхідний початковий етап, основна мета якого – це парсинг XML даних та підготовка їх до подальшої обробки. Згідно фрагмента таких даних, що були представлені вище, в процесі створення онтології необхідно послідовно переходити до кожної наступної статті (вузол page). Із піддерева вузла page для створення онтології далі становлять інтерес заголовок (вузол title) та текст статті (вузол text).

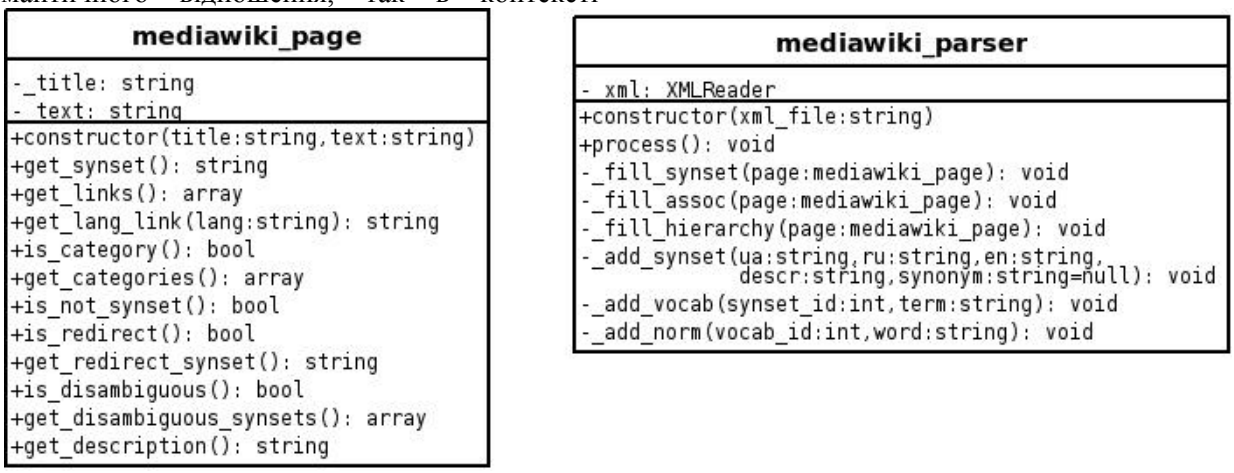


Рис. 3 – Діаграма класів для роботи з Вікіпедією у форматі MediaWiki.

На основі отриманих даних створюється екземпляр класу mediawiki_page для роботи зі сторінкою Вікіпедії у форматі MediaWiki (рис. 3). Він дозволяє оперувати даними статті, робити різні перевірки, виділяти логічні фрагменти тощо.

Відповідно до діаграми (рис. 3) клас mediawiki_page реалізує наступні важливі методи:

- get_synset – повертає символьне подання синсету для даного типу статті (якщо це можливо);
- get_links – аналізує текст статті та повертає масив посилань на інші статті Вікіпедії;
- get_lang_link(lang) – повертає посилання на відповідну статтю, написану іншою мовою відповідно до параметра lang, якщо така існує;
- is_categoгy – перевіряє, чи має поняття, що описується даною статтею, статус категорії;
- get_categories – повертає перелік категорій, до яких належить дана стаття;
- is_not_synset – перевіряє, чи може поняття, що описується даною статтею, інтерпретуватися як синсет;
- is_redirect – перевіряє, чи є дана стаття лише перенаправленням на іншу статтю;
- get_redirect_synset – повертає символьне подання синсету, на статтю якого здійснюється перенаправленням з даної статті;
- is_disambiguous – перевіряє, чи не є дана стаття описом багатозначного поняття;
- get_disambiguous_synsets – повертає перелік синсетів з однаковим написанням, але з різними тлумаченнями, сформований відповідно до вмісту статті багатозначного поняття;
- get_description – повертає тлумачення для поняття, що описується даною статтею.

Всі наведені вище методи класу активно використовуються на відповідних етапах формування онтології. Після підготовчого етапу настає черга заповнення безпосередньо бази даних онтології. Сам процес інкапсульований у методі process класу mediawiki_parser. На рис.4 наведений псевдокод цього методу.

```
/* Створюємо XMLReader */
xml = new XMLReader(xml_file);
/*Формуємо базу синсетів*/
/* Поки можливо, зчитуємо наступний вузол page */
while (xml_page = xml->next('page')) do
        /*Створюємо об'єкт статті*/
        page = new mediawiki_page(xml_page->get('title'), xml_page->get('text'));
        /*Запускаємо метод заповнення синсетів на основі поточної статті*/
        this->_fill_synset(page);
end while;
/*Повертаємо вказівник XMLReader на початок файлу*/
xml->reset();
/*Формуємо зв'язки онтології*/
/* Поки можливо, зчитуємо наступний вузол page */
while (xml_page = xml->next('page')) do
        /*Створюємо об'єкт статті*/
        page = new mediawiki_page(xml_page->get('title'), xml_page->get('text'));
        /*Запускаємо метод заповнення ієрархічних зв'язків*/
        this->_fill_hierarchy(page);
        /*Запускаємо метод заповнення асоціативних зв'язків*/
        this->_fill_assoc(page);
end while;
```

Рис.4. Псевдокод заповнення бази даних онтології

На етапі створення бази синсетів фактично формується основа онтології, її понятєва складова. В кожній статті шляхом аналізу її заголовка і тексту визначається тип поняття, яке вона описує, і відповідним чином заповнюються таблиці synset, vocab та norm. Для цього використовуються методи класу mediawiki_parser: _add_synset, _add_vocab та _add_norm. Необхідно зазначити, що метод _add_synset автоматично додає інформацію в словник синсетів за допомогою виклику _add_vocab, який в свою чергу автоматично викликає метод _add_norm. На рис.5 наведений псевдокод, який абстрактно показує послідовність дій на даному етапі (метод _fill_synset класу mediawiki_parser).

Етап створення ієрархічних зв'язків онтології, як зазначалось раніше, базується на використанні ієрархії статей Вікіпедії, що описують категорії. В кінці кожної тлумачної

статті чи статті, що описує категорію, вказується перелік батьківських категорій, до яких така стаття належить. Неважко помітити, що в рамках Вікіпедії одне поняття може відноситись одразу до декількох категорій. Ієрархічна структура, сформована таким чином, буде відрізнятись від класичної ієрархії понять, де в кожного поняття є лише одне більш загальне поняття (інакше кажучи, лише один батьківський елемент). Однак таке обмеження не відповідає реальним відношенням між поняттями, оскільки для більшості із них неможливо однозначно виявити тільки один батьківський елемент. Саме тому пряма конвертація всіх категоріальних зв'язків Вікіпедії в ієрархічну структуру синсетів лінгвістичної онтології розглядається як найоптимальніша, за умови, що при проектуванні інструментарію, який буде використовувати безпосередньо гіперонімічні зв'язки онтології буде врахована вказана вище особливість формування таких зв'язків. Відповідно до цього псевдокод методу _fill_hierarchy буде виглядати як на рис.6.

Етап створення асоціативних зв'язків принципово відрізняється від етапу створення ієрархічних зв'язків лише тим, що тут для ідентифікації зв'язків використовуються посилання на інші статті Вікіпедії (а отже й інші поняття в онтології) в тексті поточної статті. Крім того, оскільки посилання може зустрічатись неодноразово, а також можуть мати місце зворотні зв'язки (коли стаття, на яку здійснюється перехід, містить зворотні посилання), це дає змогу підвищувати вагу даного відношення на деяку величину *delta*, тим самим відображаючи посилення семантичного відношення. Чим більше таких посилань, тим сильніший зв'язок між статтями, відповідно тим більшою має бути вага відношення. Для того, щоб розподілення ваги зв'язків по онтології було достатньо рівномірним, величину *delta* варто вибирати порівняно невелику по відношенню до початкового значення ваги w_0. Заповнення асоціативних зв'язків відбувається в методі _fill_assoc, псевдокод якого подано на рис.7.

```
/*Якщо стаття описує не синсет*/
if page->is_not_synset() then
        return; /*припиняємо роботу*/
/*Якщо стаття описує перенаправлення на іншу статтю*/
else if page->is_redirect() then
        /*Додаємо новий синсет з символьним поданням поняття із статті,*/
        /* на яку йде перенаправлення, в якості синоніма подаємо поняття поточної статті*/
        this->_add_synset(page->get_redirect_synset(), '', '', '', page->get_synset());
/*Якщо стаття описує багатозначне поняття*/
else if page->is_disambiguous() then
        /*Отримуємо перелік всіх синсетів та їх тлумачень*/
        list(synset,description) = page->get_disambiguous_synsets();
        /*Для кожного запису із переліку створюємо новий синсет*/
        for each record in list(synset,description) do
                /*Додаємо новий синсет synset з тлумаченням description*/
                /*в якості синоніма передаємо поточний заголовок статті*/
                this->_add_synset(synset, '', '', description, page->get_synset());
        end for;
/*Якщо стаття описує поняття або категорію*/
else
        /*Додаємо новий синсет*/
        this->_add_synset(
                page->get_synset(),
                page->get_lang_link('ru'),
                page->get_lang_link('en'),
                page->get_description());
end if;
```

Рис.5. Псевдокод заповнення бази синсетів онтології

```
/*Якщо стаття не є описом поняття або категорії*/
if page->is_not_synset() OR  page->is_redirect() OR  page->is_disambiguous() then
        return; /*припиняємо роботу*/
/*Якщо стаття описує поняття або категорію*/
else
        /*Для кожної категорії статті*/
        for each category in page->get_categories() do
                /*Зберігаємо відношення ієрархії з вагою відношення рівне w0*/
                database->relation->insert(
                        'synset_id1' = category,
                        'synset_id2' = page->get_synset(),
                        'type' = 'H',
                        'weight' =w0
                );
        end for;
end if;
```

Рис.6. Псевдокод формування ієрархічних зв'язків онтології

```
/*Якщо стаття не є описом поняття або категорії*/
if page->is_not_synset() OR  page->is_redirect() OR  page->is_disambiguous() then
        return; /*припиняємо роботу*/
/*Якщо стаття описує поняття або категорію*/
else
        /*Для кожного посилання на іншу статтю*/
        for each link in page->get_links() do
                /*Якщо дане відношення уже було занесено в базу даних*/
                if database->relation->select_where(
                        'synset_id1' == page->get_synset(),
                        'synset_id2' == link) then
                        /*Оновлюємо вагу відношення на деяку величину delta*/
                        line = database->relation->select_where(
                                'synset_id1' == page->get_synset(),
                                'synset_id2' == link);
                        line->update(weight = weight + delta);
                else
                        /*Зберігаємо відношення асоціації з вагою відношення рівне w0*/
                        database->relation->insert(
                                'synset_id1' = link,
                                'synset_id2' = page->get_synset(),
                                'type' = 'A',
                                'weight' = w0
                        );
                end if;
        end for;
end if;
```

Рис.7. Псевдокод формування асоціативних зв'язків онтології

Нарешті, останній етап – це необов'язковий етап корекції уже створеної на попередніх стадіях онтологічної бази. Він може відбуватися як автоматично (наприклад, видалення можливих службових понять та всіх їх зв'язків), так і шляхом «ручного» редагування онтологічної бази редакторами онтології (наприклад, видалення нерелевантних або помилкових зв'язків тощо). Така корекція може знадобитись зважаючи на неповну прозорість автоматичного формування онтології та наявність службової інформації в першоджерелі, яка може відображатись у специфічних для енциклопедії статтях (напр., наявність категорії «статті на букву А» тощо), включати які в онтологію недоцільно.

4. РЕЗУЛЬТАТИ ТА ВИСНОВКИ

Для створення і аналізу експериментальної лінгвістичної онтології розроблено дослідний прототип програмного забезпечення конвертації

українського сегменту Вікіпедії у відповідну україномовну онтологічну базу знань. Результати роботи програмного модуля формування лінгвістичної онтології на основі статей українського сегменту Вікіпедії на момент проведення експерименту зведено у табл. 2. Наведені тут кількісні характеристики дають підстави стверджувати, що така онтологія охоплює досить широку область людського знання і може бути використана в роботі інформаційних систем. Необхідно зробити наголос на тому, що створена таким чином онтологія – це лише початковий етап, базис для подальшої роботи експертів і в той же час науково-інформаційний ресурс для дослідницько-експериментальної роботи в області пошукових технологій. В процесі подальшої роботи над онтологією можна вносити нові поняття, привносити нові зв'язки чи видаляти помилкові.

Табл.2. Кількісні характеристики онтології

Характеристика	Кількісний показник
Кількість синсетів	243948
Середня кількість слів у синсеті	2.24
Кількість ієрархічних зв'язків	127366
Кількість асоціативних зв'язків	2658584

Якісний аналіз створеної онтології в цілому провести фактично неможливо, зважаючи на дуже велику кількість понять та зв'язків між ними. Цей процес вимагає довготривалого використання такої онтології на практиці та потребує значних людських та часових ресурсів. Підвищення якості онтологічної бази знань має відбуватись паралельно із її використанням в розробці процедур інформаційного пошуку, аналізу контенту тощо. Втім оціночну характеристику можна визначити шляхом вибіркового аналізу. Для цього будують декілька невеликих точкових підмереж (наприклад, таку, як зображено на рис.8) з різних частин онтології та оцінюють основні зв'язки між синсетами.

Розгляд таких фрагментів дозволяє зробити висновок, що поняття та відношення між ними, за деякими незначними винятками, в основному відповідають понятєвій структурі предметних галузей. Тому якісний склад онтології та адекватність відношень можна вважати цілком прийнятними для використання в інформаційно-пошуковій діяльності.

Рис. 8. Фрагмент асоціативних зв'язків сформованої онтології

Таким чином, наведені процедури формування онтології дають можливість порівняно легко в автоматичному режимі створити достатньо об'ємну лінгвістичну онтологію, що буде відповідати основним вимогам до таких ресурсів. Крім того, оскільки Вікіпедія, як основне джерело інформації для формування онтології, має властивість багатомовності, в перспективі можливе створення кросмовних онтологій, що значно розширить сферу їх застосування.

5. СПИСОК ЛІТЕРАТУРИ

[1] B.V. Dobrov, N.V. Lukashevich, The linguistic ontology of nature sciences and technologies for information retrieving applications, S*cientists' notes of Kazan. univ.,* 2 (2007), pp. 49-72. (in Russian)

[2] E.F. Skorohodko, *Semantic Networks and Automatic Text Processing*, Naukova Dumka, Kyiv, 1983, p. 217. (in Russian)

[3] D.V. Lande, *Knowledge Retrieving at the Internet. professional work*, Williams, Moscow, 2005, p. 272. (in Russian)

[4] D.V. Lande, A.A. Snarskiy, I.V. Bezsydnov, *Internetics. Navigation in the Complex Networks: Models and Algorithms*, Librikom, Moscow, 2009, p. 264. (in Russian)

[5] D.V. Lande, *The Base of Information Flows Integration*, Engeeniring, Kyiv, 2006, p. 240. (in Russian)

[6] Z.S. Syed, T. Finin, A. Joshi, Wikipedia as an ontology for describing documents, *In Proceedings of ICWSM*, AAAI Press, 2008, pp. 136-144.

[7] A.Y. Mykhailiuk, O.V. Pylypchuk, M.V. Snizhko, V.P. Tarasenko, Quasi-semantic search of textual data in an electronic

information source, *Radioelectronics and Informatics*, KhNURE, Kharkov, 3 (2009), pp. 61-67. (in Ukrainian)

[8] D.J. Barrett, *MediaWiki*, O'Reilly Media, 2008, p. 384.

***Антон Михайлюк**, доцент кафедри Інформатики Київського університету імені Бориса Грінченка. Кандидат технічних наук, старший науковий співробітник. Наукові інтереси: методи та засоби інтелектуального аналізу природномовних текстових інформаційних об'єктів, інноваційні підходи до комп'ютеризації науково-освітньої діяльності.*

***Олена Михайлюк**, науковий співробітник кафедри системного програмування і спеціалізованих комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут». Закінчила Київський політехнічний інститут у 1989р. Наукові інтереси: експертні системи, їх застосування в задачах аналізу даних.*

***Олексій Пилипчук**, асистент кафедри системного програмування і спеціалізованих комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут». Закінчив НТУУ «Київський політехнічний інститут» у 2008р.*

Наукові інтереси: інформаційний пошук, інформаційний-моніторинг, квазісемантичний пошук текстових даних, системи транспортної логістики.

***Володимир Тарасенко**, завідувач кафедри системного програмування і спеціалізованих комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут». Доктор технічних наук, професор, заслужений діяч науки і техніки України, лауреат Державної премії України в галузі науки і техніки. Наукові інтереси: методи та засоби підвищення ефективності обробки ресурсів глобального електронного інформаційного простору.*

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

A CREATION OF THE LINGUISTIC ONTOLOGY BASED ON A STRUCTURED ELECTRONIC ENCYCLOPEDIC RESOURCE

Anton Mykhailiuk [1)], Olena Mykhailiuk [2)], Oleksiy Pylypchuk [2)], Volodymyr Tarasenko [2)]

[1)] Borys Grinchenko Kyiv University, 13-b Tymoshenko str., Kyiv, 04212, Ukraine, may-62@ukr.net
[2)] NTUU «KPI», SP SCS, 37 Peremogy avenue, Kyiv, 03062, Ukraine
mes@scs.ntu-kpi.kiev.ua, ilexcorp@ukr.net, vtarasen@scs.ntu-kpi.kiev.ua

Abstract: *Solving the problem of intelligent information retrieval requires the use of additional specialized linguistic resources. One of such type of resources is a linguistic ontology of a subject field. The paper considers an approach to develop software for an automatic creation of an ontological knowledge base by the converting structured encyclopedic resource into appropriate ontology's objects. The procedures of creation of the ontology entities base, the hierarchy of entities and the network of associative links are considered. Qualitative and quantitative analysis of the produced experimental ontology based on Ukrainian part of Wikipedia is investigated.*

Keywords: *linguistic ontology, semantic relations, structured encyclopedia.*

1. INTRODUCTION

Intellectualization of content analysis procedures and information retrieval requires the development of specialized linguistic resources such as dictionaries of synonyms, thesauri computer [1], semantic networks [2], and, in particular, linguistic ontologies. So the purpose of this paper is to develop ways of organizing software that would allow to create linguistic ontological knowledge base, based on articles of a language segment (eg, the Ukrainian segment) of Wikipedia, in automatic or semi-automatic mode for using in retrieval activities, including procedures quasisemantic search [7].

2. INPUT DATA STRUCTURE

The work of software tools of ontology building use the structural organization features of Wikipedia materials. Special format MediaWiki is used in articles creation [8]. The whole archive collection of a language segment of Wikipedia articles stored as an XML-document and it is available for download. XML-layout allows to distinguish the necessary components for further processing. Fig. 1 shows the example of a site that describes one of the articles of the Ukrainian segment. The entire set of Wikipedia articles on various criteria can be attributed to several groups, which represented in Table 1. Analysis of the Wikipedia structure allows to do the conclusion that Wikipedia articles with mutual internal links create the prototype of an ontological knowledge base [6].The main role belongs here explanatory articles and categories articles; a group of articles of multivalued concepts has a supporting role in the processing of explanatory articles on each of the values. In addition, incomplete articles can be used to identify the relationship between other articles if they contain link to it.

3. OUTPUT DATA STRUCTURE

Detected features of the Wikipedia structure allows to consider software to form a linguistic ontology as a means of converting Wikipedia elements to ontology objects. Since they will be actively used in information-retrieval activities, that issues of storage such objects with the possibility of efficient and operative access to them can be solved by means of a database system application. According to logic of the ontology, the block diagram of a relational database with basic components of ontology and relations between them is proposed in Fig. 2. Database tables with filled corresponding values are the result of information filling of the ontology. Consider more detail all relations in the relational database. Synset – a central relation that reflects an ontology synset. It contains the following attributes: a unique identifier (id), symbolic representation of a synset by Ukrainian (ua), Russian (ru) and English (en) languages, unique semantic interpretation of the synset (descr).

The fields ru and en were introduced to establish a potential connection of the Ukrainian ontology with similar ontologies created on the basis of Russian and English segments of Wikipedia. Vocab – an auxiliary relation designed to store all synonymous inputs of the synset. It consists of a unique identifier (id), a term – a symbolic representation of the synonymic input (term), an identifier of a parent synset (synset_id) and a field that contains the number of words in the term (parts). Norm – an auxiliary relation for storage of normal forms of words which comprise to the symbolic representation of the term from Vocab. Relation – a key relation that reflects the existing connections between the ontology concepts. The fields synset_id1 and synset_id2 contain synset identifiers between which there is a semantic relationship. A field weight characterizes this relationship by certain value of the weight. The more the value, the stronger is the relationship between synsets. Finally, a field type stores the type of the semantic relation, so in the context of the quasisemantic search two types of relationships are considered – association and hyponym/hypernym.

4. ALGORITHMIC ORGANISATION OF ONTOLOGY FORMING

The formation of the ontology can be divided into several stages: the preparatory stage, the stage of creation of a synsets database, the stages of building of hierarchical and associative relationships and the final stage of correction. Mediawiki_parser class ensures the execution all these steps to create the ontology. According to the diagram (Fig. 3) mediawiki_page class implements the following important methods:

- get_synset – returns the symbolic representation of the synset for this type of an article (if it possible);
- get_links – analyzes the text of the article and returns an array of links to other Wikipedia articles;
- get_lang_link (lang) – returns a reference to the relevant article written in another language in accordance with the lang parameter, if it is;
- is_category – checks whether the concept described in this article has the category status;
- get_categories – returns a list of categories, which related with this article;
- is_not_synset – checks whether the concepts described in this article, can be interpreted as a synset;
- is_redirect – checks whether a given article only redirect to another page;
- get_redirect_synset – returns the symbolic representation of the synset, of which article the redirection from this article is executed;
- is_disambiguous – checks whether this article describes a multi-valued concept;
- get_disambiguous_synsets – returns a synset list with the same spelling but with different interpretations, formed in accordance with the content of a multivalued concept article;
- get_description – returns the interpretation of a concept, described in this article.

All specified class methods are widely used at the appropriate stages of ontology forming. After the preparatory phase the ontology database is filled. The process encapsulated in the process method of mediawiki_parser class. The basis of an ontology (its conceptual component) is formed actually at the stage of the synsets database creation. For each article the type of a concept is determined by analyzing its title and text, and synset, vocab and norm tables are filled. Mediawiki_parser class methods are used: _add_synset, _add_vocab and add norm. It should be noted that _add_synset method automatically adds information to the synsets dictionary by calling _add_vocab, which in turn automatically invokes the _add_norm method.The stage of creating of hierarchical relationships for ontology, as noted earlier, is based on the using of Wikipedia articles hierarchy that describes the categories. At the end of each explanatory article or article, which describes the category, the list of parent categories to which this article refers is specified. Links to other Wikipedia articles (and hence other concepts in the ontology) in the text of the current article are used for relationships identification in the stage of creating associative connections. It is the fundamental difference from the stage of creating of hierarchical relationships. In addition, because the reference can occur repeatedly, and there may be feedbacks, it allows to increase the weight of the relationship at some value delta. The associative relationships are filled in the _fill_assoc method.

5. RESULTS AND CONCLUSIONS

The research prototype of conversion software of Wikipedia's Ukrainian segment to the appropriate Ukrainian-ontological knowledge base for creation and analyzing of experimental linguistic ontology is developed. The work results of the software module of forming the linguistic ontology of Ukrainian segment of Wikipedia articles at the time of the experiment are shown in the Table 2.

These quantitative characteristics give reason to believe that such ontology covers quite a wide area of human knowledge and can be used in the work of information systems.

Table 2. Ontology quantitative characteristics

Characteristics	Quantitative parameter
Quantity of synsets	243948
Average number of words in a synset	2.24
Quantity of hierarchical relationships	127366
Quantity of associative relationships	2658584

6. REFERENCES

[1] B.V. Dobrov, N.V. Lukashevich, The linguistic ontology of nature sciences and technologies for information retrieving applications, S*cientists' notes of Kazan. univ.*, 2 (2007), pp. 49-72. (in Russian)

[2] E.F. Skorohodko, *Semantic Networks and Automatic Text Processing*, Naukova Dumka, Kyiv, 1983, p. 217. (in Russian)

[3] D.V. Lande, *Knowledge Retrieving at the Internet. professional work*, Williams, Moscow, 2005, p. 272. (in Russian)

[4] D.V. Lande, A.A. Snarskiy, I.V. Bezsydnov, *Internetics. Navigation in the Complex Networks: Models and Algorithms*, Librikom, Moscow, 2009, p. 264. (in Russian)

[5] D.V. Lande, *The Base of Information Flows Integration*, Engeeniring, Kyiv, 2006, p. 240. (in Russian)

[6] Z.S. Syed, T. Finin, A. Joshi, Wikipedia as an ontology for describing documents, *In Proceedings of ICWSM*, AAAI Press, 2008, pp. 136-144.

[7] A.Y. Mykhailiuk, O.V. Pylypchuk, M.V. Snizhko, V.P. Tarasenko, Quasi-semantic search of textual data in an electronic information source, *Radioelectronics and Informatics*, KhNURE, Kharkov, 3 (2009), pp. 61-67. (in Ukrainian)

[8] D.J. Barrett, *MediaWiki*, O'Reilly Media, 2008, p. 384.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

МЕТОД ПОСТРОЕНИЯ УЛУЧШЕННЫХ ВЕЙВЛЕТОВ ПУТЕМ ПРЕОБРАЗОВАНИЯ ГРАФИКА СТЕПЕННОЙ ФУНКЦИИ ДЛЯ ЗАДАЧИ КОНТУРНОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Марина Полякова, Виктор Крылов, Наталья Волкова

Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, 65044, Украина,
marina_polyakova@rambler.ru

Резюме: *В данной работе разработан метод построения улучшенных вейвлетов для задачи контурной сегментации изображений путем преобразования графика степенной функции.*

Ключевые слова: *контурная сегментация изображений, улучшенные вейвлеты.*

THE METHOD OF WAVELETS CONSTRUCTION BY TRANSFORMATION OF A GRAPH OF POWER FUNCTION FOR EDGE DETECTION

Marina Polyakova, Victor Krylov, Natalya Volkova

Odessa National Politechnic University, 1, Shevchenko prospect, Odessa, 65044, Ukraine,
marina_polyakova@rambler.ru

Abstract: *In this paper the method to construct the improved wavelets by transformation of a graph of power function is developed for the edge detection.*

Keywords: *Underlining of edges, Edge detection, Improved wavelets.*

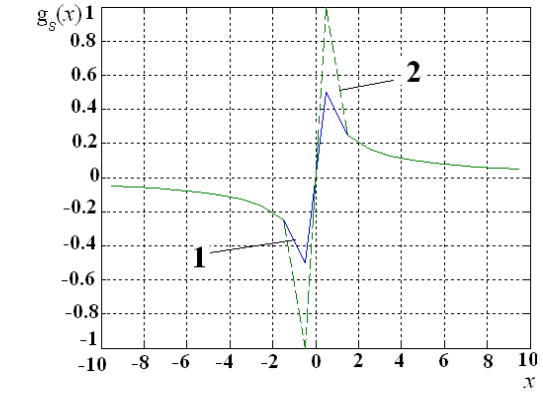
ВВЕДЕНИЕ

Процедура контурной сегментации изображений применяется в системах компьютерного распознавания зрительных образов для снижения объема обрабатываемой информации и обеспечения инвариантности к трансформациям интенсивности. Признаком, по которому производится разбиение изображения на однородные области, является наличие перепада интенсивности.

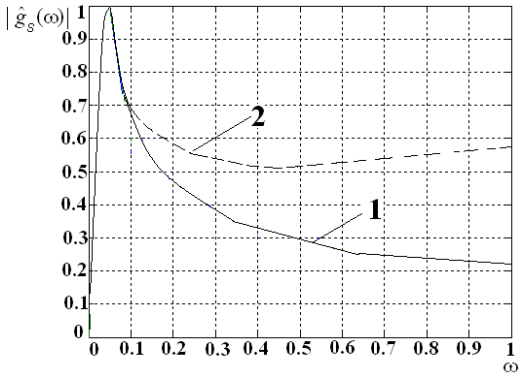
В значительной части встречающихся практических задач объекты распознавания имеют иерархическую структуру (объект-подобъект). Для таких задач или для задач анализа сцен необходимо различать объекты на изображении не только по величине перепада интенсивности, но и по их геометрическим размерам. В этом случае контурную сегментацию изображений целесообразно производить в пространстве вейвлет-преобразования (ВП).

При выборе вида ВП, как правило, учитываются требования к помехоустойчивости контурной сегментации и погрешности определения координат точек перепадов интенсивности. В [1] построена новая система вейвлет-функций на основе однородных обобщенных функций, позволяющая уменьшить погрешность определения координат точек перепадов интенсивности изображения в задаче контурной сегментации. В качестве способа построения новой системы вейвлет-функций был выбран лифтинг [2], использующий для определения пространственно-частотной локализации составляющих изображения исходную пространственную область. В результате с помощью лифтинга были модифицированы коэффициенты фильтра, полученного путем дискретизации степенной функции (рис. 1); на основе модифицированного фильтра построена система улучшенных вейвлет-функций для задачи контурной

сегментации изображений; предложен метод контурной сегментации изображений в пространстве ВП с использованием лифтинга на основе детектора Канни [3].



а



б

Рис. 1 – Импульсная (а) и амплитудно-частотная (б) характеристика исходного фильтра g_s (1) и улучшенного фильтра g_s^{new} (2)

Этот метод рекомендуется в задачах обработки изображений в случае, если помеховая ситуация не очень жесткая и необходимы методы контурной сегментации с низкой погрешностью определения координат точек перепадов интенсивности изображения.

Однако при повышении требований к помехоустойчивости контурной сегментации изображений использование системы улучшенных вейвлетов [1] нецелесообразно, т. к. расширение полосы пропускания фильтра (см. рис. 1, б) приводит к меньшему подавлению шума и более высокой детализации контурного препарата. Переход в пространство обобщенных функций ослабляет зависимость результата ВП от геометрических размеров объекта распознавания.

Заметим, что требования к помехоустойчивости сегментации и погрешности определения координат точек контуров объекта определяются как уровнем шума на изображении, так и геометрическими размерами объекта. Абсолютная погрешность определения координат точек контуров объекта чаще всего фиксирована для заданного метода сегментации и уровня шумов. Поэтому, если объект имеет большие геометрические размеры, то относительная погрешность определения координат точек границ объекта мала. В этом случае на результат распознавания объекта больше влияет разрыв контура или ложные контуры, т. е. результат распознавания объекта определяется помехоустойчивостью сегментации. Если геометрические размеры объекта распознавания малы, то относительная погрешность определения координат точек контуров объекта велика. Тогда результат распознавания объектов на изображении в основном характеризуется погрешностью определения координат точек контуров. Поэтому возникает необходимость в разработке в составе метода сегментации ВП, у которого при изменении параметра преобразования изменяется и характер обработки.

Заметим, что улучшенные вейвлеты для задачи контурной сегментации изображений по своим свойствам сходны с дифференциатором [1]. Этим и определяется низкая помехоустойчивость методов сегментации, использующих эти вейвлеты. Высокой помехоустойчивостью обладают корреляционно-экстремальные методы контурной сегментации изображений [4], в основе которых лежит согласованная фильтрация. Поэтому для реализации разного характера обработки изображений построим систему улучшенных вейвлетов, одни из которых проводят пространственное дифференцирование, другие – согласованную фильтрацию.

Конструирование системы вейвлетов с требуемыми характеристиками можно осуществлять на основе лифтинга. Последний представляет собой модификацию передаточных функций фильтров, соответствующих вейвлетам, за счет аддитивной составляющей [1]. Однако при реализации лифтинга возникают затруднения при выборе вида аддитивной составляющей передаточной функции фильтра, обеспечивающей улучшение свойств соответствующего вейвлета. Применение вместо лифтинга преобразований графиков степенных функций [5] упрощает процесс построения улучшенных вейвлетов.

Целью работы является повышение помехоустойчивости контурной сегментации изображений и снижение погрешности определения координат точек контуров объектов на разных уровнях иерархии за счет применения вейвлетов, построенных путем преобразований

графика степенной функции.

Для достижения поставленной цели решаются следующие задачи:

— разработан метод построения улучшенных вейвлетов для задачи контурной сегментации изображений путем преобразования графика степенной функции;

— с помощью вейвлетов, построенных путем преобразования графика степенной функции, система фильтров в составе подчеркивающего преобразования на основе лифтинга дополнена до согласованного фильтра;

— на основе детектора Канни предложен метод контурной сегментации изображений с применением вейвлетов, построенных путем преобразования графика степенной функции.

1. ЛИФТИНГОВАЯ СХЕМА ДЛЯ ЗАДАЧИ КОНТУРНОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

При решении задачи контурной сегментации изображений ВП используется в качестве преобразования, подчеркивающего перепады интенсивности изображения. Последнее обеспечивает помехоустойчивость сегментации и снижает погрешность определения координат точек контуров. Очевидно, что требования помехоустойчивости и низкой погрешности определения координат точек перепадов интенсивности изображения противоречивы, поэтому результат подчеркивающего преобразования представляется иерархией изображений с подчеркнутыми контурами объектов разного размера.

Для получения результата иерархического подчеркивающего преобразования используется свертка с каждой из функций $\{f_{i\min}(x), ..., f_i(x), ..., f_{i\max}(x)\}$, которые определены для всех вещественных x и $f_{i\min}(x) \to 2\theta(x)-1$ при $i_{\min} \to -\infty$, а $f_{i\max}(x) \to \delta'(x)$ при $i_{\max} \to \infty$ [6]. Если в качестве подчеркивающего преобразования изображения используется ВП, под системой функций $\{f_{i\min}(x), ..., f_i(x), ..., f_{i\max}(x)\}$ подразумевается множество вейвлет-функций. Чтобы удовлетворить предельному соотношению $f_{i\max}(x) \to \delta'(x)$ при $i_{\max} \to \infty$ для подчеркивающего преобразования и этим уменьшить погрешность определения координат точек перепадов интенсивности при увеличении масштаба ВП в [1] к вейвлет-функции применялся лифтинг.

В результате обобщенная вейвлет-функция на основе улучшенного фильтра g_s^{new} (см. рис. 1) определялась формулой

$$\psi_s^{new}(x) = \psi(sx) + \sum_{k=-\infty}^{\infty} l_s(k)\delta(x-k), \qquad (1)$$

где $\psi(x) = 1/x$, $l_s(k)$ – вещественные коэффициенты, $\delta(x)$ – дельта-функция. ВП с вейвлет-функциями (1) является нестационарным, т. к. на разных масштабах преобразования используются разные анализирующие вейвлеты.

Оценим поведение обобщенных вейвлет-функций (1), полученных на основе лифтинга, в контексте свойств иерархического подчеркивающего преобразования. По построению обобщенной вейвлет-функции на основе лифтинга параметр масштаба $s > 1$.

В процессе реализации подчеркивающего преобразования нас интересует поведение функций $\psi(sx)$ при $s \to \infty$ и фиксированном x. Эти функции монотонно убывают при $s \to \infty$ и фиксированном x, а также $\psi(sx) = \frac{1}{sx} \to +0$ $\left(\psi(sx) = \frac{1}{sx} \to -0\right)$, если $x > 0$ ($x < 0$). При $x \to 0$ в случае реализации подчеркивающего преобразования нас интересует поведение $\psi(sx) = \frac{1}{sx}$ при $s \to \infty$. Это неопределенное выражение типа $0 \cdot \infty$. Заметим, что $x \to 0$ быстрее, чем $s \to \infty$ при выполнении подчеркивающего преобразования. Последнее объясняется тем, что изображение сначала обрабатывается на одном уровне (s фиксировано, изменяется x), а затем происходит смена уровня обработки (изменяется s). Тогда $\psi(sx) \to +\infty$ $(\psi(sx) \to -\infty)$ при $x \to +0$ ($x \to -0$). Следовательно, $\psi(sx) \to \delta'(x)$ при $s \to \infty$. В случае $s \to 1$ $\psi(sx) = \frac{1}{sx} \to \frac{1}{x}$.

Сходимость $\psi(sx)$ к $\delta'(x)$ при $s \to \infty$ и к $\frac{1}{x}$ при $s \to 1$ поточечная и имеет место также для улучшенных вейвлетов $\psi_s^{new}(x)$ задачи контурной сегментации. Последнее справедливо для любого $x \in R$, $x = \pm 1$, т. к. по построению улучшенных вейвлетов в этих точках их значения стремятся к бесконечности.

2. ПОСТРОЕНИЕ ВЕЙВЛЕТОВ С ПОМОЩЬЮ ПРЕОБРАЗОВАНИЙ ГРАФИКА СТЕПЕННОЙ ФУНКЦИИ

Дополнение системы улучшенных вейвлетов $\psi_s^{new}(x)$ для задачи контурной сегментации изображений вейвлетами, реализующими согласованную фильтрацию, выполним на основе геометрического преобразования графика функции [7].

Под графиком функции понимается множество точек плоскости с прямоугольными координатами (x, y), где $y = f(x)$ – действительная функция одной действительной переменной x, которая принимает значения из области определения этой функции. С помощью линейных преобразований параллельного переноса, растяжения или сжатия, симметрии в некоторых случаях график функции можно построить по заданной его части или по графику другой функции. Так, графики функций $y = f(x + a)$ и $y = f(x) + b$ можно построить с помощью параллельного переноса вдоль оси Ox и Oy соответственно. С помощью растяжения или сжатия по оси Ox или Oy можно построить график функции $y = f(kx)$ и $y = mf(x)$. Чтобы построить график функции $y = mf(kx + a) + b$ последовательно выполняют перечисленные преобразования. Путем преобразования симметрии относительно биссектрисы первого координатного угла получают график обратной функции $y = g(x) = f^{-1}(x)$. График функции $y = -f(x)$ может быть получен из графика функции $y = f(x)$ отражением относительно оси Ox, а график функции $y = f(-x)$ – из графика функции $y = f(x)$ отражением относительно оси Oy. График функции $y = |f(x)|$ получается отражением относительно оси Ox частей графика $y = f(x)$ при $y < 0$.

С учетом приведенного анализа процесс построения улучшенных вейвлетов можно представить при помощи преобразований графиков степенных функций [5] следующим образом. График гиперболы $y = 1/x$ (рис. 2, а) преобразуется с помощью с помощью параллельного переноса вдоль оси Ox на $b_0 > 0$: $y_1 = 1/(x + b_0)$ (рис. 2, б). От графика $y_1 = 1/(x + b_0)$ оставляем фрагмент при $x > 1$, а от графика $y = x$ фрагмент при $x \in [0, 1]$ (рис. 2, в). Оба графика соединяются в точке разрыва и подвергаются отражению относительно оси Oy, а затем относительно оси Ox (рис. 2, г). График функции $y_1 = 1/(x + b_0)$ может дополнительно подвергаться растяжению или сжатию относительно оси Ox перед выделением фрагмента при $x > 1$. В результате преобразований графиков степенных функций получен график локально интегрируемой функции.

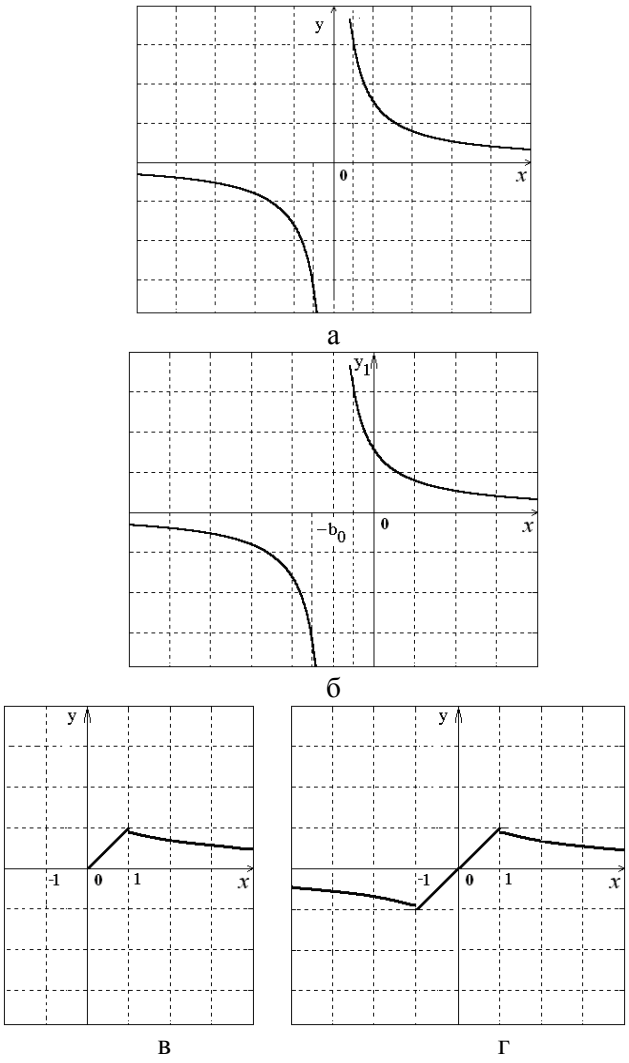


Рис. 2 – График гиперболы $y = 1/x$ (а), $y_1 = 1/(x + b_0)$ (б), построение вейвлет-функции (в, г)

На основе приведенной последовательности преобразований графиков степенных функций можно сформулировать метод построения улучшенных вейвлетов для задачи контурной сегментации изображений. Он заключается в следующем.

1. Выполняется сдвиг влево графика степенной функции на положительную величину, нелинейно зависящую от масштаба. Такое преобразование графика степенной

функции позволяет сконструировать улучшенные вейвлеты, способные выделять контуры объектов разных геометрических размеров с разной погрешностью определения координат точек контуров.

2. Функции, полученные на предыдущем этапе метода, подвергаются растяжению или сжатию с коэффициентом, равным значению масштаба преобразования. Это обеспечивает возможность обработки объектов разных геометрических размеров с разной помехоустойчивостью.

3. Фрагменты графиков полученных функций при $x < 0$ удаляются. Оставшиеся фрагменты графиков подвергаются преобразованию симметрии относительно оси Oy. Результат преобразования симметрии относительно оси Oy при $x < 0$ повторно подвергается преобразованию симметрии, но уже относительно оси Ox. Такие преобразования предполагают равноценность элементов строки (столбца) изображения относительно вертикальной (горизонтальной) оси.

В качестве примера применения метода построения вейвлетов с помощью преобразований графиков степенных функций покажем, что система вейвлет-функций [7]

$$\psi_s(x) = \begin{cases} \dfrac{1}{s(x+\dfrac{1}{s}-1)}, & \varepsilon \le x \le \gamma s, \\ \dfrac{1}{s(x-\dfrac{1}{s}+1)}, & \gamma s \le x \le -\varepsilon, \\ 0, & |x| > \gamma s, \quad |x| < \varepsilon. \end{cases} \qquad (2)$$

Рассмотрим случай $\psi(x) = \dfrac{1}{x}$. Пусть значения параметра масштаба $s < 1$. График гиперболы $y = 1/x$ (рис. 3, а) подвергаем преобразованию сдвига на $(1/s - 1) > 0$: $y_1 = 1/(x + 1/s - 1)$ (рис. 3, б). От графика $y_1 = 1/(x + 1/s - 1)$ оставляем фрагмент при $x > 0$, к которому применяем преобразование подобия с коэффициентом $1/s$ (рис. 3, в). Отображаем антисимметрично полученный график на отрицательную полуось Ox (рис. 3, г).

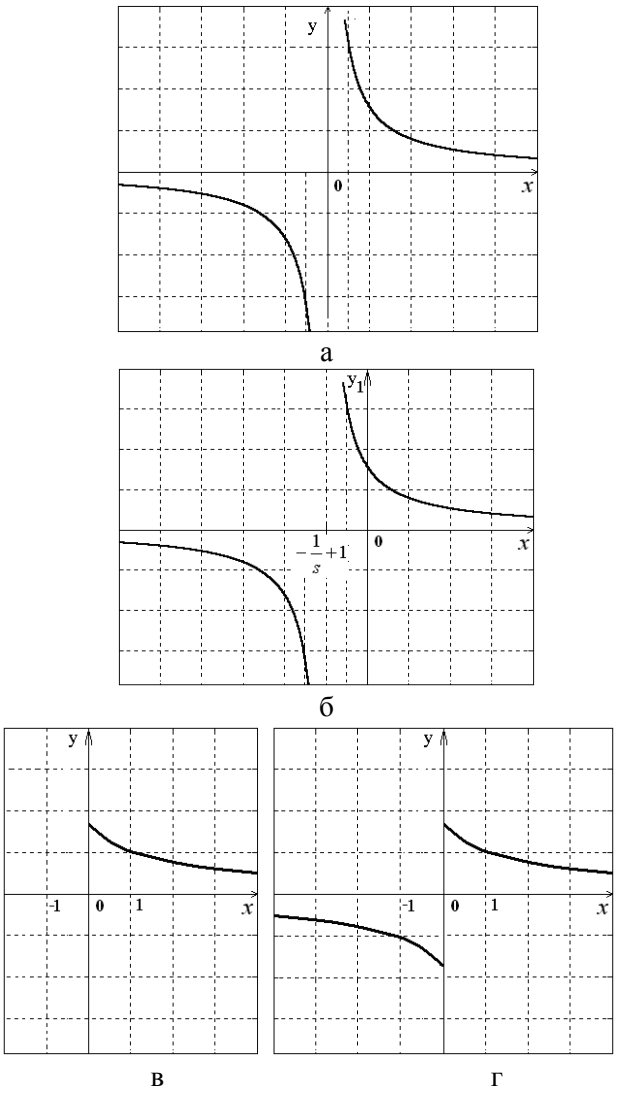


Рис. 3 – График гиперболы $y = 1/x$ (а), $y_1 = 1/(x + 1/s - 1)$ (б), построение вейвлета (в, г)

Полученные, в результате геометрических преобразований вейвлет-функции описываются формулой:

$$\psi_s(x) = \begin{cases} \dfrac{1}{s(x+\dfrac{1}{s}-1)}, & x > 0, \\ \dfrac{1}{s(x-\dfrac{1}{s}+1)}, & x < 0, \\ 0, & x = 0. \end{cases} \qquad (3)$$

Графики этих функций для разных значений s изображены на рис. 4.

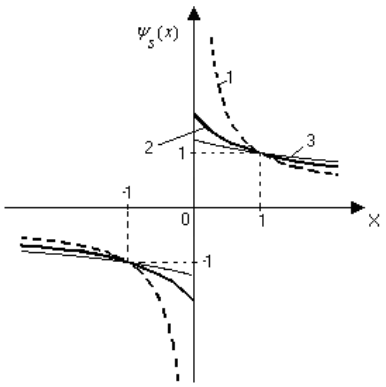


Рис. 4 – Графики вейвлет-функций, определяемых формулой (3) при $s = 1$ (1); $0{,}5$ (2); $0{,}25$ (3)

Система вейвлет-функций (3) отличается от (2) значениями в окрестности нуля и на бесконечности. Последнее может быть скорректировано заменой фрагментов графиков функций системы (3) в окрестности нуля и на бесконечности соответствующими фрагментами графика функции $y = 0$. Т. о. система вейвлет-функций (2) получена путем преобразований графиков степенных функций. Однако система вейвлет-функций (3) в отличие от (2) удовлетворяет требованиям помехоустойчивости и низкой погрешности определения координат точек перепадов интенсивности для иерархического подчеркивающего преобразования, т. е. $\psi_s(x) \to 2\theta(x) - 1$ при $s \to 0$ и $\psi_s(x) \to 1/x$ при $s \to \infty$.

Для того, чтобы это доказать, исследуем поведение функций $\psi_s(x)$ при $s \to 0$ и $s \to 1$. Если $x > 0$, то

$$\lim_{s\to 0}\psi_s(x) = \lim_{s\to 0}\frac{1}{s\left(x + \frac{1}{s} - 1\right)} = \lim_{s\to 0}\frac{1}{s(x-1)+1} = 1.$$

В случае $x < 0$

$$\lim_{s\to 0}\psi_s(x) = \lim_{s\to 0}\frac{1}{s\left(x - \frac{1}{s} + 1\right)} =$$

$$= \lim_{s\to 0}\frac{1}{s(x+1)-1} = -1.$$

Если $x = 0$, то $\lim_{s\to 0}\psi_s(x) = 0$. Тогда $\psi_s(x) \to \operatorname{sgn}(x)$ при $s \to 0$. При $s \to 1$ $\lim_{s\to 1}\psi_s(x) = \lim_{s\to 1}\frac{1}{s(x-1)+1} = \frac{1}{x}$, если $x > 0$, и также $\lim_{s\to 1}\psi_s(x) = \lim_{s\to 1}\frac{1}{s(x+1)-1} = \frac{1}{x}$, если $x < 0$.

Покажем, что функция $\operatorname{sgn}(x)$ представляет собой импульсную характеристику фильтра, согласованного с перепадом интенсивности изображения. Импульсная характеристика согласованного фильтра представляет собой масштабную копию входного сигнала, которая располагается в зеркальном порядке вдоль оси пространственной координаты [8]. Кроме этого импульсная характеристика согласованного фильтра смещена относительно зеркальной копии входного сигнала на величину x_0:

$$h_{\text{согл}}(x) = k s_{\text{вх}}(x_0 - x), \qquad (4)$$

где $h_{\text{согл}}(x)$ – импульсная характеристика согласованного фильтра, $s_{\text{вх}}(x)$ – входной сигнал, x_0, k – постоянные.

В качестве входного сигнала, соответствующего перепаду интенсивности изображения, выберем

$$s_{\text{вх}}(x) = a_0 + a_1\theta(-x), \qquad (5)$$

где $\theta(x)$ – функция Хевисайда, $\theta(x) = \begin{cases} 1, & x \ge 0, \\ 0, & x < 0. \end{cases}$ Функция $\operatorname{sgn}(x) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases}$ Тогда

$$\operatorname{sgn}(x) = 2\theta(x) - 1 \qquad (6)$$

всюду на R, кроме точки $x = 0$. Учитывая (4), (5), (6), получаем

$$h_{\text{согл}}(x) = \operatorname{sgn}(x)$$

всюду на R, кроме точки $x = 0$. При этом $k = 2$, $x_0 = 0$ в (4), а также $a_0 = -1$, $a_1 = 1$ в (5).

Таким образом, функция $\operatorname{sgn}(x)$ представляет собой импульсную характеристику фильтра, согласованного с перепадом интенсивности изображения на $R \setminus 0$. На значение интеграла свертки не влияет значение функции в одной точке $x = 0$. Следовательно, система вейвлет-функций (3) при $s \to 0$ сходится к импульсной

характеристике согласованного фильтра для перепада интенсивности изображения.

В результате равномерной дискретизации при $x > 0$ и $x < 0$ вейвлет-функции (3) получаем фильтр g_s^{new} с коэффициентами

$$g_s^{new} = \left\{ -\frac{1}{s \cdot k + 1}, ..., -\frac{1}{s \cdot 2 + 1}, -\frac{1}{s + 1}, -1, 1, \frac{1}{s + 1}, \frac{1}{s \cdot 2 + 1}, ..., \frac{1}{s \cdot k + 1} \right\}, \quad (7)$$

для которых предполагается, что $s < 1$. Этот фильтр является полосовым с подъемом в области верхних частот (рис. 5). Такое свойство фильтра позволяет регулировать соотношение между помехоустойчивостью и погрешностью определения координат точек перепадов интенсивности изображения в методах контурной сегментации [5].

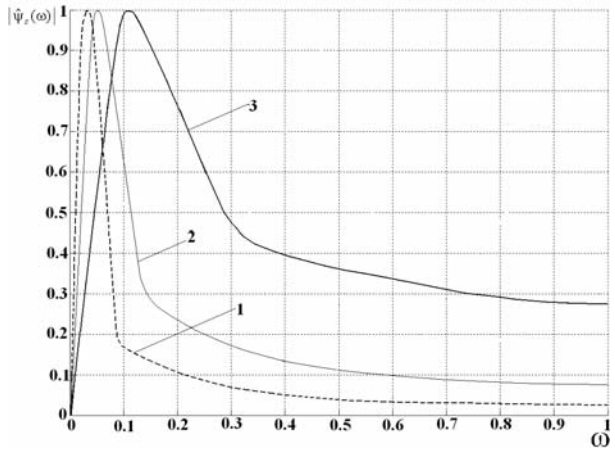


Рис. 5 – Амплитудно-частотная характеристика фильтра (7) при $s = 0{,}01$ **(1);** $0{,}1$ **(2);** $0{,}5$ **(3)**

3. МЕТОД КОНТУРНОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ С ПРИМЕНЕНИЕМ ВЕЙВЛЕТОВ, ПОЛУЧЕННЫХ ПУТЕМ ПРЕОБРАЗОВАНИЙ ГРАФИКА СТЕПЕННОЙ ФУНКЦИИ

В работе [1] сформулирован метод контурной сегментации изображений в пространстве ВП с использованием лифтинга с параметрами: отношение верхнего порога к нижнему и верхний порог, т. е. процент точек в распределении величин вейвлет-коэффициентов, которые будут включены в контурный препарат. Этот метод заключается

- задаются значения масштаба ВП s;
- для результата ВП с каждым значением масштаба s:
- вычисляются верхний и нижний пороги с учетом выбранных параметров;
- для каждой строки исходного изображения выполняется быстрое многофазное ВП с использованием лифтинга; получается контрастированное изображение $R_1(x, y)$ такого же размера, как и исходное;
- к контрастированным изображениям $R_1(x, y)$ и $R_2(x, y)$ применяется морфологическая обработка контура изображения метода Канни [3].

В данной работе метод контурной сегментации изображений в пространстве ВП с использованием лифтинга обобщен на случай применения вейвлетов, полученных путем преобразований графика степенной функции. В результате для каждой строки и каждого столбца изображения выполняется быстрое многофазное ВП с использованием лифтинга в случае, когда значение параметра масштаба $s > 1$. Если $s < 1$, то для каждой строки и каждого столбца изображения выполняется свертка с вейвлетами, полученными путем преобразований графика степенной функции. Остальные этапы метода работы [1] переносятся в более общий метод данной работы без изменений.

4. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ И ВЫВОДЫ

В ходе экспериментальных исследований оценивалась эффективная длительность подчеркнутого перепада интенсивности изображения, качество, помехоустойчивость и эффективность предложенного метода контурной сегментации изображений с применением вейвлетов, полученных путем преобразований графика степенной функции. Эффективная длительность подчеркнутого перепада интенсивности изображения – показатель, функционально связанный с погрешностью определения координат точек перепадов интенсивности. Он определялся по ослаблению модуля свертки фильтра с моделью перепада интенсивности $\theta(x)$ до 0,5 от максимального значения (табл. 1).

Показателем качества контурной сегментации изображения выбран показатель близости между контурами тестового идеально

сегментированного изображения $I^{эт}$ и изображения I^{t}, сегментированного исследуемым методом обработки [10]:

$$F = \frac{\sqrt{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(I^{t}(x,y) - I^{эт}(x,y)\right)^{2}}}{P}, \quad (8)$$

где P – длина выделенных контуров в пикселях, $M \times N$ – размеры изображения.

Таблица 1. Значения эффективной длительности подчеркнутого перепада интенсивности изображения для разных фильтров

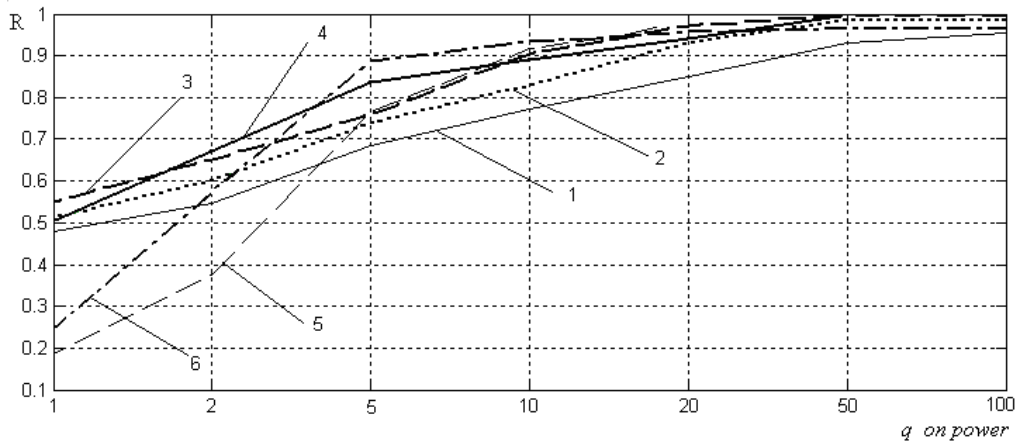
Фильтр	Эффективная длительность подчеркнутого перепада интенсивности изображения
g_s^{new}, s=0,01	83
g_s^{new}, s=0,02	67
g_s^{new}, s=0,05	45
g_s^{new}, s=0,1	33
g_s, s=1	7
g_s, s=2	3
g_s, s=4	1

Для оценки эффективности сегментации использовался показатель [10]

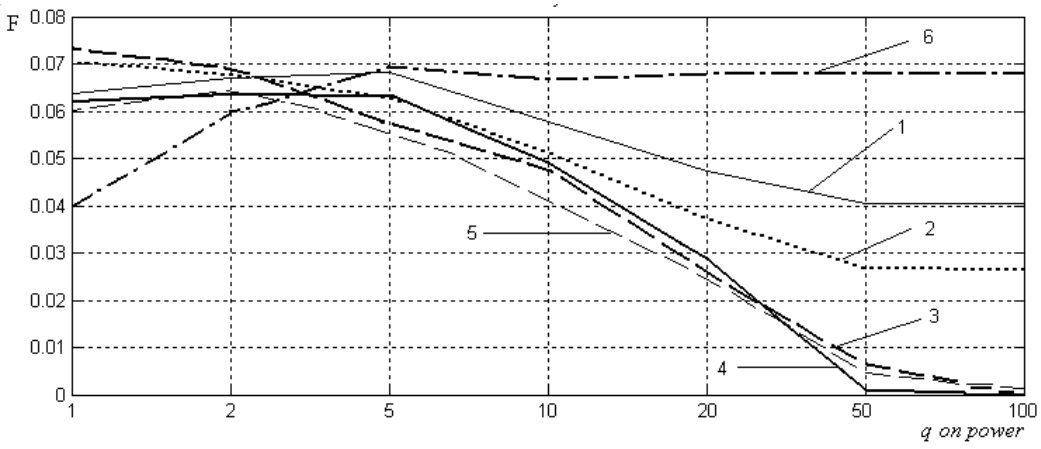
$$E = \frac{n \log_2 q}{k+1}, \quad (9)$$

где n – количество пикселей в обрабатываемом полутоновом изображении, q – количество градаций интенсивности, k – количество значащих пикселей в контурном препарате.

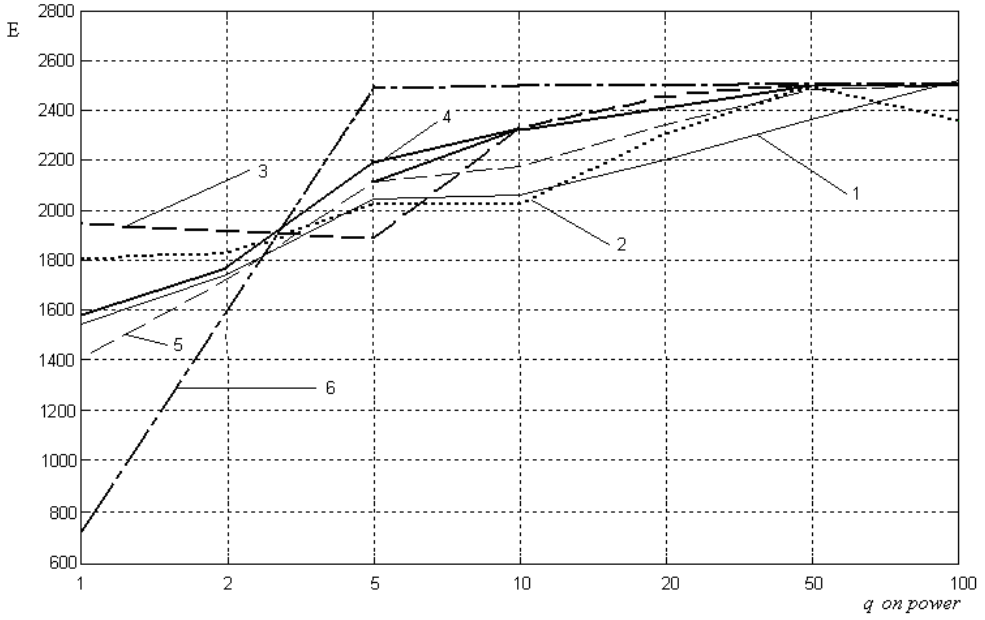
Получены графики зависимости значения критерия Прэтта и показателей (8), (9) от отношения сигнал/шум q по мощности для тестового изображения размером 256x256 пикселей, представляющего собой черно-белый перепад с наложенным на него аддитивным независимым гауссовским шумом (рис. 6). Отношение сигнал/шум по мощности определялось как $q = h^2 / \sigma_{вх}^2$, где h – высота перепада интенсивности; $\sigma_{вх}^2$ – дисперсия шума.



а



б



в

Рис. 6 – Зависимости критерия Прэтта (а), показателей качества (б) и эффективности сегментации (в) от отношения сигнал/шум q по мощности при s=0,01 (1); 0,02 (2); 0,05 (3); 0,1 (4); 0,5 (5) для предложенного метода и метода Канни (6)

Результаты исследований предложенного метода контурной сегментации изображений с применением вейвлетов, полученных путем преобразований графика степенной функции, показали, что при росте параметра преобразования *s* снижается до 1,6 раз эффективная длительность подчеркнутого перепада интенсивности изображения, функционально связанная с погрешностью определения координат точек перепадов интенсивности. При этом показатель помехоустойчивости при значениях отношения сигнал/шум 2 и менее по мощности снижается до 3 раз, а по сравнению с методом Канни до 2,2 раз. Показатель качества с ростом *s* ухудшается до 8 раз при значениях отношения сигнал/шум более 5 по мощности, а по сравнению с методом Канни до 14 раз. По эффективности сегментации с ростом параметра *s* получены сходные результаты.

Таким образом, разработанный метод контурной сегментации изображений с применением вейвлетов, полученных путем преобразований графика степенной функции, можно рекомендовать в задачах обработки изображений с повышенными требованиями к помехоустойчивости, в случае, если необходимы методы контурной сегментации с регулируемой детализацией объектов изображения.

Приведенные результаты исследований позволяют разработчику систем компьютерного распознавания зрительных образов в зависимости от свойств изображений и целей обработки выбрать адекватный вид вейвлет-преобразования и значение его параметра.

5. СПИСОК ЛИТЕРАТУРЫ

[1] V. Krylov, M. Polyakova, Contour segmentation of images in space of wavelet transform with the use of lifting, *Optoelectronic Information-Power Technologies*, 12 (2006), pp. 48-58. (in Russian)

[2] S. Mallat, *Wavelets in Signal Processing*, Moscow, Mir publishers, 2005, 671 p. (in Russian)

[3] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8 (1986), p. 679-698.

[4] V. Krylov, M. Maksimov, *The Second Transformers of Images Signals*, Odessa, Astroprint, 1997, 176 p. (in Russian)

[5] M. Polyakova, V. Krylov, Classification of methods of signal semantic wavelet transform for image contour segmentation, *International Journal of Computing*, (7) 1 (2008), pp. 51-57.

[6] M. Polyakova, V. Krylov, Hierarchical signal-semantic transform in the distribution space for the image segmentation, *Odes'kyi Politechnichnyi Universytet. Pratsi,* 26 (2006), pp. 161-167. (in Russian)

[7] *Mathematical Encyclopedic Dictionary* / Yu. Prokhorov; S. Adyan, N. Bahvalov, V. Bityuckov, A. Ermov, L. Kudryavcev, A.

Onischuk, A. Yushkevich (eds). Moscow, Soviet encyclopedia, 1988, 847 p. (in Russian)
[8] O. Babilunga, *Iterative methods of hierarchical contour segmentation of images in the space of hyperbolic wavelet transform*, Avtoref. dis... kand. techn. sciences: 05.13.23. Odessa, Odessa National Polytechnical University, 2008. 20 p. (in Ukrainian)
[9] S. Baskakov, *Radiotechical Chains and Signals*, Moscow, Vysshaya shkola, 1988, 488 p. (in Russian)
[10] V. Abakumov, V. Krylov, S. Antoschuk, Increase of efficiency of processing of pattern information in automatic sysems, *Electronics and Communications. Thematic issue of «Problems of electronics»*, 1 (2005), pp. 100-105. (in Russian)
[11] W. Pratt, *Digital Image Processing*, Moscow, Mir publishers, 1982, 790 p. (in Russian)

Полякова Марина Вячеславовна – *специалист (1994), прикладная математика, Одесский государственный университет, к.т.н. (2004), автоматизированные системы управления и прогрессивные информационные технологии, доцент кафедры «Прикладная математика и информационные технологии в бизнесе» (2006), Одесский национальный политехнический университет.*

Научные интересы: вейвлет-анализ, фракталы, теория обобщенных функций, функциональный анализ.

Крылов Виктор Николаевич – *специалист (1978), радиотехника, Одесский политехнический институт, к.т.н. (1986), радиотехнические и телевизионные системы и устройства, д.т.н. (2003), автоматизированные системы управления и прогрессивные информационные технологии, профессор кафедры «Прикладная математика и информационные технологии в бизнесе» (2005), Одесский национальный политехнический университет.*

Научные интересы: цифровая обработка изображений, распознавание образов.

Волкова Наталья Павловна – *специалист (1992), механика, Одесский государственный университет, старший преподаватель кафедры «Прикладная математика и информационные технологии в Бизнесе» (2004), Одесский национальный политехнический университет.*

Научные интересы: цифровая обработка изображений, распознавание образов.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

THE METHOD OF WAVELETS CONSTRUCTION BY TRANSFORMATION OF A GRAPH OF POWER FUNCTION FOR EDGE DETECTION

Marina Polyakova, Victor Krylov, Natalya Volkova

Odessa National Politechnic University, 1, Shevchenko prospect, Odessa, 65044, Ukraine,
marina_polyakova@rambler.ru

Abstract: *In this paper the method to construct the improved wavelets by transformation of a graph of power function is developed for the edge detection.*

Keywords: *Underlining of edges, Edge detection, Improved wavelets.*

In considerable part of computer recognition of visual patterns the objects of recognition have a hierarchical structure (object-subobject). For such problems or for the scene analysis it is necessary to distinguish objects on an image not only on the strength of edge but also on geometrical size of objects. In this case it is expedient to make edge detection of images in space of wavelet transform (WT) coefficients.

In this paper the noise stability of edge detection of images is increased and the error of determination of co-ordinates of position of objects edges on the different levels of hierarchy is lowered due to application of wavelets built by transforms of graph of power function. The method of construction of improved wavelets by transform of graph of power function is developed for the edge detection. The function graph is understood as a points set of a plane with rectangular co-ordinates (x, y), where $y = f(x)$ is a real-valued function of one real variable x which accepts values from a range of definition of this function [1] means of linear transforms of parallel shifting, scaling, and symmetry in some cases the function graph can be constructed with its parts or under a drawing of other function.

On the basis of the parallel shifting, scaling, and symmetry of graphs of power functions it is possible to formulate the method of construction of improved wavelets for the edge detection. It consists in the following.

1. To the graph of power function the left-shift on the positive value nonlinearly depending on a scale is applied. Such transform of the graph of power function allows constructing improved wavelets underlines the edges of objects of different geometrical size with a different error of determination of co-ordinates of position of edges.

2. The functions got at the previous stage of a method are scaled with the factor equal to value of a scale of transform. It ensures a possibility of processing of objects of different geometrical size with different noise stability.

3. Fragments of graphs of the got functions at x <0 leave. The remained fragments of graphs are transformed by symmetry to axis Oy. The outcome of symmetry to axis Oy at x <0 repeatedly is exposed to symmetry transform, but already concerning an axis Ox. Such transforms assume an equivalence of pixels of a line (column) of the image concerning a vertical (horizontal) axis.

The method of construction of wavelets by transforms of graph of power functions [1] applied on the wavelets

$$\psi_s(x) = \begin{cases} \dfrac{1}{s(x+1/s-1)}, & x > 0, \\ \dfrac{1}{s(x-1/s+1)}, & x < 0, \\ 0 \qquad\qquad , & x = 0. \end{cases} \qquad (1)$$

Let's consider the case $\psi(x) = 1/x$. Let values of parameter of scale $s < 1$. The graph of hyperbola $y = 1/x$ (Fig. 1, a) we shift on $(1/s - 1) > 0$: $y_1 = 1/(x + 1/s - 1)$ (Fig. 1, b). From graph $y_1 = 1/(x + 1/s - 1)$ we leave a fragment at $x > 0$ to which we apply a scaling with factor $1/s$ (Fig. 1, c). We map antisymmetrically the resulted graph on negative semiaxis Ox (Fig. 1, d).

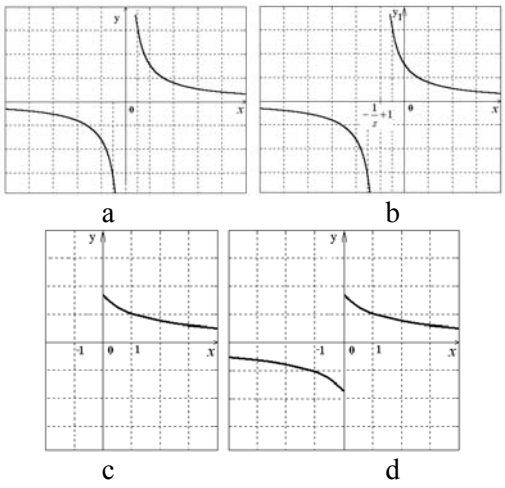


Fig. 1 – the graph of hyperbola $y = 1/x$ (), $y_1 = 1/(x+1/s-1)$ (b), construction of wavelet (c, d)

The edge detector on images based on the Canny detector with the use of wavelets built by transform of the graph of power function is proposed.

Wavelets got as a result of geometrical transforms are described a formula:

$$\psi_s(x) = \begin{cases} \dfrac{1}{s(x+\dfrac{1}{s}-1)}, & x > 0, \\ \dfrac{1}{s(x-\dfrac{1}{s}+1)}, & x < 0, \\ 0, & x = 0. \end{cases} \qquad (2)$$

Graphs of these functions for different values s are represented on Fig. 2.

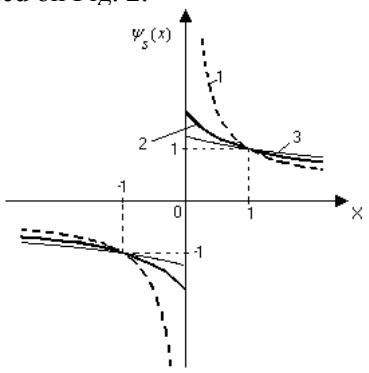


Fig. 2 – Graphs of the wavelets defined by the Eq. 3 at $s = 1$ (1); $0{,}5$ (2); $0{,}25$ (3)

The wavelets from Eq. 2 differs from Eq. 1 by values in a neighbourhood of zero and on infinity. The last can be corrected by replacement of fragments of graphs of functions of Eq. 1 in a neighbourhood of zero and on infinity by corresponding fragments of the graph of function $y = 0$. Therefore the wavelets from Eq. 1 is got by transforms of graphs of power functions. However the wavelets from Eq. 2 unlike Eq. 1 satisfies to requirements of a noise stability and a low error of determination of co-ordinates of position of edges for hierarchical underlining transform, i.e. $\psi_s(x) \to 2\theta(x) - 1$ at $s \to 0$ and $\psi_s(x) \to 1/x$ at $s \to \infty$.

As a result of uniform discretization at $x > 0$ and $x < 0$ of wavelets from Eq. 2 we get filter g_s^{new} with coefficients

$$g_s^{new} = \left\{ -\frac{1}{s \cdot k + 1}, \ldots, -\frac{1}{s \cdot 2 + 1}, -\frac{1}{s+1}, -1, 1, \frac{1}{s+1}, \frac{1}{s \cdot 2 + 1}, \ldots, \frac{1}{s \cdot k + 1} \right\}, \qquad (3)$$

for which it is supposed, that $s < 1$. This is band filter with lifting of the upper frequencies (Fig. 3). Such property of the filter allows to regulating a relation between noise stability and an error of definition of co-ordinates of position of edges of the image in the methods of edge detection [2].

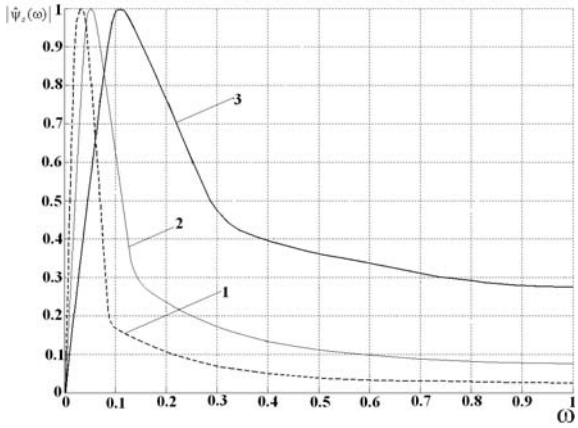


Fig. 3 – Frequency response of filter (3) at (1); (2); (3)

Thus elaborated edge detector for images with the use of wavelets got by transforms of graph of power function it is possible to recommend in processing of images with the high requirements to noise stability in the case if the edge detector with the different processing of objects details on image is needed.

REFERENCES

[1] *Mathematical Encyclopedic Dictionary*, edited by F. V. Prokhorov, Moscow, Soviet encyclopedia, 1988.

[2] M. Polyakova, V. Krylov, Classification of methods of signal semantic wavelet transform for image contour segmentation, *International Journal of Computing*, (7) 1 (2008), pp. 51-57.

[3] V. G. Abakumov, V. N. Krylov and S. G. Antoschuk, Increasing of efficiency of processing of pattern information in the automated systems, *Electronics and communication: «Problems of electronics»*, 1 (2005), pp. 100-105.

[4] W. K. Pratt, *Digital Image Processing*, New York, John Wiley and Sons, 1991.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

INTEGRATED EFFECT OF DATA CLEANING AND SAMPLING ON DECISION TREE LEARNING OF LARGE DATA SETS

Dipak V. Patil [1)], Rajankumar S. Bichkar [2)]

[1)] Department of Computer Engineering
Sandip Institute of Technology and Research Centre, Nasik, M.S., India.
dvspatil@yahoo.co.in
[2)] Department of Computer Engineering
G.H. Raisoni College of Engineering & Management, Pune, M.S., India.
bichkar@yahoo.com

Abstract: *The advances and use of technology in all walks of life results in tremendous growth of data available for data mining. Large amount of knowledge available can be utilized to improve decision-making process. The data contains the noise or outlier data to some extent which hampers the classification performance of classifier built on that training data. The learning process on large data set becomes very slow, as it has to be done serially on available large datasets. It has been proved that random data reduction techniques can be used to build optimal decision trees. Thus, we can integrate data cleaning and data sampling techniques to overcome the problems in handling large data sets. In this proposed technique outlier data is first filtered out to get clean data with improved quality and then random sampling technique is applied on this clean data set to get reduced data set. This reduced data set is used to construct optimal decision tree.*

Experiments performed on several data sets proved that the proposed technique builds decision trees with enhanced classification accuracy as compared to classification performance on complete data set. Due to use of classification filter a quality of data is improved and sampling reduces the size of the data set. Thus, the proposed method constructs more accurate and optimal sized decision trees and it also avoids problems like overloading of memory and processor with large data sets. In addition, the time required to build a model on clean data is significantly reduced providing significant speedup.

Keywords: *Large data sets, decision tree, data cleaning, data sampling, speedup and classification accuracy.*

1. INTRODUCTION

The volume of data in databases is growing to large sizes, both in the number of attributes and instances. Data mining provides tools to inference knowledge from databases. This knowledge is used to boost the businesses. Data mining on a very large data set may overload a computer systems memory and processor making the learning process very slow. Data sets used for inference may be very large, may be up to terabytes. The data mining methods are faster when used on smaller data sets. T. Oates, D. Jensen [1] proved that removing randomly selected training instances;often results in smaller trees that are equally accurate to those built on all available training instances. Random data reduction technique is one of the solutions to handle the large data sets. [2].

The noise or outliers in data also affect the classification performance of the classifier on that data. An outlier is an instance that is significantly divergent with rest of the data set. [3]. Gamberger and Lavarac [4] suggested that effect of erroneous data on hypothesis could be avoided by eliminating it from training data before induction.

Data cleaning and sampling reduces time complexity of decision tree learning. Let n be the number of training instances and m be the number of attributes in a instance, computational cost of building tree is $O(m\ n\ log\ n)$ [5]; as n is reduced due to sampling and filterring, the corresponding cost is reduced. Similarly the cost of decision tree pruning process is $O(n(logn)^2)$ [5]. Due to data cleaning operation overfitting can be reduced and as erroneous data is removed the time complexity of pruning process can be reduced significantly.

The proposed technique integrates data cleaning and data sampling techniques to overcome these problems. The classification filter is used to filter training data to improve data quality, subsequently incremental random sampling is applied on this filtered data. These cleaned and sampled data sets

are used to learn classifiers.

1.1 Decision Tree Construction

Decision tree is a classifier in the form of tree that contains a decision node and leaves. A decision node specifies a test to be carried on single attributes value. A leave specifies a class. A solution is at hand for each probable outcome of the test in the form of child node. A tree is traversed from root to node to find out the class of an instance. A performance measure of a decision tree is the number of correctly classified instances called classification accuracy of the tree. It is defined in terms of the percentage of correctly classified instances. [6] - [8]. A decision tree algorithms construct accurate decision trees for classification, but they often experience the drawback of excessive complexity that can make them incomprehensible to human experts [9].

Hybrid learning methodologies that integrate genetic algorithms (GAs) and decision tree learning for evolving optimal decision trees have been proposed by different authors. Although the approaches are different the objective is to obtain optimal decision trees. The decision tree is called optimal if it is accurate and has minimum number of leafs. The GAIT algorithm proposed by Z. Fu [10] proposed generation of the set of diverse decision trees from different subsets of the original data set by using a decision tree algorithm C4.5, on small samples of the data. These decision trees are used as the initial populations to genetic algorithm. The fitness criterion for evaluation is the classification accuracy on test data. A. Papagelis, and D. Kalles proposed GATree, a genetically evolved decision tree [11]. The Genetic Algorithm is used to directly evolve binary decision trees. Decision trees those have one decision node with to two different leaves are operated by genetic operators. The constructed trees are called genetically evolved decision trees. Similar approaches are available in [12] - [13].

2. RELATED WORK

The proposed approach is combination of two techniques first is data cleaning and second is data sampling, these techniques are presented in two separate subsections.

2.1 DATA CLEANING

The outlier detection and noise elimination is an important issue in data analysis. The exclusion of outliers improves data quality and therefore classification performance. Several researchers have proposed various approaches for data cleaning. G.H. John [14] proposed a technique that removes a misclassified training instance from training data and reconstructs the trees, the process is repeated till all such instances are removed from training data. Misclassified instances are recognized using tree classifier as a filter. The resulting classifier enhances classification performance accuracy. Broadly and Friedl [15] proposed a method for detecting mislabeled instances. The method uses a set of learning algorithms to construct classifiers that act as a filter for the training data. The technique removes outliers in regression analysis. Arning et al. [16] proposed framework for the problem of outlier detection. Similar to human beings, it observes all instances for similarity with data sets and it treats dissimilar data set as an exception. A dissimilarity function is used to find out outliers.

Raman and Hellerstein [17] proposed an interactive framework for data cleaning that integrates transformation and discrepancy detection. Guyon et al [18] proposed training of convolutional neural networks with local connections and shared weights. Gamberger and Lavrac [19] proposed conditions for Occam's razor applicability in noise elimination. Knorr and Ng [20] proposed unified outlier detection system. Subsequently, they proposed and analyzed some algorithms for detecting distance-based outliers. Tax and Duin [21] proposed outlier detection that is based on the instability of the output of simple classifiers on new objects. Gamberger et al [22] proposed saturation filter. It is based on principle that detection and removal of noisy instances from training data induces less complex and more accurate hypothesis. The saturated training data set can be employed for induction of stable target theory.

Schwarm and Wolfman [23] proposed Bayesian methods for data cleaning which detects errors and corrects them. Ramaswamy et al [24] proposed algorithm for distance-based that ranks each point based on its distance to its nearest neighbor. Kubika and Moore [25] presented system for learning explicit noise. The system detects corrupted fields and uses non-corrupted fields for consequent modeling and analysis. Verbaeten and Assche [26] proposed three ensemble based methods for noise elimination in classification problems. Loureiro et al [27] proposed a method that applies hierarchical clustering methods for outlier detection. Xiong et al. [28] proposed a hyperclique-based noise removal system to provide superior quality association patterns. Patil and Bichkar [29] proposed use of evolutionary decision tree as classification filter and found that, with use of genetic algorithm optimal trees can be built.

2.2 DATA SAMPLING

Sampling is the procedure to obtain a subset of

instances that represents the entire data set. It is necessary to have sufficient sample size to validate statistical analysis. Sampling is done because it is impracticable to test every single individual in the data set. Moreover it saves time, resources and effort while executing the research. The representativeness is the most important issue in statistical sampling. The sample obtained from the set of data instances must be representative of the same. Probability sampling and Non-probability sampling are two types of sampling. In Probability sampling, all the instances in set have equal probability of being selected. The approach assures completely randomized selection process which is unbiased. The hypothesis is accurate when this sampling is used. In Non-Probability sampling all the instances in data set do not have equal probability of being selected. Thus sample does not completely represent the target data set. The representativeness can be achieved by using simple randomized statistical sampling techniques.

Jenson and Oates [1] experimented with random data sampling and proved that as size of the training dataset increases size of tree also increases where as classification accuracy does not increase significantly.

S. Vucetic and Z. Obradovic [30] proposed an effective data reduction method founded on guided sampling for determining a minimal size representative subset, it was followed by a model-sensitivity analysis for determining a suitable compression level for each attribute.

A. Lazarevic and Z. Obradovic [31] proposed several efficient techniques based on the idea of progressive sampling. The sampling process combines all the models constructed on previously considered data samples. The authors also proposed controllable sampling based on the boosting algorithm, where the models are combined using a weighted voting. Another contribution by authors is sampling procedure for spatial data domains, where the data instances are selected in accordance with the performance of previous models as well as in accordance with the spatial correlation of data.

Patil and Bichkar [32] proposed use of evolutionary decision tree along with random sampling of data to optimize the decision trees and concluded that the proposed method builds trees that are accurate and relatively smaller in size.

3. PROPOSED MODEL

The proposed model is based on combination of removal of outliers from data as first stage and incremental random sampling of data to evolve decision trees as second stage to obtain compact representation of large data set. The data set is first filtered using classification filter and then sampled randomly in different subsets by random sampling, initially, sample size is 5% and reaches up to 100% of the instances available in clean data set. The size of training data set is increased by 5% each time using random selection of instances and using this data tree are evolved. We have set 5% sample size as minimum Threshold size. The tree is first on built sample of size 5% of clean data and the classification accuracy on clean sampled data is compared with accuracy on complete clean data set. If classification accuracy on sample data is equal or greater than accuracy complete data then, the sample size is representative of complete data set, otherwise next incremental sample size is 10% of data and so on.

Let T_F be a set of all available n training instances classification filter is applied and we get clean data set T_C and unclean data set T_E. Let X_C is classification accuracy on T_C. Let a training instance be denoted by I and let the sampling size Threshold be denoted by α, and $\alpha = 5$. Random sampling is done on T_C, let T_i be subset of instances in T_C where $T_i \in T_C$ and $i = 1$ to 20, let hypothesis H_i is induced on T_i and classification accuracy on this training data be X_i.

$\forall T_i$ on H_i if $X_i \geq X_C \Rightarrow T_i$ is Final training set for constructing classifier and is denoted as reduced data set T_R and accuracy on T_R is denoted by X_R.

The proposed algorithm works as follows**.**

1. Induce*(H, T_F)*.
2. Classify *(H, T_F)*. // Classification Filter
3. $\forall I$ in T_F If $X(I) = 1 \Rightarrow I \in T_C$.
4. Else del*(I)*.
5. Sample*(T_C, T_i)* //Random sampling on T_C.
6. For *($i = 1$; $i \leq 20$; $i = i+1$)*
7. Induce*(H_i, T_i)*.
8. Compare*(X_i, X_C)*.
9. If $X_C \geq X_i \Rightarrow T_i$ is final reduced data set T_R.
10. Induce hypothesis H_R on T_R with decision tree as a final classifier.
11. Else repeat step 7 to 10 with next incremental sample until we get $X_C \geq X_i$.
12. End.

4. METHOD OF EXPERIMENTATION

Prior to applying the proposed technique on large data sets, we found it appropriate to first test it on the normal sized benchmark data set from UCI repository [33]. Experiments were performed on 24 benchmark data sets to explore classification accuracy of decision tree at reduced training data using proposed approach and results are validated using standard decision tree algorithm J48, CART [34] and GATree[11]. A significant enhancement in

classification performance on all data sets is observed, in order to make the presentation concise we are presenting only few cases with lower, moderate and higher classification accuracy on sample basis; these results are presented in Table 1, 2 and 3 respectively. For these data sets value of α = 25; i.e. initial sample size is 25% as data sets are smaller to midsize data sets. Whenever average results are presented in result discussion, it means it is average of all 24 data sets.

In test method, the data set is first filtered using classification filter. Then clean data is incrementally sampled in different subsets and trained using decision trees algorithms until we get classification accuracies comparable to that obtained on complete data set. The same method without filtering data is also followed in experimentation to get results with unclean data.

In order to analyse effect of cleaning on data and effect of sampling on data we have separate subsections followed by analysis on combined effect of cleaning and sampling on data by proposed method.

4.1 EFFECT OF CLEANING ON DATA

Data cleaning process improves the quality of data. This section presents effect of data cleaning on classification performance of the classifier on clean data sets. To study the effect, the complete data set T_F is filtered using classification filter and clean data T_C is obtained. The classification accuracies on data sets T_F and T_C are obtained using k-fold cross validation method and are presented in Table 1; X_F is classification accuracy on T_F. It is observed that for J48 classifier, classification accuracies on cleaned data sets in Table 1 are around 99% except Monks data set. Monks data set is having lower classification accuracy on complete data set T_F amongst all, it is 70.16% and classification accuracy on cleaned data is 90.16% which is a significant enhancement in classification performance on cleaning data. Similar results are observed on CART classifier. In case of GATree classifier, exceptional case is data set Mfeat-Factor, it has lower classification accuracy on complete data set which is 38.80% and is significantly enhanced to 80.67% with 41.87% enhancement in accuracy. The enhancement in accuracy ΔX is the absolute difference in accuracy between the tree build from the cleaned data set and the tree build from complete data training data which is unclean data. The average results for enhancement in accuracy (average ΔX) due to data cleaning for all 24 data sets are 8.87%, 8.43%, and 25.49% for J48, CART and GATree respectively. The previous results available so far by G. H. John [12] indicate enhancement in accuracy on J48 by 2% to 4% in average whereas here we have enhancement of around 8% in accuracy. Thus, with data cleaning with proposed we get enhanced classification performance, the reason is removal of anomalies in data.

Table 1. Effect of cleaning on data

Sr. No.	Data Set	J48			CART			GATree		
		X_F	X_C	ΔX	X_F	X_C	ΔX	X_F	X_C	ΔX
1	Australian	90.72	100.00	9.28	91.01	100.00	8.99	87.971	100.00	12.02
2	Breast –w	94.85	98.21	3.36	96.42	98.51	2.09	94.10	99.40	5.30
3	German	79.80	99.72	19.92	85.50	99.44	13.94	70.20	99.58	29.38
4	kr-vs-kp	99.41	100.00	0.59	99.62	100.00	0.38	91.05	98.55	7.50
5	Mfeat-Factor	93.50	99.73	6.23	93.25	98.66	5.41	38.80	80.67	41.87
6	Monks	70.16	90.16	20.00	64.52	93.44	28.92	70.00	90.00	20.00

4.2 EFFECT OF SAMPLING ON DATA

As it is impracticable to test every single individual instance in the data set, tests are conducted on samples of data. In this section we present analysis effect of sampling on data, and hence data cleaning process is not exercised here. The results are presented in Table 2. The complete data set T_F is sampled. Initial sample size is 25% of data, train decision tree on it, if classification accuracy on this classifier is equal or more than classification accuracy on complete dataset; the sample size is final sample size T_{RU}. Otherwise increment sample size by 5% and repeat the process. The data is incrementally sampled in different subsets and trained using decision trees algorithms successively until we get classification accuracies comparable to that obtained on complete data set. The sample size is abbreviated as SS and classification accuracy on reduced unclean data set T_{RU} is abbreviated as X_{RU} in Table 2. Here the enhancement in accuracy ΔX is the absolute difference in accuracy between the tree build from the reduced unclean data set and the tree build from complete data training data which is unclean data.

The average classification accuracies on all 24 complete datasets T_F are 90.02%, 89.31% and 72.25% as compared to accuracy of 90.32%, 90.05% and 74.75% on sampled data set T_{RU} for J48, CART

and GATree classifiers respectively. Thus, with data sampling, we get comparable accuracies on reduced data set.

The average percentage sample sizes for all 24 data set are 72.70%, 74.17% and 33.33% on complete data sets for J48, CART and GATree classifiers respectively. The sample size on J48 and CART are similar in average, whereas it is significantly smaller for GATree. Although average sample size is between 72% to 74% for J48 and CART, sampling was not so successful on some unclean data sets. The exceptional cases are, Monks data set on J48 classifier, German and Mfeat-Factor data sets on CART. These data sets could get required comparable accuracy with 100% data samples on unclean data set. The anomalies present in the data set are the most probable reason for it.

The sample size required for GATree is smaller because, GAtree incorporates genetic algorithm for global search in the problem space with classification performance in terms of accuracy as fitness function without being biased towards a local optimum and the sample size required is optimised. Thus we get reduced data set with comparable classification performance.

Table 2. Effect of sampling on data

Sr. No.	Data Set	J48				CART				GATree			
		X_F	X_{RU}	ΔX	SS	X_F	X_{RU}	ΔX	SS	X_F	X_{RU}	ΔX	SS
1	Australian	90.72	90.72	0.00	85	91.01	91.52	0.51	65	87.97	87.50	-0.47	35
2	Breast –w	94.85	95.22	0.37	45	96.42	97.42	1.00	50	94.10	97.65	3.55	25
3	German	79.80	80.15	0.35	65	85.50	85.50	0.00	100	70.20	74.40	4.20	25
4	kr-vs-kp	99.41	99.30	-0.11	40	99.62	99.43	-0.19	55	91.05	91.07	0.02	25
5	Mfeat-Fact.	93.50	93.00	-0.50	75	93.25	93.25	0.00	100	38.80	41.00	2.20	25
6	Monks	70.16	70.16	0.00	100	64.52	67.50	2.98	65	70.00	70.00	0.00	90

4.3 EFFECT OF CLEANING AND SAMPLING ON DATA

In the experimentation with the proposed approach, the data is first filtered and then sampled incrementally and trained using decision trees algorithms until we get classification accuracies comparable to that obtained on clean data set T_C. The comparison of accuracies on complete data set T_F with cleaned and reduced data set T_{RC} is presented here in Table 3. The enhancement in accuracy ΔX is the absolute difference in accuracy between the tree build from the reduced clean data set and the tree build from complete data training data which is unclean data.

The accuracies on clean and reduced data set T_{RC} are around 99% for all data set in Table 3 for J48 and CART. Exceptional case is Monks data set, where we could get around 93% accuracy with enhancement in accuracy of around 23% and 28% for J48 and CART respectively. Monks data set has lower accuracy on T_F hence, there is scope for enhancement. The data set kr-vs-kp is having around 99% accuracy on T_F and hence there is very less scope for enhancement in classification performance on the classifiers J48 and CART.

Similarly in case of GATree accuracies on reduced data sets are 99% on all data sets in table with exceptional case Mfeat-Factor data, where the minimum accuracy is 38.80% which is enhanced to 83.78%.

The average enhancement in accuracy for all 24 data set is 9.28%, 8.84%, and 22.07% for J48, CART and GATree; we get significant enhancement in accuracy with proposed approach.

In analysis of sample data size or reduced data set size, we found that data size in terms of number of training instances gets reduced and required sample size varies from 25% to 45% for J48, CART and GATree. Exceptional case is Breast-w on J48 with 70% sample size.

The average percentage sample data set size for all 24 clean data sets T_{RC} are 38.13%, 39.38% and 30.42% as compared to 72.70%, 74.17% and 33.33% for for unclean data set T_F on J48, CART and GATree respectively. We could observe very less deviation in sample size for GATree classifier from T_F to T_{RC}, the reason is already mentioned in above section.

The numbers of training instances required are less for clean data set as compared to unclean data set. As anomalies are removed the numbers of instances required for induction are also less. Thus with proposed approach we get reduced training data set for induction and upon induction enhanced classification performance is available.

Table 3. Effect of cleaning and sampling on data

Sr. No.	Data Set	J48				CART				GATree			
		X_F	X_{RC}	ΔX	SS	X_F	X_{RC}	ΔX	SS	X_F	X_{RC}	ΔX	SS
1	Australian	90.72	100.00	9.28	25	91.01	100.00	8.99	25	87.97	100.00	12.03	25
2	Breast –w	94.85	99.15	4.3	70	96.42	98.01	1.59	45	94.10	100.00	5.9	30
3	German	79.80	100.00	20.2	25	85.50	100.00	14.5	25	70.20	99.64	29.44	40
4	kr-vs-kp	99.41	100.00	0.59	35	99.62	100.00	0.38	40	91.05	98.67	7.62	25
5	Mfeat-Fact	93.50	99.66	6.16	40	93.25	98.93	5.68	25	38.80	83.78	44.98	25
6	Monks	70.16	93.33	23.17	25	64.52	93.33	28.81	50	70.00	100.00	30	25

5. EXPERIMENTS WITH LARGE DATA SET

The validation of proposed method on big data was done with test on two big benchmark data sets namely, Census-income (KDD) and KDDCup99 data set with data set size of 299,000 and 494,020 instances of data respectively. The classifiers used are J48, CART and GATree. The method of experimentation on big data has exception that the data sampling threshold is 5% instead of 25%. The threshold value for minimum sampling size is 5% for big data sets because the adequate data samples are available for training even at lower sampling percentage with big data set.

5.1. ANALYSIS ON ACCURACY OF THE TREE AND SIZE OF TRAINING DATA

Table 4 presents a comparison of percentage sample size, accuracies on clean and unclean data sets with and without sampling. It also presents the enhancement in accuracy ΔX. Here ΔX is the absolute difference in accuracy between the tree build from the cleaned data set and the tree build from complete data training data which is unclean data. The classification accuracies obtained on clean data T_C are higher, and are around 99% on clean data set, similar results are obtained clean sampled data T_{RC} and thus, classification performance of the proposed method in terms of accuracy is enhanced.

It its observed that the average percentage sample size is lower on clean data set as compared to unclean data. Due to filtering of data, outlier data is removed from data set and as anomalies in data are removed and the data set size required to build the model is also reduced. In case of unclean data the average sample set size around 24% where as for clean data set it around 13%. Thus, significant reduction in data set size is achieved using proposed technique.

Table 4. Comparison on percentage sample size and enhancement in accuracy

Data Set	Classifier	SS		X_F	X_{RU}	X_C	X_{RC}	ΔX
		Unclean	Clean					
Census-income	J48	30	10	95.39	95.26	100.00	99.99	4.51
	CART	30	10	95.50	95.24	100.00	99.99	4.49
	GATree	25	15	93.70	94.23	99.99	99.99	6.29
KDDCup99	J48	25	10	99.96	99.91	99.99	99.98	0.02
	CART	25	10	99.95	99.90	99.99	99.99	0.04
	GATree	10	25	88.56	93.67	98.06	98.30	9.74
Average		**24.17**	**13.33**	**95.51**	**96.37**	**99.67**	**99.71**	**4.18**

Table 5 provides comparison of number of instances in T_F and reduced data set T_{RU}, after cleaning of number of instances in T_C, and reduced data set T_{RC}.

In this table N_F indicates number of training instances in T_F, similarly N_C, N_{RU} and N_{RC} for T_C, T_{RU} and T_{RC}. The Table shows percentage reduction in required training set size ΔN with proposed method. It is the difference in number of training instances in unclean complete data set and cleaned, reduced data set. Where

$$\Delta N = (N_F - N_{RC}) * 100 / N_F \quad (1)$$

As anomalies are removed the required sample size is also reduced for clean data and the reduction is around 90%. This is the most significant achievement with proposed method because the size of data set is reduced considerably along with enhanced classification performance; the reduced data set help to deal with the memory and processor limitations. Thus reduction in training set size achieved is significant.

Table 5. Comparison data reductions and speed up in case unclean and clean large data sets.

Data Set	Classifier	Unclean Data				Clean Data				Δ t	Δ N
		N_F	$t(T_F)$	N_{RU}	$t(T_{RU})$	N_C	$t(T_C)$	N_{RC}	$t(T_{RC})$		
Census	J48	299K	137.84	89.7K	17.78	224.1K	19.88	22.4K	2.98	97.84	92.51
	CART	299K	8521.48	89.7K	1440.48	224.1K	1470.23	22.4K	79.24	99.07	92.51
	GATree	299K	1593.00	74.7K	352.00	224.1K	252.00	33.6K	32.00	97.99	88.76
KDDCup	J48	494K	305.70	123.5K	30.67	280.0K	48.17	28.0K	2.20	99.28	94.33
	CART	494K	1791.00	123.5K	242.22	280.0K	368.2	28.0K	23.84	98.67	94.33
	GATree	494K	2523.00	49.4K	550.00	280.0K	492.00	70.0K	60.00	97.62	85.83

5.2. ANALYSIS ON TIME REQUIRED TO BUILD THE MODEL

In this experimentation with large data sets one more parameter «time required to build the model» is added for analysis as it is significant in case of large data sets. In this analysis $t(T_F)$ indicate time required to build model on T_F, similarly $t(T_C)$, $t(T_{RU})$ and $t(T_{RC})$ indicate time required to build model on T_C, T_{RU} and T_{RC}.

The speedup or percentage reduction in average time required to build the model Δt with proposed method is the difference between average time required to build the model with all available data instances T_F and average time required to build the model on cleaned and reduced data set T_{RC}. The percentage of reduction in time required to build the model is given by

$$\Delta t = (t(t_F) - t(t_{RC})) * 100 / t(T_F) \qquad (2)$$

The results are presented in Table 5. Due to filtering of data, as anomalies in data are removed, time complexity of tree pruning process $O(n(\log n)^2)$ is reduced and thus decision tree learning process is accelerated. Here it is observed that the time required to build the model on clean data set is significantly lower than the time required building model on unclean data of same size. When we combine data cleaning and data sampling, time complexity of tree induction process $O(m\ n \log n)$ is also reduced due to reduction in n, significant reduction in time to build the model is observed for reduced clean data set T_{RC} as compared to complete data set T_F and there is around 98% reduction in time required to build the model for all cases in Table 5. Thus significant speedup is achieved with the proposed method.

6. CONCLUSION

The proposed approach is a combination of removal of noise from training data and incremental random sampling. The aim is to evolve decision trees in order to obtain compact representation of large data sets. The removal of outlier data improves the quality of training data. The results with this combined approach indicate significant improvements in classification performance along with data reduction.

The possible reduction in sample size depends on nature of data set and it cannot be fixed, but with significant improvement in data quality, significant reduction in required sample size with reference to complete data set is observed. The problems like memory and processor overloading can be handled due to reduced data set.

Since, noise is removed from training data, the time required to build the model is also reduced. Overfitting occurs due to noise in data. For n training instances with m attributes computational cost of building tree is $O(m\ n \log n)$, as n is reduced due to sampling and filtering the corresponding cost is reduced. Similarly the cost of decision tree pruning process is $O(n(\log n)^2)$. Due to data cleaning operation overfitting can be reduced and the time complexity of pruning process is reduced significantly. It is verified that the time required to build model on clean data is very low as compared to time required to build the model on unclean data of same size. The proposed approach improves data quality and reduces data set size while maintaining the classification performance and by virtue of improved data quality and reduced data size the proposed method gives excellent speedup. In conclusion with proposed method the size of data set is reduced, classification accuracy is improved, speedup is acquired and the problem of limitations of memory and processor can also be solved.

7. REFERENCES

[1] Tim Oates, David Jensen, The effect of training set size on decision tree complexity, *Proceedings 14th International Conference on Machine Learning*, (1997), pp. 254-262.

[2] G.H. John and Pat Langley, Static versus dynamic sampling for data mining, *In Proceedings of the Second International*

Conference on Knowledge Discovery in Databases and Data Mining, 1996.

[3] V. Barnett and T. Lewis, *Outliers in Statistical Data*, John Wiley and Sons, 1978.

[4] Gamberger, N. Lavrac, and S. Dzeroski, Noise elimination in inductive concept learning: a case study in medical diagnosis, *In proceedings of 7th International Workshop on Algorithmic Learning Theory*, Sydney, 1996.

[5] Ian H. Witten and Eibe Frank, *Data Mining Practical Machine Learning Tools and Techniques*, Morgan Kaufmann publications, 2005.

[6] Quinlan J. R., *C4.5: Programs for Machine Learning*, Morgan Kaufman, San Mateo, 1993.

[7] Quinlan J.R., Decision trees and decision making, *IEEE Transaction on Systems, Man, and Cybernetics,* (20) 2 (1990), pp. 339-346.

[8] Salvatore Ruggieri. Efficient C4.5: IEEE Transaction On Knowledge and Data Engineering, (14) 2 (2002), pp. 438-444.

[9] Quinlan J.R., Simplifying decision trees, International Journal of Man Machine Studies, 1987.

[10] Zhiwei Fu, Fannie Mae, A computational study of using genetic algorithms to develop intelligent decision trees, *Proceedings of the 2001 IEEE congress on evolutionary Computation*, 2001.

[11] Athanassios Papagelis, Dimitrios Kalles, GATree: genetically evolved decision trees, *Proceedings 12th International Conference on Tools with Artificial Intelligence*, (13-15 November 2000), pp. 203-206.

[12] A. Niimi and E. Tazaki, Genetic programming combined with association rule algorithm for decision tree construction, *Proceedings of Fourth International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies*, 2 (2000), pp. 746-749.

[13] Y. Kornienko and A. Borisov, Investigation of a hybrid algorithm for decision tree generation, *Proceedings of the Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing*, 2003.

[14] G. H. John, Robust decision trees: removing outliers from databases, *In Proceedings of the First ICKDDM*, (1995), pp. 174–179.

[15] C. E. Brodley and M. A. Friedl, Identifying mislabelled training data, *Journal of Artificial Intelligence Research*, (1999), pp. 131-167.

[16] Arning, R. Agrawal, and P. Raghavan, A linear method for deviation detection in large databases, *In KDDM*, 1996, pp. 164-169.

[17] V. Raman and J.M. Hellerstein, An interactive framework for data transformation and cleaning, *Technical report University of California Berkeley*, California, September 2000.

[18] A. I. Guyon, N. Matic and V. Vapnik, Discovering informative patterns and data cleaning, *Advances in knowledge discovery and data mining, AAAI*, (1996), pp. 181-203.

[19] D.Gamberger and N. Lavrac, Conditions for Occam's razor applicability and noise elimination, *Proceedings of the 9th European Conference on Machine Learning*, Springer, (1997), pp. 108-123.

[20] E. M. Knorr and R. T. Ng, Algorithms for mining distance-based outliers in large datasets, *In Proceedings 24th VLDB*, (1998), pp. 392-403.

[21] D. Tax and R. Duin, Outlier detection using classifier instability, *Proceedings of the workshop Statistical Pattern Recognition*, Sydney, 1998.

[22] D. D. Gamberger, N. Lavrac, and C. Groselj, Experiments with noise filtering in a medical domain, *In Proceedings 16th ICML*, Morgan Kaufman, San Francisco, CA, (1999), pp. 143-151.

[23] S. Schwarm and S. Wolfman, Cleaning data with Bayesian methods, *Final project report for University of Washington Computer Science and Engineering CSE574*, March 16, 2000.

[24] S. Ramaswamy, R. Rastogi, and K. Shim, Efficient algorithms for mining outliers from large data sets, *ACM SIGMOD*, (29) 2 (2000), pp. 427-438.

[25] J. Kubica and A. Moore, Probabilistic noise identification and data cleaning, *Third IEEE International Conference on Data Mining*, (19-22 Nov. 2003).

[26] V. Verbaeten and A. V. Assche, *Ensemble Methods for Noise Elimination in Classification Problems*, In Multiple Classifier Systems. Springer, 2003.

[27] J. A. Loureiro, L. Torgo, and C. Soares, Outlier detection using clustering methods: a data cleaning application, *In Proceedings of KDNet Symposium on Knowledge-based Systems*, Bonn, Germany, 2004.

[28] H. Xiong, G. Pande, M. Stein and Vipin Kumar, Enhancing data analysis with noise removal, *IEEE Transaction on knowledge and Data Engineering*, (18) 3 (2006), pp. 304-319.

[29] D. V. Patil and R. S. Bichkar, Improving classification performance of evolutionary decision tree using classification filter, *Third International Conference on Info. Processing*, Bangalore, India, (2009), pp.696-700.

[30] Slobodan Vucetic and Zoran Obradovic, Performance controlled data reduction for

knowledge discovery in distributed databases, *Proceedings 4th Pacific-Asia Conference, PADKK 2000*, Kyoto, Japan, (18-20 April 2000), pp. 29-39.

[31] A. Lazarevic and Z. Obradovic, Data reduction using multiple models integration: principles of data mining and knowledge discovery, *5th European Conference, PKDD 2001*, Freiburg, Germany, (3-5 September 2001), pp. 301-313.

[32] D. V. Patil and R. S. Bichkar, A hybrid evolutionary approach to construct optimal decision trees with large data sets, *In Proceedings IEEE ICIT06*, Mumbai, India (15-17 December 2006), pp. 429-433.

[33] Frank A. and Asuncion A., UCI Machine learning repository Irvine. CA, [http://archive.ics.uci.edu/ml]. University of California, School of Information and Computer Science, 2010.

[34] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, *The WEKA Data Mining Software: An Update; SIGKDD Explorations*, (11) 1 (2009).

Dipak V. Patil *received B.E. degree in Computer engineering in 1998 from University of North Maharashtra India and M. Tech. degree in Computer Engineering in 2004 from Dr. B. A. Technological University, Lonere, India. Currently he is an Associate Professor in Computer Engineering Department at Sandip Institute of Technology and Research Centre Nasik, India. His research interests are in data mining specifically handling noise in data and Handling large data sets.*

Rajankumar S. Bichkar *received B.E. (Electronics Engineering) and M.E. (Electronics Engineering) degrees from S.G.G.S. Institute of Engineering and Technology, Nanded India in 1986 and 1991 respectively, and Ph.D. degree in Engineering from prestigious Indian Institute of Technology, Kharagpur India in 2000.*

He has published his research papers in various International Journals. Currently he is a Professor (Electronics and Telecommunication Engineering) and Dean (R&D) at G.H. Raisoni College of Engineering & Management, Pune, India. His research interests are in genetic algorithms, image processing, scheduling, timetabling and data mining.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ОБОБЩЕННАЯ МОДЕЛЬ ОРГАНИЗАЦИИ ВЗАИМОДЕЙСТВИЙ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ

Антон Кабыш, Владимир Головко

Брестский государственный технический университет
ул. Московская, 267, 224017, Брест, Республика Беларусь
www.robotics.bstu.by
e-mail: anton.kabysh@gmail.com, gva@bstu.by

Резюме: *В данной работе описывается модель для нахождения оптимального поведения многоагентной структуры через организацию в ней оптимальных взаимодействий между агентами. Модель включает две основные техники. Модель графов координации позволяет явно выразить зависимость между агентами, что позволяет разбить целевую функцию поведения в линейную сумму индивидуальных целевых функций. Модель оценки влияний позволяет оценить влияния других агентов на действия друг друга и как результат позволяет им координировать свои действия. В работе приведена реализация данной модели на основе обучения с подкреплением и экспериментальные результаты применения данной модели.*

Ключевые слова: *многоагентные системы, обучение с подкреплением, Q-Learning, обучение через влияние, граф координации.*

GENERAL MODEL FOR ORGANIZING INTERACTIONS IN MULTI-AGENT SYSTEMS

Anton Kabysh, Vladimir Golovko

Brest State Technical University
Moskovskaya str. 267, 224017, Brest, Republic of Belarus
www.robotics.bstu.by
e-mail: anton.kabysh@gmail.com, gva@bstu.by

Abstract: *In this article we describe a model for finding optimal for learning the behavior of a group of agents in a collaborative multiagent setting. This model contains a set of scalable techniques that organize behavior of a multi-agent system. As a basis we use the framework of coordination graphs which exploits the dependencies between agents to decompose the global payoff function into a sum of local terms. To estimate a quality of interactions between agents we are using the concepts of the influence value learning paradigm. In last section we present the implementation of the considered model via reinforcement learning and experimental results of the use of this paradigm.*

Keywords: *multi-agent system, reinforcement learning, Q-Learning, coordination graph, influences learning.*

ВВЕДЕНИЕ

Огромный класс задач сводится к теории многоагентных систем, где рассматриваются способы взаимодействия между агентами. Типовые задачи, решаемые в теории многоагентных, систем включают в себя [1] автоматизированный трейдинг, ведение переговоров, управление нагрузкой сети, организация коллектива роботов-футболистов, распределенная оптимизация, распределенное планирование задач [2] и другие.

Традиционно, **многоагентная система** состоит из совокупности агентов, разработанных для кооперации друг с другом для достижения некоторой цели. **Агент** понимается как сущность, обладающая состоянием, способная воспринимать окружающую среду и выполнять в ней какие-то действия.

Свойства и поведение многоагентной системы

зависят от свойств входящих в неё агентов. В зависимости от алгоритма коллективного поведения многоагентная система выдвигает различные требования к агентам, которые могут быть использованы ею для данного алгоритма. Например, большинство алгоритмов стайного поведения не предполагает наличия коммуникации между агентами, а алгоритмы ведения переговоров требуют, что бы агенты были разными и по возможности, держали в секрете свои модели.

Представленная в статье модель так же предъявляет ряд требований к агентам. Агент рассматривается как **актор** [3] – активный агент, существующий параллельно (и одновременно) другим таким же агентам, взаимодействующий с другими посредством стандартного протокола внутреннего взаимодействия. Другими словами, агенты должны обладать способностью к коммуникации, а многоагентная система обеспечивать параллельность работы агентов.

Стоит отметить, что коммуникация подразумевает обмен сообщениями, в каком либо виде, а не восприятие другого агента посредством сенсоров. Агент должен иметь восприятие среды и восприятие сообщений, как нечто разделенное. Модель не специфицирует конкретный способ коммуникации между агентами. Это может быть как прямой обмен сообщениями, так и использование среды, в качестве посредника для передачи сообщений.

При обучении многоагентных систем либо изобретаются новые, гибридные алгоритмы, например в [4], либо адаптируются уже существующие. При этом разрабатывается какой-либо новый алгоритм, оперирующий уже существующими алгоритмами и регулирующий способ совместной работы этих алгоритмов. Существует 4 основные парадигмы адаптации алгоритмов обучения к многоагентным системам [5]:

- **Командное обучение** (*team learning*) развивает идею, что команду агентов можно обучать так, как если бы это был один агент. Этот подход не требует никаких изменений в алгоритмах обучения, но имеет ограниченную применимость. При увеличении количества агентов, экспоненциально растет и размер пространства состояний/действий, в котором ведется поиск решения (проблема проклятья размерности).
- **Индивидуальное обучение** (*individual learning*) рассматривает применение индивидуального алгоритма обучения каждому агенту, игнорируя дополнительные данные от других агентов. Проблемой данного подхода является то, что совокупность оптимальных индивидуальных стратегий не обязательно представляет оптимальную командную стратегию.
- **Совместное обучение** (*joint action learning*) фокусируется на обучении лучшим действиям в ответ на действия других агентов. Каждый агент обучается выполнять наилучшие действия в контексте объединенных действий других агентов. Этот подход позволяет построить оптимальную коллективную стратегию, но обладает тем же проклятьем размерности, что и командное обучение.
- **Обучение через влияние** (*influence-value learning*) основано на идее изменения поведения агента под влиянием мнения других агентов. Данный подход берет на вооружение множественные социальные факторы и использует их в качестве эвристик для конкретизации отношений внутри многоагентной системы.

В данной работе будет рассмотрена обобщенная, адаптивная модель организации взаимодействий внутри многоагентной системы на основе обучения через влияние для формирования в ней целенаправленного поведения посредством нахождения оптимальных взаимодействий между агентами. Под ожидаемым поведением многоагентной системы понимается совокупное поведение входящих в неё агентов, которое можно оценить как одно целое, приводящее к достижению некоторой цели. Примерами таких целей могут быть формирование и поддержание формации, коллективное управление нагрузкой или ресурсами, построение целостной карты окружающего пространства и др.

Модель обладает следующими свойствами:

- Данная модель является обобщенной, т.е. он не накладывает никаких первоначальных требований к агентам, кроме способности получать и отправлять сообщения другим агентам.
- Модель является адаптивной в том, что она позволяет изменить структуру многоагентной системы и поведение входящих в неё агентов, если изменилась целевая функция поведения.
- Модель фокусируется на определении тех оптимальных взаимодействий между агентами, которые приводят к решению поставленной задачи, сохраняя аспекты индивидуального обучения.

1. МОДЕЛЬ ОРГАНИЗАЦИИ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ

Ключевой концепцией является понятие **взаимодействия** или **влияния** между агентами. В реальном мире, при взаимодействии людей друг с другом, все действия отдельного индивида оцениваются, как и им самим, так и другими людьми в разрезе получаемого опыта. Примерами таких оценок являются ожидания похвалы или наказания до или после совершенного действия.

Все агенты, в коллаборативной многоагентной системе могут потенциально влиять друг на друга. Под **влиянием** (*influence*) понимается оценка другими агентами, направленная на текущего агента. Другие агенты могут оценивать поведение агента и его регулировать, с той целью, чтобы действия, выбранные индивидуальным агентом, соответствовали оптимальным решениям группы в целом. Любой агент, инициирующий влияние на другого агента называется *отправителем*. Агент, принимающий влияние и как-то на него реагирующий называется *получателем*.

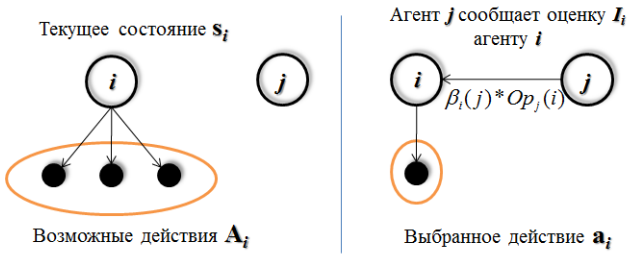
Каждый агент i выбирает индивидуальное действие, a_i из множества A_i находясь в некотором состоянии. Значение влияния между агентом i и группой из N агентов относительно выбранного агентом действия определяется следующей формулой [5]:

$$I_i = \sum_{j=1, i \neq j}^{N} \beta_i(j) * Op_j(i) \qquad (1)$$

где $\beta_i(j)$ – коэффициент влияния агента j на агента i $(0 \le \beta \le 1)$, $Op_j(i)$ – коэффициент оценки агентом j действия агента i.

Коэффициент влияния β определяет, будет или нет, будет ли агент подвержен влиянию других агентов. Так, если β будет стремиться к нулю, то агент предпочтет действовать индивидуально. *Op* – это оценка другими агентами действия выполняемого текущим агентом.

Рис.1. показывает пример формирования влияния агентом j после выбора агентом i действия.



(a) (b)

Рис. 1 – Пример модели влияний. (а) – Агент i выбирает действие. Выбрав действие агент j сообщил свою оценку i сформировав влияние

Многоагентная система формирует результирующее объединенное действие (joint action) $a = \{a_1, a_2, ..., a_N\}$. *Проблема координации* [6] поведения состоит в том, что бы найти такое оптимальное объединенное действие a^*, которое максимизирует полезность объединённого действия коллективного поведения системы $u(a)$, что $a^* = \arg\max_a u(a)$.

В большинстве реальных проблем, действия одного агента зависят лишь от небольшого числа других агентов. Например, в задаче коллективного сбора ресурсов, агенты могут координировать свои действия только с соседями. Для явного отслеживания таких зависимостей предложена модель **графов координации** (coordination graphs, (CG)) [7].

Пускай, действия агента i зависят только от некоторого подмножества других агентов, $j \in M(i)$. Тогда, целевая функция многоагентной системы $u(a)$ может быть разбита на линейную комбинацию локальных целевых функций по следующей формуле:

$$u(a) = \sum_{i=1}^{N} f_i(sub_i) \qquad (2)$$

где sub_i подмножество всего множества действий a, соответствующее действиям агентов $j \in M(i)$ от которых зависит агент i. Таким образом, глобальная проблема координации заменена на совокупность локальных проблем координации, каждая из которых включает малое подмножество агентов.

Зависимость между агентами может быть отражена в виде графа координации, где вершины соответствуют агентам, а ребра связывают зависимых агентов, которые обязаны координировать свои действия. По ребрам зависимостей агенты могут обмениваться

сообщениями, а сами ребра специфицировать конкретный протокол общения между агентами. Пример графа координации изображен на рис. 2.

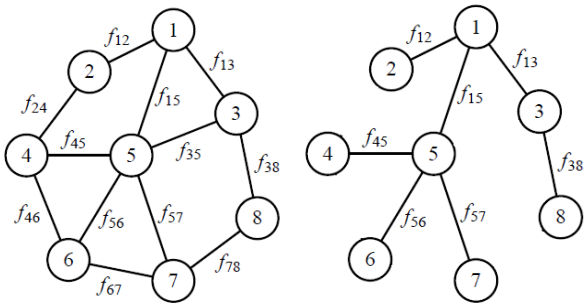


Рис. 2 – Пример графа координации для восьми агентов до (a) и после (b) декомпозиции зависимостей; каждое ребро представляет зависимость координации

В данном разделе не были специфицированы следующие основные пункты:

- Как происходит расчет коэффициента $Op_j(i)$ между агентами
- Как происходит перерасчет составляющих формулы (1) с течением времени.
- Как агент изменяет свое поведение под воздействием влияний от других агентов.
- Как агент оценивает влияние, оказываемое на него другими агентами.
- Каков алгоритм нахождения оптимальных зависимостей в многоагентной системе.

Эти пункты специально были оставлены открытыми для максимальной обобщенности алгоритма под конкретную задачу и алгоритм.

Например, в работе [5] коэффициент $Op_j(i)$ увеличивался, если агент j выполнял действие с меньшей ценностью, чем агент i, и наоборот, уменьшался, если агент j выполнял более выгодное действие. Это отражает факт того, что люди думают хорошо о тех действиях, которые приносят другим больше прибыли. В следующем разделе представлена реализация данной модели на основе подкрепляющего обучения.

Сформулируем задачу организации многоагентной системы следующим образом:

- Пусть, имеется N агентов, каждый их которых может выполнять некоторые действия (агенты могут быть как гетерогенными, так и гомогенными). Агенты имеют некоторое начальное состояние. Агенты объединены в многоагентную систему, обеспечивающую их параллельность и коммуникацию.
- Пусть имеется некоторая известная цель, достижение которой одним агентом либо невозможно, либо неэффективно по сравнению с многоагентным подходом. Агент, или многоагентная система может узнать, достигнута цель или нет.
- Требуется определить такое поведение агентов, которое приводит к достижению цели оптимальным образом. В частных случаях, данная задача уже может решаться различными способами [1], без коммуникации между агентами.
- С учетом коммуникации между агентами требуется определить оптимальные взаимодействия между ними, приводящие к коллективному решению задачи.
- Необходимо декомпозировать граф координации, при условии, что не имеется априорной информации о взаимодействиях между агентами, или она ограниченна.

Решение поставленной задачи выполняется посредством двух процессов, которые могут выполняться как последовательно, так и параллельно, в зависимости от используемых алгоритмов:

- Определение оптимальной структуры графа координации многоагентной системы с целью декомпозиции пространства решений.
- Определение оптимальных взаимодействий между агентами порождает оптимальное поведение.

2. ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ ДЛЯ НАХОЖДЕНИЯ ОПТИМАЛЬНОЙ СТРУКТУРЫ ВЗАИМОДЕЙСТВИЙ

Особенностью поставленной задачи является динамическое определение оптимальных влияний между агентами и графа координации. Для достижения этой цели требуется итеративный процесс настройки структуры многоагентной системы. За время функционирования многоагентной системы можно оценить разные взаимодействия между агентами, приводящими к решению задачи и усилить конструктивные и ослабить деструктивные. Если взаимодействие между агентами не является необходимым то, следовательно, координация между ними может не выполняться и зависимость на графе координации может быть устранена.

В качестве итеративного процесса оценки и формирования структуры многоагентной системы в большинстве работ [1] выступает **обучение с подкреплением** (*Reinforcement Learning*, (RL)) [9].

Ключевой особенностью данного метода является то, что он является активным методом

обучения, направленным на взаимодействие со средой и с другими агентами в случае многоагентной системы. В классической трактовке, обучение с подкреплением – это нахождение оптимального поведения агента методом проб и ошибок через его взаимодействие со средой. В данной работе, по отношению к многоагентной системе, обучение с подкреплением – это нахождение оптимальной структуры системы методом проб и ошибок через взаимодействия агентов.

Модель обучения с подкреплением формулируется следующим образом. Агент i выбирает действие $a_i(t)$, из множества A_i находясь в состоянии $s_i(t)$. После выполнения действия агент переходит в следующее состояние $s_i(t+1)$ и получает из внешней среды числовое значение награды $r_i(t+1)$ являющееся оценкой выполненного действия и совершенного изменения состояния. В состоянии $s_i(t+1)$ агент выбирает и выполняет следующее действие $a_i(t+1)$. Разница между ценностью следующего состояния с учетом награды и ценностью текущего состояния называется **ошибкой временной разности** (temporal-difference error).

Последовательность изменений состояния агента изображена на рис. 3 в виде **графа переходов** (*transition graph*). Вершинами графа переходов являются состояния агента, а ребрами отмечаются выбранные действия и переходы между состояниями вместе с ассоциированной с переходом значением награды. Глобальное, вся среда в которой действует агент описывается в виде Марковского процесса принятия решений.

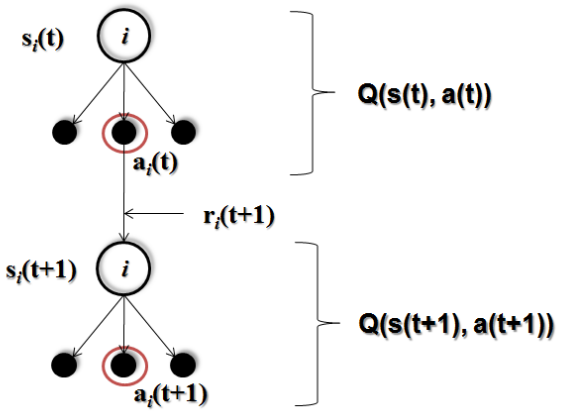


Рис. 3 – Граф переходов агента. Изображен переход агента из состояния s(t) в состояние s(t+1)

С любой парой состояние-действие (s,a) ассоциировано значение ценности $Q(s,a)$, обозначающее полезность выполнения действия a в состоянии s. Значения $Q(s,a)$ рассчитываются Q-функцией и заранее неизвестны. Цель обучения – аппроксимировать истинные значения оценок Q- функции путем последовательного посещения пар (s,a) и получения ассоциированных с ними награды r используя значение награды в качестве фактора изменяющего ценность состояния. Способ отображения состояния s на ассоциированное с ним действие a называется политикой $\pi(s,a)$. Следовательно, задача обучения с подкреплением сводится к нахождению оптимальной политики, которая на каждом шаге выбирает оптимальные действия, ведущие к максимальной суммарной награде в будущем.

Самым популярным алгоритмом обучения с подкреплением является Q-*learning* [9], обновляющий значения функции ценности на каждой итерации по формуле (3):

$$Q(s(t),a(t)) = Q(s(t),a(t)) + \alpha[r(t+1) + \gamma \max_{a\in A} Q(s(t+1),a(t+1)) - Q(s(t),a(t))] \quad (3)$$

где в квадратных скобках [.] рассчитывается ошибка временной разности, α – шаг обучения $(0<\alpha\le 1)$, а γ – коэффициент обесценивания $(0<\gamma\le 1)$ отдаленных ценностей.

Из алгоритма видно, что последующие значения ценностей влияют на предыдущие значения. Следовательно, в момент времени t(t+1) могут быть обновлены значения ценностей всех пар состояние действие которые, привели агента к текущему состоянию. Такая последовательность пар состояние действие называется **следами преемственности** (*eligibility traces*). Модифицированная версия алгоритма называется Watkins-Q(λ) [9].

Поскольку на начальном этапе обучения не известно, какие взаимодействия между агентами являются оптимальными, а какие нет, действия выбираются случайно, методом проб и ошибок (фаза исследования). В конце обучения истинные значения ценности аппроксимированы, и агент использует их как руководство для оптимального поведения (фаза использования) во внешней среде в контексте других агентов.

3. ПОИСК ОПТИМАЛЬНЫХ КОМАНД В МНОГОАГЕНТНОЙ СИСТЕМЕ

В разделе 1 было введено понятие влияния как контекста оценок создаваемого другими агентами по отношению целевому. Рассмотрим частный случай модели влияний, где влияние

выступает в виде частного случая – активной команды.

В отличие от влияния, **команда** – это (1) активное сообщение от отправителя к клиенту, способное изменить его состояние, принятое или проигнорированное, а так же (2) оценка этого сообщения. Если влияние принято, то оно может изменить текущее состояние агента или его последующие действия. Если влияние не принято, то оно возвращается с соответствующей пометкой и отправитель может отметить данное влияние как неконструктивное либо не влияющее на получателя.

Конструктивность команды – это характеристика, оценивающая ценность данного влияния по отношению к задаче решаемой многоагентной системой. Конструктивные влияния имеют положительную оценку полезности, с некоторыми ограничениями, в рамках которых эта полезность сохраняется.

Получатель может проигнорировать влияние либо ввиду того, что отправитель не имеет достаточно веса с точки зрения получателя, или получатель выполняет собственные «эгоистичные» действия, ведущие его к цели и нечувствителен к командам.

Рассмотрим пример двух агентов i и j. Оба агента в момент времени t находятся в некоторых состояниях и готовы выбрать действие для перехода в следующее состояние. Пускай, агент i может повлиять на агента j, указав какое ему действие ему выполнить. Посылаемую команду можно трактовать как выбор агентом i некоторого действия $a_i(t)$, которое, не переводит его в новое состояние $s_i(t+1)$. Агент j принимает команду в качестве указания и выполняет указанное действие, $a_j(t)$ переходя в новое состояние $s_j(t+1)$ и получая награду за выполнение действия $r_j(t+1)$.

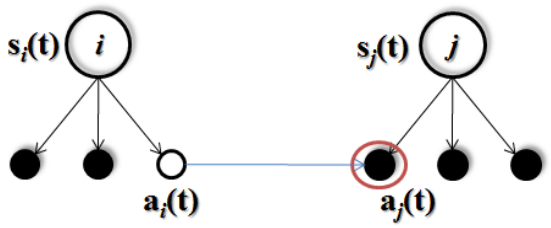


Рис. 4 – Диаграмма переходов для агентов *i* и *j*. Агент *i* инициирует команду над агентом *j*

После перехода, агент j может скорректировать ценность перехода $Q(s_j(t), a_j(t))$ по формуле (3), а так же выполнить оценку команды. В простейшем случае предполагаем, что коэффициент $\beta_i(j)$ равен 1. Значение $Op_j(i)$ содержит оценку агентом i действия агента j. Примером оценки являются качественные показатели следующего состояния агента j: награда $r_j(t+1)$ и обновленное значение ценности $Q'(s_j(t), a_j(t))$. Эти показатели передаются в качестве обратной связи агенту i, который может использовать их для обновления значения ценности команды.

Если обратная связь постоянно отрицательна, то ценность команды $Q(s_i(t), a_i(t))$ падает. Таким образом, выполняя разные команды над агентом j и получая обратную связь, агент i обучается оптимальному управлению над ним, формируя диапазон состояний, при котором сохраняется конструктивность команды.

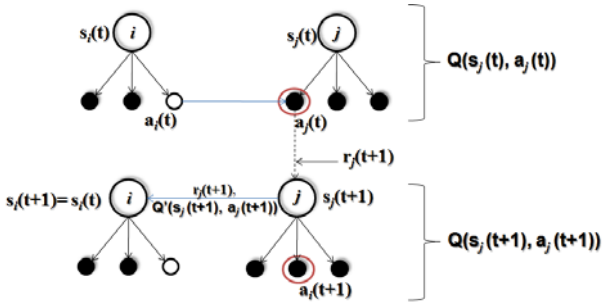


Рис. 5 – Диаграмма переходов для агентов *i* и *j*. Продолжение. Агент *j* выполнил команду, совершил переход, обновил значение ценности и вернул обратную связь агенту *i*

Модифицированная формула, для агента i обновляющая ценность команды по оценке, полученной от агента j имеет следующий вид:

$$\Delta Q(s_i(t), a_i(t)) = \alpha(Op_j(i) - Q(s_i(t), a_i(t))) \quad (4)$$

$$Op_j(i) = r_j(t+1) + \gamma Q_j^*(s_j(t+1), a_j(t+1)) \quad (5)$$

$$Q_j^*(s_j(t+1), a_j(t+1)) = \arg\max_{a_j \in A_j} Q(s_j(t+1), a_j(t+1)) \quad (6)$$

Если агенты не имеют конструктивных отношений друг с другом, то зависимость в поведении между этими агентами не является необходимой, что ведет к естественной декомпозиции графа координации.

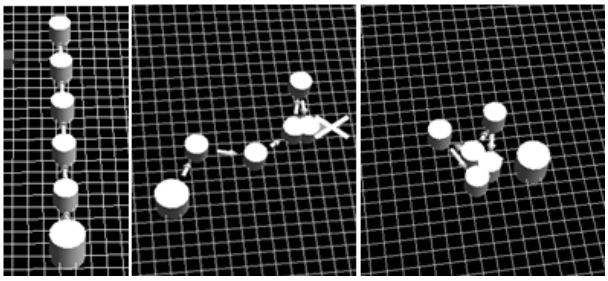
В этом примере агент i не мог выполнить одновременно команду и действие, т.к. выполнение команды трактовалось как отдельное действие. В зависимости от модели многоагентной системы, агенты могут вести переговоры и смешивать команды и действия в один такт времени.

4. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Для описания графа координации, графа переходов и графа Марковских процессов принятия решений был разработана библиотека моделирования на графах [10] предназначенная для задач робототехники и многоагентного моделирования.

Описанный в главе 1 подход применялся для задачи формирования и поддержания формации заданной формы на модельных агентах [11].

Описанная в главе 3 адаптация на основе команд применялась в задаче моделирования многозвенного робота обладающего 5-ю степенями свободы [12]. Среда моделирования и многоагентная система, имитирующая робота, изображена на рис. 6.



(a) (b) (c)

Рис. 6 – Моделирование многозвенного «робота-манипулятора». (a) Начальное положение. (b) Оптимальная последовательность команд найдена. (c) «Организационный хаос»

Каждый сегмент, кроме последнего мог командами изменять положение всех последующих сегментов. Итоговое расстояние до цели определялось после выполнения действий всеми сегментами. Целью эксперимента обучение робота достижению цели путем его самоорганизации посредством поиска оптимальных влияний между сегментами. Обученный робот оказался робастным к удалению или добавлению необученных сегментов, а так же быстро переучивался при смене цели. Поскольку граф координации был задан заранее, явная декомпозиция пространства поиска по сегментам (текущий сегмент зависел только от предыдущего) показала эффективность в скорости сходимости алгоритма по сравнению с командным и совместным обучением подверженным проклятью размерности. Рис. 7 иллюстрирует графики уменьшения суммарной (по всем сегментам) ошибки временной разности для трех алгоритмов обучения с подкреплением в сравнении с совместным обучением (JAL).

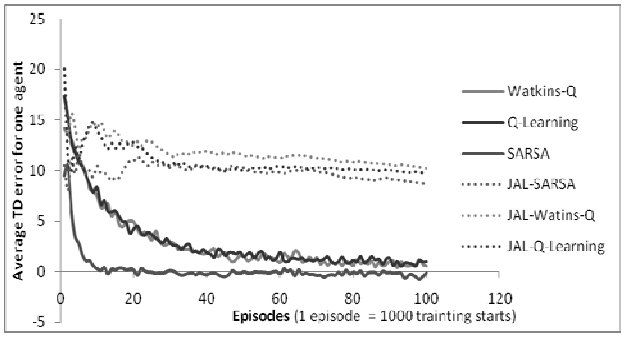


Рис. 7 – Графики изменения ошибки временной разности между совместным обучением и моделью команд

5. ВЫВОДЫ

Представленная в работе модель организации взаимодействий в многоагентной содержит набор техник, которые позволяют конкретизировать, оценить и оптимизировать отношения между агентами в системе с целью координации их действий. Представленный подход к адаптации рассмотренных концепций не единственный. В работах [5,6] представлены адаптации данного подхода для задач теории игр, коллективного фуражирования, коллективного принятия решений и ряда других.

Эксперимент с использованием модели показали, что эвристики, конкретизирующие отношения между агентами, применимы для широкого класса задач и могут быть более эффективны, чем изобретение нового алгоритма.

6. СПИСОК ЛИТЕРАТУРЫ

[1] L. Panait and S. Luke, Cooperative multi-agent learning: the state of the art, *Autonomous Agents and Multi-Agent Systems*, (11) 3 (2005), pp. 387-434.

[2] T. Gabel, *Multi-Agent Reinforcement Learning Approaches for Distributed Job-Shop Scheduling Problems*. PhD Thesis, University of Osnabrueck, 2009, 175 p.

[3] E. Lee, S. Neuendorffer, M.J. Wirthlin, Actor-oriented design of embedded hardware and software systems, *Journal of Circuits, Systems and Computers*, 12 (2003), pp 231-260.

[4] N. Monekosso, P. Remagnino, The analysis and performance evaluation of the pheromone-Q-learning algorithm, *Expert Systems,* (21) 2 (2004), pp. 80-91.

[5] B. Dennis, Luiz M. G. Goncalves, *Influence Value Q-Learning: A Reinforcement Learning Algorithm for Multi Agent Systems*, in: Meng Joo Er and Yi Zhou (Eds.), Theory and Novel Applications of Machine Learning, Book, I-Tech, Vienna, Austria, 2009, pp. 376.

[6] J. R. Kok, N. Vlassis, Sparse cooperative q-learning, *In Proceedings of the XXI international conference on Machine Learning*. Banff, Alberta, Canada, (2004), pp. 61.

[7] C. Guestrin, D. Koller, R. Parr, Multiagent planning with factored MDPs, *In Advances in Neural Information Processing Systems* (NIPS) 14. The MIT Press, (2002), pp. 1523-1530.

[8] J Kok, N Vlassis, Using the max-plus algorithm for multiagent decision making in coordination graphs, *RoboCup 2005: Robot Soccer World Cup IX*, 2006.

[9] R. S. Sutton, A. G. Barto, *Reinforcement Learning: an Introduction*, MIT Press, 1998.

[10] A. Kabysh, Graph Modeling Framework, BrSTU Robotics Wiki (www.robotics.bstu.by/mwiki), (2012), http://robotics.bstu.by/mwiki/index.php?title=Библиотека_моделирования_на_графах (in Russian).

[11] V. A. Golovko, A. S. Kabysh, Collective behavior in multiagent systems based on reinforcement learning, *Proceedings of the Tenth International Conference «Pattern recognition and image processing»* (PRIP–2009), Minsk, Belarus, (19-21 May 2009), pp. 260-264.

[12] A. Kabysh, V. Golovko, A. Lipnickas, Influence learning for multi-agent system based on reinforcement learning, *International Journal of Computing*, (11) 1 (2012) pp. 39-44.

Антон Сергеевич Кабыш, *Брестский государственный технический университет, Брест, Беларусь. Аспирант, кафедра ИИТ. Научные интересы: многоагентные системы, машинное обучение, data mining, робототехника, collective intelligence.*

Проф. Владимир Адамович Головко, *получил степень магистра по компьютерной инженерии в 1984 году в Московском государственном техническом университете им. Баумана. В 1990 он защитил кандидатскую диссертацию в Белорусском государственном университете и в 2003 году он защитил докторскую диссертацию по компьютерным наукам в Объединенном институте проблем информатики Национальной Академии Наук Беларуси. В настоящее время он работает заведующим кафедрой интеллектуальных информационных технологий и лаборатории искусственных нейронных сетей Брестского государственного технического университета.*

Научные интересы включают искусственный интеллект, нейронные сети, обучение автономных роботов, обработка сигналов, хаотические процессы, выявление вторжений и эпилепсии.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

GENERAL MODEL FOR ORGANIZING INTERACTIONS IN MULTI-AGENT SYSTEMS

Anton Kabysh, Vladimir Golovko

Brest State Technical University
Moskovskaya str. 267, 224017, Brest, Republic of Belarus
www.robotics.bstu.by
e-mail: anton.kabysh@gmail.com, gva@bstu.by

Abstract: *In this article we describe a model for finding optimal for learning the behavior of a group of agents in a collaborative multiagent setting. This model contains a set of scalable techniques that organize behavior of a multi-agent system. As a basis we use the framework of coordination graphs which exploits the dependencies between agents to decompose the global payoff function into a sum of local terms. To estimate a quality of interactions between agents we are using the concepts of the influence value learning paradigm. In last section we present the implementation of the considered model via reinforcement learning and experimental results of the use of this paradigm.*

Keywords: *multi-agent system, reinforcement learning, Q-Learning, coordination graph, influences learning.*

A multiagent system (MAS) consists of a group of agents that reside in an environment and can potentially interact with each other. In this article we are interested in collaborative multiagent systems in which the agents have to work together in order to optimize a shared performance measure. In detail, we focus on inherently cooperative tasks involving a large group of agents in which the success of the team is measured by the specific combination of actions of the agents.

This kind of problems requires special techniques for producing effective behavior in multi-agent system. This article describes two kinds of such techniques: influence value as a measure about executing actions from another agent, and coordination graphs that's decomposes a coordination problem into a combination of simpler problems.

In influence value paradigm, agents estimate the values of their actions based on the reward obtained and a numerical value called influence value. The influence model can specify two side relationships between agents that behavior requires coordination. Influence value determines the measure of quality of agent's interactions. E.g. it can determine whether or not an agent will be influenced by opinions of other agents.

All agents in a collaborative multiagent system can potentially influence each other. It is therefore important to ensure that the actions selected by the individual agents result in optimal decisions for the group as a whole. This is often referred to as the coordination problem.

Fortunately, in many problems the action of one agent does not depend on the actions of all other agents, but only on a small subset. For example, in many real-world domains only agents which are spatially close have to coordinate their actions. The framework of coordination graphs is a recent approach to exploit these dependencies. This framework assumes the action of an agent i only depends on a subset of the other agents, $j \in \Gamma(i)$. The global payoff function is then decomposed into linear combination of local payoff function.

The described techniques in combination with reinforcement learning helps to solve almost all the difficulties facing when need to find optimal multi-agent structure producing optimal interactions.

REFERENCES

[1] L. Panait and S. Luke, Cooperative multi-agent learning: the state of the art, *Autonomous Agents and Multi-Agent Systems*, (11) 3 (2005), pp. 387-434.

[2] T. Gabel, *Multi-Agent Reinforcement Learning Approaches for Distributed Job-Shop Scheduling Problems*. PhD Thesis, University of Osnabrueck, 2009, 175 p.

[3] E. Lee, S. Neuendorffer, M.J. Wirthlin, Actor-oriented design of embedded hardware and software systems, *Journal of Circuits, Systems*

and Computers, 12 (2003), pp 231-260.

[4] N. Monekosso, P. Remagnino, The analysis and performance evaluation of the pheromone-Q-learning algorithm, *Expert Systems,* (21) 2 (2004), pp. 80-91.

[5] B. Dennis, Luiz M. G. Goncalves, *Influence Value Q-Learning: A Reinforcement Learning Algorithm for Multi Agent Systems*, in: Meng Joo Er and Yi Zhou (Eds.), Theory and Novel Applications of Machine Learning, Book, I-Tech, Vienna, Austria, 2009, pp. 376.

[6] J. R. Kok, N. Vlassis, Sparse cooperative q-learning, *In Proceedings of the XXI international conference on Machine Learning*. Banff, Alberta, Canada, (2004), pp. 61.

[7] C. Guestrin, D. Koller, R. Parr, Multiagent planning with factored MDPs, *In Advances in Neural Information Processing Systems* (NIPS) 14. The MIT Press, (2002), pp. 1523-1530.

[8] J Kok, N Vlassis, Using the max-plus algorithm for multiagent decision making in coordination graphs, *RoboCup 2005: Robot Soccer World Cup IX*, 2006.

[9] R. S. Sutton, A. G. Barto, *Reinforcement Learning: an Introduction*, MIT Press, 1998.

[10] A. Kabysh, Graph Modeling Framework, BrSTU Robotics Wiki (www.robotics.bstu.by/mwiki), (2012), http://robotics.bstu.by/mwiki/index.php?title=Библиотека_моделирования_на_графах (in Russian).

[11] V. A. Golovko, A. S. Kabysh, Collective behavior in multiagent systems based on reinforcement learning, *Proceedings of the Tenth International Conference «Pattern recognition and image processing»* (PRIP–2009), Minsk, Belarus, (19-21 May 2009), pp. 260-264.

[12] A. Kabysh, V. Golovko, A. Lipnickas, Influence learning for multi-agent system based on reinforcement learning, *International Journal of Computing*, (11) 1 (2012) pp. 39-44.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ФОРМУВАННЯ ТА ОПРАЦЮВАННЯ ЕНТРОПІЙНО-МАНІПУЛЬОВАНИХ СИГНАЛЬНИХ КОРЕКТУЮЧИХ КОДІВ У БЕЗПРОВІДНИХ СЕНСОРНИХ МЕРЕЖАХ

Артур Воронич [1)], Ярослав Николайчук [2)], Володимир Гладюк [2)]

[1)] Івано-Франківський національний технічний університет нафти і газу,
76019, м. Івано-Франківськ, вул. Карпатська 15, archy.bear@gmail.com
[2)] Тернопільський національний економічний університет, 46020, Тернопіль, вул. Львівська 11

Резюме: *Викладена систематизація теоретичних основ та аналітичних виразів оцінки мір ентропії. Наведено характеристику сигнальних коректуючих кодів поля Галуа. Описано можливість використання ентропійної маніпуляції та сигнальних коректуючих кодів у безпровідних сенсорних мережах. Встановлено, що енторопійний підхід з застосування коректуючих кодів поля Галуа є найбільш ефективним та перспективним для формування і опрацювання сигналів при передачі інформації в безпровідних сенсорних мережах.*

Ключові слова: *безпровідна сенсорна мережа, ентропія, сигнальні коректуючі коди, Галуа.*

FORMATION AND PROCESSING OF ENTROPY-MANIPULATED CORRECTING SIGNAL CODES IN WIRELESS SENSOR NETWORKS

Artur Voronych [1)], Yaroslav Nykolaychuk [2)], Volodymyr Hladyuk [2)]

[1)] Ivano-Frankivsk National Technical University of Oil and Gas
15 Carpathian Str., 76019 Ivano-Frankivsk, archy.bear @ gmail.com
[2)] Ternopil National Economic University, 11 Lvivska Str., 46020 Ternopil

Abstract: *It is described a systematization of theoretical bases and analytical expressions for evaluation of entropy measures. It is showed a characteristic of signal corrective codes of Galois field. It is described the use of entropic manipulation and corrective signal codes in wireless sensor networks. It is founded that entropy approach of applying the corrective codes of Galois field is the most effective and promising for the formation and processing of signals during transmission in wireless sensor networks.*

Keywords: *wireless sensor networks, entropy, corrective signal codes, Galois.*

ВСТУП

Безпровідні сенсорні мережі (wireless sensor networks) – одні з передових комп'ютерних мережевих технологій, розмір ринку яких в 2011р. становив пів мільярда доларів[1]. Основою таких мереж є мініатюрні обчислювально-комунікаційні пристрої – моти (від англ. motes – порошинки), або сенсори. Сенсор є платою розміром не більш за один кубічний дюйм. На платі розміщуються процесор, пам'ять флеш і оперативна, цифро-аналогові і аналого-цифрові перетворювачі, радіочастотний приймач, джерело живлення і сенсори. Сенсори можуть бути найрізноманітнішими, вони підключаються через цифрові і аналогові конектори. Частіше за інших використовуються сенсори температури, тиску, вологості, освітленості, вібрації, магніто-електричні, хімічні, звукові і деякі інші. Набір використаних сенсорів залежить від функцій, що виконуються безпровідними сенсорними мережами. Живлення мотів здійснюється від невеликої батареї [2]. Моти використовуються тільки для збору, первинної обробки і передавання сенсорних даних.

1. ТЕОРЕТИЧНІ ОСНОВИ ЕНТРОПІЙНИХ МЕТОДІВ ОПРАЦЮВАННЯ СИГНАЛІВ

В світовій практиці для оцінки ентропії інформації найширше застосування отримали інформаційні міри Р. Хартлі та К. Шеннона. В той же час на практиці застосовується багато інших оцінок ентропії(табл. 1), які можуть бути теоретичною основою для методів формування та опрацювання сигналів [3,4,5].

Таблиця 1. Формули оцінки ентропії

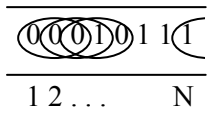
№	Функція	Міра ентропії
1.	$H = \log S^n = n \cdot \log S$, де H – ентропія; S – число незалежних рівноймовірних станів джерел інформації(ДІ); n – довжина вибірки.	Р. Хартлі
2.	$I_x = \log_2 \sqrt{2\pi e \sigma_x^2}$, $I_x = \hat{E}[\log_2 3\sigma_x]$ $\hat{E}[\ ]$ – цілочисельна функція з округленням до більшого	К. Крампа
3.	$H_\varepsilon \le T/\Delta t + \log(C/\varepsilon)$,де Δt – крок дискретизації, що забезпечує точність квантування ε; С – діапазон квантування; T – інтервал часу спостереження ДІ.	А. Колмогор
4.	$H = -k\sum_{i=1}^{n} p_i \cdot \log p_i$, де k – додатній коефіцієнт, що враховує основу логарифма; p_i – ймовірність *i*-го стану дискретного ДІ.	К. Шенно-на
5.	$h_\Delta = \lvert f'_{cep}(t)\rvert / \lvert f'_{\max}(t)\rvert$, де $f'_{cep}(t)$, $f'_{\max}(t)$ – відповідно середнє і максимальне значення похідних зміни кількості станів джерела.	В. Боюна
6.	$I_x = n \cdot \hat{\mathrm{E}}\left[\frac{1}{2}\log_2 \frac{1}{m}\sum_{j=1}^{m}\left(D_x^2 - R_{xx}^2(j)\right)\right]$, де D_x – дисперсія значень x_i; $R_{xx}(j)$ – автокореляційна функція; m – число точок функції $R_{xx}(j)$ на інтервалі кореляції.	Я. Николай-чука
7.	$H(u,p) = -k\sum_{i=1}^{n}[u_i p_i \cdot \log p_i]$, де u_i – коефіцієнт корисності; k – стала величина; $p = p_i$ – імовірність *i*-го стану.	Дж. Лонго
8.	$H(P,W) = -\sum_{i=1}^{n} \frac{P_i W_i}{\sum_{j=1}^{n} P_j W_j} \cdot \log \frac{P_i W_i}{\sum_{j=1}^{n} P_j W_j}$, де $P_i W_j$ – оціночні коефіцієнти.	Г. Шульца
9.	$H = \lim[(\log N)/n] = -\sum p(j) \cdot \log p(j)$, де $N = n! \Big/ \prod_j^n S_j$ - загальна кількість можливих комбінацій станів	Б. Олівера
10.	$H(X) = -\sum_{l_1}^{L} \ldots \sum_{l_n}^{L} p(X) \log p(X)$, де $(l \le l_i \le L;\ i = 1,\ 2,\ \ldots,\ n)$. X – апріорна невизначеність ; X_i, y_i – статистично залежні стани ДІ	Д. Мідл-тона
11.	$H \le k2BT(1 + S/N)$ $H = k \cdot n \log S_{ave}$, де S_{ave} – середнє значення станів ДІ; BT – інформаційна база повідомлень, що формується; N – значення шуму.$1/S$ – інтервал кореляції між відліками	В. Таллера

Детальний аналіз властивостей оцінок ентропії проведений в [5], де показано, що реалізація ентропійних моделей на основі функцій автокореляції(табл. 1(6) дозволяє зняти існуючі функціональні обмеження, які присутні при розрахунку ентропії на основі інформаційних мір Р. Хартлі та К. Шеннона. Оскільки перша не враховує статистичного розподілу потоку інформаційних даних, а друга враховує тільки ймовірності станів джерел інформації і не враховує ймовірностей їх переходу з одного стану в інший, що реалізується в інформаційній мірі ентропії на основі формули, запропонованої проф. Николайчуком Я.М [3,4,5].

Прикладом кодів, які володіють високими ентропійними характеристиками є сигнальні коректуючі коди в базисі Галуа [3].

2. СИГНАЛЬНІ КОРЕКТУЮЧІ КОДИ В БАЗИСІ ГАЛУА

Коди поля Галуа (рис. 1)[3, 6] за загальною класифікацією відносяться до підкласу циклічних блокових кодів, які володіють всіма основними властивостями завадозахищених кодів. В блокових кодах послідовність елементарних повідомлень розбиваються на блоки символів (B_1, B_2, B_3, …, B_n) фіксованої довжини К, кожному з яких ставиться в відповідності певна комбінація символів кодового слова (b_1, b_2, b_3, …, b_n).



$V=N$

V –об'єм виродженої матриці

Рис. 1 – Представлення коду Галуа

Циклічні коди відносяться до класу систематичних кодів. Для даних кодів можна записати відповідний їм аналітичний вираз, чи деяке логічне співвідношення, яке визначається правилами створення цих кодів. Найбільш зручною формою представлення циклічних кодів – використання алгебраїчного виразу [3, 6]:

$$G(x) = a_{n-1} \cdot x^{n-1} + a_{n-2} \cdot x^{n-2} + ... + a_1 \cdot x + a_0 \,(1),$$

де $a_0,\ldots,a_{n-1}$ – числа, що дорівнюють «0» чи «1», які визначають відповідні значення розрядів кодових комбінацій.

Таким чином дія над циклічними кодами зводиться до дії над відповідними математичними виразами. Дані коди є одними з найбільш досконалою упаковкою інформації.

Найбільш ефективно переваги даного базису можна використати при кодуванні інтегральних значень, оскільки при інтегруванні кожне наступне значення збільшується на одиницю. Тому на відміну від базису Радемахера кожне дискретне значення інтеграла функції замість n–розрядного двійкового коду фіксується одним бітом Галуа (рис. 2).

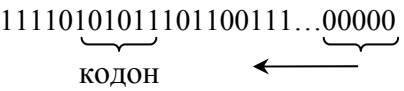


Рис. 2 – Формування коду Галуа

У загальному випадку код поля Галуа генерується згідно рівнянь незвідних поліномів (табл. 2) на основі узагальненого рекурентного виразу[7]:

$$X_{i+1} = \left(x_i \cdot a_i \oplus x_{i-1} \cdot a_{i-1} \oplus ... \oplus x_{i-n} \cdot a_{i-n}\right), \; (2)$$

де $a_i \in 0,1$ - двійкові значення незвідного алгебраїчного поліному, що формує код рекурентного ключа для послідовності.

В окремих випадках код Галуа G_2^n формується згідно рекурентної послідовності:

$$G_{i+1} = G_i \oplus G_{i-n}, \qquad (3)$$

де n – розрядність кодона

Таблиця 2. Незвідні поліноми $\pi(x) = x^r + f(x)$ степенів r і характеристик p

p	r	$\pi(x)\,;\ \left\lfloor x^r = f(x) \right\rfloor$
2	2	$x^2 + x + 1$
	3	$x^3 + x + 1\,; x^3 + x^2 + 1$
	4	$x^4 + x + 1$
	5	$x^5 + x^2 + 1$
	6	$x^6 + x + 1$
	7	$x^7 + x + 1\,; x^7 + x^3 + 1$
	8	$x^8 + x^4 + x^3 + x^2 + 1$
	9	$x^9 + x^4 + 1$
	10	$x^{10} + x^3 + 1$
	11	$x^{11} + x^2 + 1$
	12	$x^{12} + x^6 + x^4 + x + 1$
	13	$x^{13} + x^4 + x^3 + x + 1$
	14	$x^{14} + x^{10} + x^6 + x + 1$
	15	$x^{15} + x + 1$
	16	$x^{16} + x^{12} + x^3 + x + 1$
	17	$x^{17} + x^3 + 1$
	18	$x^{18} + x^7 + 1$
	19	$x^{19} + x^5 + x^2 + x + 1$
	20	$x^{20} + x^3 + 1$
	21	$x^{21} + x^2 + 1$
	22	$x^{22} + x + 1$
	23	$x^{23} + x^5 + 1$
	24	$x^{24} + x^7 + x^2 + x + 1$
	25	$x^{25} + x^3 + 1$
	26	$x^{26} + x^6 + x^2 + x + 1$

При передаванні та прийманні інформації на основі сигнальних кодів використовуються маніпульовані сигнали сформовані на основі чотирьох ознак, які поставлені у відповідність до елементів інформаційного повідомлення відповідно до кодів поля Галуа [6].

Відомий спосіб передавання, та приймання інформації на основі модифікованої частотної модуляції (MFM) [8]. В модифікованій частотній модуляції використовуються 4 сигнальні ознаки: фронт наростання($\wedge$), фронт спаду($\vee$), які відповідають символу «1», і потенціал «+», потенціал «–», які відповідають символу «0». При повторенні символу «0» для бітової синхронізації також використовують фронт наростання($\wedge$) чи спаду($\vee$).

Проте такий спосіб не дозволяє виявляти і виправляти помилки.

Інший спосіб передавання та приймання інформації[9] дозволяє виявити та виправити помилки при прийманні біторієнтованих даних за рахунок коректуючих властивостей кодів Галуа без додаткового передавання контрольних сум.

При передаванні та прийманні інформації використовуються маніпульовані сигнали сформовані на основі чотирьох ознак, які поставлені у відповідність до елементів інформаційного повідомлення відповідно до кодів поля Галуа. Такий принцип формування сигнального коду полягає в тому, що біти одиниць в пакеті даних нумеруються рекурентним кодом Галуа G^2_k. Причому для одиниць в пакеті даних біт Галуа «1» передається фронтом наростання($\wedge$), а біт Галуа «0» передається фронтом спаду($\vee$). Біти нулів в пакеті даних також нумеруються рекурентним кодом Галуа G^2_k,. Причому для нулів в пакеті даних біт Галуа «1» передається потенціалом «+», а біт Галуа «0» передається потенціалом «–». В якості чотирьох ознак маніпуляції на сигнальному рівні може бути відповідно використані набори з чотирьох фаз, частот, шумоподібних сигналів та їх комбінацій.

Приклад сигналу маніпульованого за допомогою сигнальних коретуючих кодів, при якому об'єм коду Галуа відповідає об'єму даних наведено на рис. 3. З таблиці (рис. 3) видно, що в блоці об'ємом $N=2^4$ завершення послідовності нулів відповідає коду Галуа 1101 і завершується символами $+\,+\,-\,+\,\vee\,\vee\,\wedge\,\vee$, тобто N=6, згідно $G_2{}^4$. А завершення послідовності одиниць в коді Галуа відповідає символам $\wedge\,\vee\,\vee\,\wedge$, тобто коду Галуа 0110, N=10.

Позиції бітів	d1	d2	d3	d4	d5	d6	d7	d8	d9	d10	d11	d12	d13	d14	d15	d16
Біти даних	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Біти Галуа $G_4^2(1)$	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0	0
Символьний код	$\wedge$	$\wedge$	$\wedge$	$\wedge$	$\vee$	$\wedge$	$\vee$	$\wedge$	$\wedge$	$\vee$	$\vee$	$\wedge$	$\vee$	$\vee$	$\vee$	$\vee$
Сигнальний код																
Біти даних	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Біти Галуа $G_4^2(0)$	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0	0
Символьний код	+	+	+	+	−	+	−	+	+	−	−	+	−	−	−	−
Сигнальний код																
Біти даних	1	1	0	1	1	0	1	1	1	0	0	0	1	1	0	1
Біти Галуа $G_4^2(1)$	1	1		1	1		0	1	0				1	1		0
Біти Галуа $G_4^2(0)$			1			1				1	1	0			1	
Символьний код	$\wedge$	$\wedge$	+	$\wedge$	$\wedge$	+	$\vee$	$\wedge$	$\vee$	+	+	−	$\wedge$	$\wedge$	+	$\vee$
Сигнальний код																

Рис. 3 – Приклад формування сигнального коректуючого коду для інформаційного повідомлення

Таким чином забезпечується ефективне симетричне кодування у вигляді кодів Галуа послідовності нулів і одиниць блоку даних з однозначним визначенням їх числа $N_0+N_1=N$, яке використовується для виявлення та виправлення помилок після передавання даних.

На рис. 4 зображена реалізація потоку даних маніпульованих за допомогою сигнальних коректуючи кодів, з виявленням помилок на сигнальному рівні. В таблиці приведено приклад виникнення помилок на сигнальному рівні в 7-ій та 17-ій позиції нулів, а також 10-ій та 21-ій позиції одиниць.

Виявлення помилок ґрунтується на біт-орієнтованій нумерації послідовності нулів і одиниць, які передаються за допомогою кодових послідовностей Галуа. Якщо помилка виявлена, використовується формула, де рекурентним шляхом перевіряється, в якій саме позиції відбулася заміна символу нуля(одиниці), в процесі передавання даних і даний символ замінюється на правильний [9].

Номер позиції бітів	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	…
Біти даних	1	0	0	1	0	1	0	1	1	1	0	1	1	0	1	1	0	1	1	1	1	0	1	0	…
Біти Галуа $G_4^2(1)$	1			1		1		1	0	1		0	1		1	0		1	0	0	0		0		…
Біти Галуа $G_4^2(0)$		1	1		1		1				0			1			0					1		1	…
Сигнальний код																									
Помилка *							*			*							*				*				…
Символьний код з помилкою	$\wedge$	$\vee$	$\vee$	$\wedge$	$\vee$	$\wedge$	$\frac{\wedge}{-}$	$\wedge$	−	$\frac{+}{\vee}$	+	−	$\wedge$	$\vee$	$\wedge$	−	$\frac{\wedge}{\vee}$	−	−	−	$\frac{\wedge}{+}$	$\vee$	$\wedge$	$\vee$	…

Рис. 4 – Реалізація потоку даних за допомогою сигнальних коректуючих кодів, з виявленням помилок на сигнальному рівні

3. СПІРАЛЬНІ ВЛАСТИВОСТІ КОРЕКТУЮЧИХ КОДІВ В БАЗИСІ ГАЛУА

Теоретична основа підвищення завадозахищеності сигнальних коректуючих кодів показана в роботі [10], на основі виявленої наявності в послідовності Галуа черезрівневої рекурентної властивості.

Нехай маємо код Галуа, який представляється рекурентним кодом:

$$G_2^4 = 11110101\ \ 10010001111010110\ldots$$

Можна показати, що код генерований на основі виразу (3) можна упакувати в спіраль (рис. 5), причому по кожній з чотирьох твірних формується рекурентна послідовність, яка має відповідні рекурентні властивості коду в базисі Галуа.

Отримана спіраль закодована рекурентним кодом розкручується зі збереженням рекурентності через 12 символів, що дозволяє виправляти помилки по твірних спіралі згідно виразу:

$$G_{i+1} = G_i \oplus G_{i-13} \tag{4}$$

З врахуванням спіральних властивостей сигнальний рекурентний код можна ефективно використати при виявлені пакетів помилок. Оскільки спочатку перевіряється і виправляється помилки згідно виразу (1) а потім, згідно виразу (4) по твірних спіралі.

На рис. 5 показано виявлення помилок при використанні спіральних властивостей сигнальних коректуючих кодів поля Галуа. Як видно з рис. 6 при виникненні помилок в п’яти підряд позиціях виявлення помилок завдяки рекурсивному виразу (3) стає неможливим, оскільки можлива неправильна корекція, тому потрібно виявляти дані помилки завдяки спіральним властивостям, тобто виразу (4).

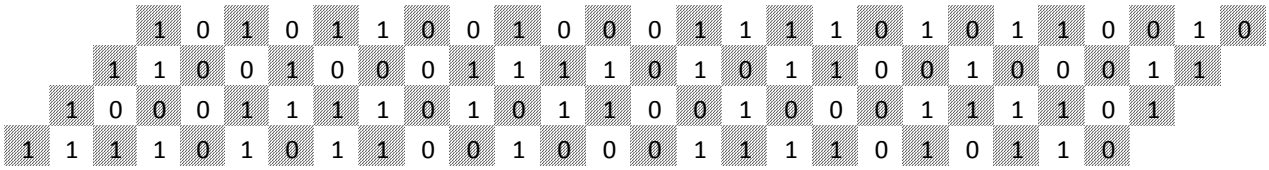


Рис. 5 – Сигнальний коректуючий код упакований у вигляді спіралі

Номер позиції бітів	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	…
G_4^2	1	1	1	1	0	1	0	1	1	0	0	1		0	0	1	1	1	1	0	1	0	…
G_{i-4}										*	*	*	*	*									
G_{i-13}													*										

Рис. 6 – Виявлення помилок при використанні спіральних властивостей сигнальних коректуючих кодів поля Галуа

4. БЕЗПРОВІДНА СЕНСОРНА МЕРЕЖА

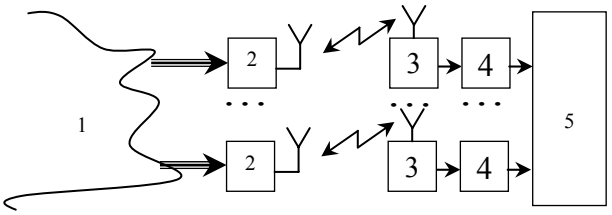
У мережі компанії Dust Networks[2] вузли працюють в шумоподібному режимі із стрибкоподібною перестройкою частоти в діапазоні 902-928 МГц. Відстань між вузлами складає 30-60 м в приміщенні і до 150 м поза приміщенням. Проте внаслідок малих розмірів мот розмір антени також повинен бути малий, і передавання даних необхідно вести на високій частоті, а це не завжди сумісно з вимогою малої споживаної потужності. До того ж, високочастотні трансивери – достатньо складні пристрої з схемами модуляції і демодуляції, смуговими фільтрами. Для передавання даних від великого числа вузлів потрібні додаткові схеми мультиплексованої передавання з часовим, частотним або кодовим розділенням. Все це істотно ускладнює завдання зниження споживаної потужності до необхідного рівня в декілька мікроват.

Безпровідна сенсорна мережа, розроблена авторами, належить до розподілених інформаційно-вимірювальних комп’ютеризованих програмно-апаратних засобів перетворення, дистанційного передавання, реєстрації та відображення даних про характеристики складних просторово-розподілених промислових об’єктів контролю та управління. Такі сенсорні мережі можуть бути

використані для віддаленого контролю та обліку витрати енергоносіїв (електроенергії, газу, стисненого повітря, води, пари, сипучих матеріалів), ідентифікації, станів технологічних установок, які характеризуються квазі–стаціонарними станами (електричні двигуни, крани, засувки, станки по обробці металів, та інших матеріалів), установки нафтогазового комплексу (буріння, видобутку, підготовки, зберігання, переробки та транспортування нафти та нафтопродуктів), а також транспорту, переробки, фасування та переробки продуктів харчування.

Дана мережа володіє кращими властивостями завадозхищенності та надійності, а також апаратна складність є меншою, ніж у відомих безпровідних сенсорних мережах.

Безпровідна сенсорна мережа, на передавальній стороні містить *k*-сенсорних вузлів, на входах яких є давачі (рис. 7). Вихід кожного давача підключений до сенсорного входу пристрою формування та цифрового опрацювання сигналів. Крім цього сенсорний вузол містить батарею автономного живлення. Реєстраційний вихід підключений до флеш-пам'яті реєстрації, а інформаційний вихід через трансивер підключений до передавальної антени, а на приймальній стороні містяться k-приймальних антен. Виходи антен підключені до входів відповідних пристроїв опрацювання отриманих сигналів. Виходи цих пристроїв підключені до відповідних входів концентратора. На передавальній стороні пристрій формування та цифрового опрацювання сигналів містить інтегрально-імпульсний ентропійний перетворювач Галуа, який включає перетворювач напруга/частота, генератор Галуа, вихід якого підключений до ентропійного перетворювача сигналів, вихід якого є інформаційним виходом пристрою, і до флеш пам'яті реєстрації даних (реєстраційний вихід). На приймальній стороні пристрій опрацювання сигналів містить пристрій визначення автокореляційної міри ентропії і пристрій демодуляції бітів Галуа.

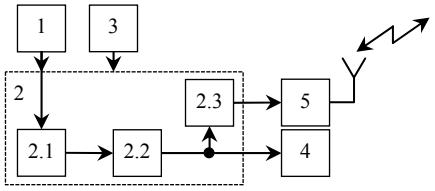


1 – розподілений об'єкт управління, 2 – сенсорні вузли з передавальною антеною, 3 – пристрій визначення ентропії з приймальною антеною, 4 –пристрій демодуляції бітів Галуа, 5 –концентратор.

Рис. 7 – Структурна схема сенсорної мережі

В основу роботи безпровідної сенсорної мережі покладено принцип передавання інформації на основі ортогональних ентропійно-маніпульованих сигналів та кодовому розділені сигналів у безпровідній лінії зв'язку та використанні сигнальних коректуючих входів у базисі Галуа.

Інформаційно-вимірювальні дані з об'єкту управління поступають на сенсорні вузли 1 з передавальної антеною 2(рис. 7), а саме на входи давачів 1 сенсорних вузлів безпровідної сенсорної мережі(рис. 8).



1 – давач, 2 – інтегрально-імпульсний ентропійний перетворювач Галуа 2.1 – перетворювач напруга/частотота, 2.2 – генератор Галуа; 2.3 – ентропійний перетворювач сигналу, 3 – блок живлення вузла, 4 – флеш-пам'ять для реєстрації даних, 5 – передавальна антена

Рис.8 – Структурну схему вузла безпровідної сенсорної мережі

Вихід кожного давача 1 підключений до відповідного входу інтегрально-імпульсного ентропійного перетворювача Галуа 2, який містить перетворювач типу напруга/частота 2.1, імпульси якого тактують роботу генератор Галуа 2.2, який з використанням кодової шкали Галуа по осі ординат формує асинхронний потік бітів Галуа в моменти пересічення квантованих значень осі ординат(рис. 9) [7].

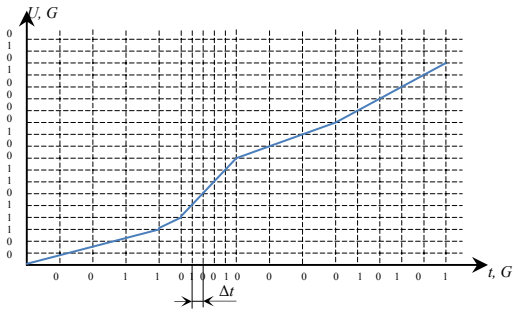


Рис. 9 – Інтегральне представлення сигналу з використанням кодової шкали Галуа

Після цього кодовані дані поступають на флеш-пам'ять реєстрації даних 4 та ентропійний перетворювач 2.3, який за допомогою шумоподібних сигналів кодує одиничні та нульові біти Галуа.

Суть методу формування та опрацювання

квазітрійкових сигналів зі змінною ентропією (рис. 10) полягає в тому, що двійковим символам інформаційного повідомлення ставиться у відповідність значення розподілу ентропії сигналу[11]. Так, в каналі зв'язку є постійна складова (рис. 10,а), яка використовується для позначення повторення, початку і кінця повідомлення. Інформаційним символам «0» (рис. 10,б) і «1» (рис. 10,в) ставиться у відповідність значення ентропії шумоподібного сигналу з маніпульованим математичним сподіванням.

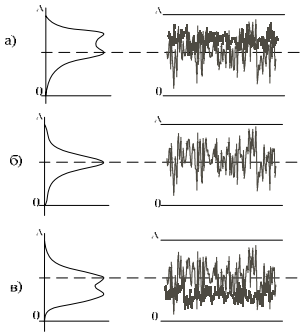


Рис. 10 – Представлення інформаційних повідомлень при маніпуляції квазітрійкових сигналів зі змінною ентропією: а) «1»; б) «синхро»; в) «0»

Приклад квазітрійкового сигналу зі змінною ентропією для інформаційного повідомлення розміром в 1 байт зображено на рис. 11, де: I – структура фрейма, II – реалізація фізичного рівня ентропійноманіпульованих сигналів, III – квазітрійковий код маніпульованих сигналів, IV – характеристики Гаусівського розподілу сигналів зі змінною ентропією.

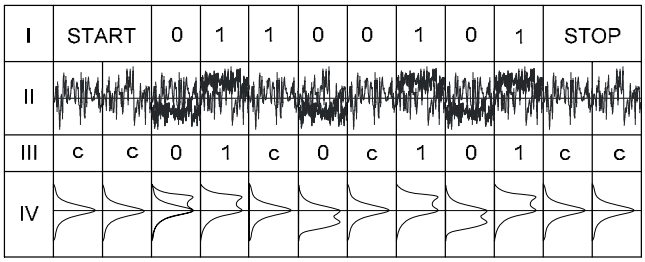


Рис. 11 – Приклад квазітрійкового сигналу зі змінною ентропією для інформаційного повідомлення розміром в 1 байт

Як видно із рис. 11. За допомогою квазітрійкового сигналу можна не змінюючи структуру сигналу організувати «старт» і «стоп» біти, а також виключити повторення інформаційних символів, що забезпечує якісну бітову синхронізацію.

Сформовані таким чином ентропійно-маніпульовані сигнали через трансивер 5 з передавальною антеною 6 дистанційно передаються на приймальну сторону.

На приймальній стороні (фіг. 8) міститься k-приймальних антен 3, кожна з яких здійснює приймання переданих сигналів і їх передавання на пристрій опрацювання автокореляційної міри ентропії 4 [12]. Опрацьований сигнал передається на пристрій демодуляції бітів Галуа 5, де відбувається перевірка даних на помилки, їх корекція у випадку наявності помилок у переданих даних та передавання інформації на відповідні входи концентратора 6.

В концентраторі *n* послідовно прийнятих бітів Галуа перетворюються в цифрові позиційні коди, які представляють покази відповідного сенсора.

5. ВИСНОВКИ

Викладені дослідження в галузі створення безпровідних сенсорних мереж є перспективним напрямом у їх розвитку, Крім цього застосування ентропійного підходу при реалізації методів маніпуляції сигналів на основі ШПС зі змінною ентропією забезпечує високий рівень завадозахищеності передавання даних в умовах інтенсивних промислових завад. Використання сигнальних коректуючих кодів для формування та передавання інформації створюють перспективу їх широкого застосування на в комп'ютерних системах, для підвищення завадостійкості і захисту від несанкціонованого доступу.

6. СПИСОК ЛІТЕРАТУРИ

[1] http://www.idtechex.com/research/reports/wireless-sensor-networks-2011-2021

[2] Y. M. Boyko, V. M. Lokazyuk, V. V. Mishan, Conceptual features of wireless sensor networks, *Bulletin of the Khmelnitsky National University*, 2 (2010), pp. 94-98 (in Ukrainian)

[3] Nykolaychuk Y.M. *Theory Sources*, Second edition, revised, Ternopil: JSC «Ternograph», 2010, 536 p. (in Ukrainian).

[4] Pohonets I. O., Voronych A. R., The method of construction of entropy model quasi-stationary sources, *Proceedings of International Symposium: Issues of Calculation Optimization*, Crimea, Katsiveli, Vol. 2, (2009), pp. 214-219 (in Ukrainian)

[5] Nykolaychuk Y. M., Voronych A. R., Theoretical basis of entropy measures and their applications in information technology signal formation and processing, *Optoelectronic information and energy technologies, International Science and Technology magazine*, (19) 1 (2010), pp. 50-64. (in Ukrainian)

[6] Y. M Nykolaychuk, A. R. Voronych, T. M. Hrynchyshyn, The theoretical foundations, principles of information formation and transmission base on signal corrective code, *Progress in science. Scientific Papers of Buchach Institute of Management*, Buchach, (6) 1 (2010), pp.41-49. (in Ukrainian)

[7] Nykolaychuk Y. M., Vozna N. Y., Pituh I. R., *Design of Specialized Computer Systems. Study Guide*, Ternopil: LLC «Terno-graph», 2010, 392 p. (in Ukrainian)

[8] US Pat.4376958 G11B5/09, *Modified Frequency Modulation*, Archibald M. Pettigrew (Glenrothes, GB6) / Elcomatic Limited (Glasgow, GB6). Appl. No.: 06/166, 777. Filed: Jul 8, 1980(in English)

[9] Patent for invention 96853 Ukraine IPC (2011.01) H03M13/00, *Method of Information Transmitting and Receiving*, Ivano-Frankivsk National Technical University of Oil and Gas, Nykolaychuk Y. M., Hrynchyshyn T. M., Voronych A. R. (Ukraine). Publ. 12.12.2011, Bull. № 23 (in Ukrainian)

[10] Petryshyn L. B., Nykolaychuk Y. M., Rolls synchronization and receiving digital messages methods in the basis of the Galois, *Proceedings of the 3rd Ukrainian Conf. with the automatic. management «Avtomahyka-96»*. Sevastopol SDTU, UAAU, (1996), pp. 76-77. (in Ukrainian)

[11] Voronych A. R. Entropy methods of formation and signal processing in specialized distributed computer systems, *Bulletin of Khmelnitsky National University*, Khmelnitsky, 4 (2010), pp. 69-72. (in Ukrainian)

[12] Patent for useful model 58743 Ukraine IPC (2006) G06F 17/15 (2011/01), *A device for determination of autocorrelation entropy measure*, Ivano-Frankivsk National Technical University of Oil and Gas, Nykolaychuk Y. M., Voronych A. R., Pohonets I. A. (Ukraine). Publ. 26.04.2011, Bull. № 8. (in Ukrainian)

Артур Воронич, *магістр з відзнакою за спеціальністю «Автоматизоване управління технологічними процесами» (2008) Івано-Франківського національного технічного університету нафти і газу, з 2010 р. аспірант кафедри Комп'ютерних систем та мереж за спеціальністю «Комп'ютерні системи та компоненти». Працює у напрямку застосування ентропійного підходу до методів формування та опрацювання сигналів та використання сигнальних коректуючих кодів в базисі Галуа.*

Ярослав Николайчук*, спеціаліст (1967), електрифікація та автоматизація видобутку, транспортування та зберігання нафти і газу, Львівський політехнічний інститут, к.т.н. (1980), елементи та пристрої обчислювальної техніки та систем керування, д.т.н. (1989), елементи та пристрої обчислювальної техніки та систем керування, професор кафедри автоматизованого управління (1993), Івано-Франківський інститут нафти і газу, директор Карпатського державного центру інформаційних засобів і технологій Національної академії наук України (1994), дійсний член Української академії національного прогресу (1995), завідувач кафедри спеціалізованих комп'ютерних систем (1999), керівник 18 захищених кандидатських дисертацій, консультант 1 захищеної докторської дисертації, член IEEE (2000), заступник голови спеціалізованої вченої ради К58.082.02 при ТНЕУ (2002).*

Наукові інтереси: спеціалізовані комп'ютерні системи, системи передавання даних, низові обчислювальні мережі.

Володимир Гладюк *магістр з відзнакою за спеціальністю «Комп'ютерні технології в управлінні та навчанні» (2009) Тернопільського національного економічного університету, з 2009 р. аспірант за спеціальністю «Електротехнічні комплекси та системи».*

Працює у напрямку методів та засобів опрацювання інформаційних характеристик пневмотранспортних систем у квазістаціонарних режимах.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

FORMATION AND PROCESSING OF ENTROPY-MANIPULATED CORRECTING SIGNAL CODES IN WIRELESS SENSOR NETWORKS

Artur Voronych [1)], Yaroslav Nykolaychuk [2)], Volodymyr Hladyuk [2)]

[1)] Ivano-Frankivsk National Technical University of Oil and Gas
15 Carpathian Str., 76019 Ivano-Frankivsk, archy.bear @ gmail.com
[2)] Ternopil National Economic University, 11 Lvivska Str., 46020 Ternopil

Abstract: *It is described a systematization of theoretical bases and analytical expressions for evaluation of entropy measures. It is showed a characteristic of signal corrective codes of Galois field. It is described the use of entropic manipulation and corrective signal codes in wireless sensor networks. It is founded that entropy approach of applying the corrective codes of Galois field is the most effective and promising for the formation and processing of signals during transmission in wireless sensor networks.*

Keywords: *wireless sensor networks, entropy, corrective signal codes, Galois.*

1. INTRODUCTION

In this article the wireless sensor networks are observed – one of the leading computer networking technology, which size of markets in 2011 is amounted to half a billion dollars [1]. The basis of such networks is a miniature computing and communication devices – motes (dust particles), or sensors. A set of sensors used depends on the functions performed by the wireless sensor networks [2].

2. THEORETICAL BASIS OF ENTROPY SIGNAL PROCESSING METHODS

In world practice to estimate information entropy of widely used information measure received by R. Hartley and K. Shannon. At the same time in practice, many estimates of entropy, which can be theoretic basis for methods of formation and signal processing [3, 4, 5].

Implementation of models based on autocorrelation entropy functions allows you to remove the existing functional limitations that are present in the calculation of entropy-based information measures R. Hartley and K. Shannon. Since the former does not take into account the statistical distribution flow of information data and the second considers only the probability of information sources states and does not consider the probability of transition from one state to another, which is implemented in information which assists in innovation as entropy based on the formula proposed by prof. Nykolaychuk Y.M. [3, 4, 5].

3. SIGNAL CORRECTIVE CODES IN GALOIS BASIS

An example of code that have high entropy features are corrective signal codes in Galois basis [3].

In some cases, Galois code G_2^n [3, 6] is formed by recurrent sequence: $G_{i+1} = G_i \oplus G_{i-n}$, where n – bit codon.

When transmitting and receiving information through signaling codes used manipulated signals are generated based on four characteristics that set in line with the elements of notification in accordance with codes Galois field [6].

It is known method of transmitting and receiving information based on the modified frequency modulation (MFM) [8], which uses four signal features: front increase ($\wedge$), front drop ($\vee$), corresponding to the character «1» and potential «+» the potential «-», corresponding to the character «0».When character «0» repetition for bit synchronization is also used front increase ($\wedge$) or decrease ($\vee$). However, this method does not allow to detect and correct errors. An other way to transmit and receive information [9] allows to detect and correct errors when receiving data bit-oriented by corrective properties of Galois codes without additional transmission checksum.

When transmitting and receiving information using manipulated signals formed on the basis of

four features which are set in line with the elements of an information message accordingly codes of Galois field. This principle of formation the signal code is what would go in the packet data units are numbered recurrent source Galois $G^2_{k.}$. And for items in the package data bit Galois «1» is transmitted front rise ($\wedge$) and Galois bits «0» is transmitted and do joint front ($\vee$). Bits of zeros in the data packets are numbered as a recurrent source of Galois $G^2_{k,}$. And for the district left in the packet data bit Galois «1» is transmitted potential «+» and Galois bit «0» is transmitted potentials scrap «-». As the four signs of manipulation in signal level can be properly used sets of four phases, frequency, noise-like signals and their combinations.

Error detection is based on bit-oriented numbering sequence of zeros and odes down and transmitted using code sequences Galois. If error found, the formula, where recurrent by check, which is held positions change the character zero (one), in the process of data and this symbol is replaced with the correct [9].Theoretical basis for improvement noise immunity corrective signal code shown in [10], based on the detected presence of a sequence of Galois across level recurrent properties.

4. SPIRAL PROPERTIES CORRECTIVE CODES IN GALOIS BASIS

May have the Galois code, which seems recurrent code: $G_2^4 = 1111010110010001111010110...$ We can show that code generated on the basis of above expression can be packed into a spiral, and for each of the four generators formed recurrent sequence that has the appropriate recurrent properties of the code in Galois basis.

The resulting spiral coded recurrent source spins preserving rekurentnosti over 12 characters, which allows correct errors in generating a spiral by the expression: $G_{i+1} = G_i \oplus G_{i-13}$.

Given the properties of spiral signal recurrent code can be effectively used for packet errors detected.

5. WIRELESS SENSORS NETWORKS

Wireless sensor networks, on transmission side, has *k*-sensor nodes at the inputs of which are sensors. Output of each sensor is connected to a touch input device formation and digital signal processing. Registration output connected to the flash memory registration, and information output by transceiver connected to the transmitting antenna and on receiving side are *k*-receiving antennas. The outputs of antennas connected to the inputs of the respective devices processing the received signals. The outputs of these devices are connected to corresponding concentrator inputs. On the transmitter side of the device formation and digital signal processing includes the integral-pulse entropy Galois converter, which includes the converter voltage / frequency, Galois generator, whose output is connected to the entropic transducer signal output which is an information output device. On the receiving side signal processing device includes a device definition autocorrelation measure of entropy and demodulation device of Galois bits.

The wireless sensor network work was based on the principle of information transmission based on orthogonal entropy-manipulated signals and code separated signals in a wireless communication line and use corrective signal inputs in Galois basis.

Information-measuring data with the object of arriving at the sensor nodes from transmitting antenna, namely sensors inputs of sensor node wireless sensor network.

Output of each sensor is connected to the corresponding integral-pulse converter entropic Galois containing transformer type voltage/frequency pulses whose work tact Galois generator that code using Galois scale on the vertical axis forms an asynchronous stream of bits in the Galois points of intersection quantized values vertical axis [7].

Then coded data enter in flash memory data recording and entropy converter that using noise-like signals and encodes a single zero bit Galois.

The method of formation and processing quasi-three signals with variable entropy is binary symbols that alert is assigned the value of entropy distribution of the signal [11].

Formed thus Entropy-manipulated signals via transceiver transmitting antenna of remotely transmitted to the receiving side.

On the receiving side contains *k*-receiving antennas, each of which performs receiving of incoming signals and transmitting them to the device processing autocorrelation entropy measure [12]. Processing signal is transmitted to the device demodulation Galois bit, where data errors test and their correction in case of errors in transmitted data and information transfer to the respective inputs of the concentrator.

In the concentrator *n* consistently taken bits Galois converted into digital positional codes that represent the impressions corresponding sensor.

6. SUMMARY

Described researches in the field of wireless sensor networks is promising direction in their development, addition of application entropic approach to the implementation of manipulation signals methods based noise-shaped signals variable

entropy provides a high level noise immunity data in intensive industrial noise. The use of signal correlation correcting codes for information formation and transmitting make a wide use of computer systems in order to increase immunity and protect against unauthorized access.

7. REFERENCES

[1] http://www.idtechex.com/research/reports/wireless-sensor-networks-2011-2021

[2] Y. M. Boyko, V. M. Lokazyuk, V. V. Mishan, Conceptual features of wireless sensor networks, *Bulletin of the Khmelnitsky National University*, 2 (2010), pp. 94-98 (in Ukrainian)

[3] Nykolaychuk Y.M. *Theory Sources*, Second edition, revised, Ternopil: JSC «Ternograph», 2010, 536 p. (in Ukrainian).

[4] Pohonets I. O., Voronych A. R., The method of construction of entropy model quasi-stationary sources, *Proceedings of International Symposium: Issues of Calculation Optimization*, Crimea, Katsiveli, Vol. 2, (2009), pp. 214-219 (in Ukrainian)

[5] Nykolaychuk Y. M., Voronych A. R., Theoretical basis of entropy measures and their applications in information technology signal formation and processing, *Optoelectronic information and energy technologies, International Science and Technology magazine*, (19) 1 (2010), pp. 50-64. (in Ukrainian)

[6] Y. M Nykolaychuk, A. R. Voronych, T. M. Hrynchyshyn, The theoretical foundations, principles of information formation and transmission base on signal corrective code, *Progress in science. Scientific Papers of Buchach Institute of Management*, Buchach, (6) 1 (2010), pp.41-49. (in Ukrainian)

[7] Nykolaychuk Y. M., Vozna N. Y., Pituh I. R., *Design of Specialized Computer Systems. Study Guide*, Ternopil: LLC «Terno-graph», 2010, 392 p. (in Ukrainian)

[8] US Pat.4376958 G11B5/09, *Modified Frequency Modulation*, Archibald M. Pettigrew (Glenrothes, GB6) / Elcomatic Limited (Glasgow, GB6). Appl. No.: 06/166, 777. Filed: Jul 8, 1980(in English)

[9] Patent for invention 96853 Ukraine IPC (2011.01) H03M13/00, *Method of Information Transmitting and Receiving*, Ivano-Frankivsk National Technical University of Oil and Gas, Nykolaychuk Y. M., Hrynchyshyn T. M., Voronych A. R. (Ukraine). Publ. 12.12.2011, Bull. № 23 (in Ukrainian)

[10] Petryshyn L. B., Nykolaychuk Y. M., Rolls synchronization and receiving digital messages methods in the basis of the Galois, *Proceedings of the 3rd Ukrainian Conf. with the automatic. management «Avtomahyka-96»*. Sevastopol SDTU, UAAU, (1996), pp. 76-77. (in Ukrainian)

[11] Voronych A. R. Entropy methods of formation and signal processing in specialized distributed computer systems, *Bulletin of Khmelnitsky National University*, Khmelnitsky, 4 (2010), pp. 69-72. (in Ukrainian)

[12] Patent for useful model 58743 Ukraine IPC (2006) G06F 17/15 (2011/01), *A device for determination of autocorrelation entropy measure*, Ivano-Frankivsk National Technical University of Oil and Gas, Nykolaychuk Y. M., Voronych A. R., Pohonets I. A. (Ukraine). Publ. 26.04.2011, Bull. № 8. (in Ukrainian)

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

СТРУКТУРНИЙ СИНТЕЗ ДИНАМІЧНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

Світлана Антощук, Дмитро Маєвський

Одеський національний політехнічний університет
просп. Шевченка, 1, 65044, м. Одеса,
e-mail: asg@ics.opu.ua, Dmitry.A.Maevsky@gmail.com

Резюме: *Розглянуто теоретичні основи та створено метод синтезу структури динамічних інформаційних систем із здатністю до адаптування при змінах предметної області. Здатність до адаптування забезпечується за рахунок виділення варіативної частини алгоритмів системи, що змінюються при змінах предметної області. Варіативна частина алгоритмів зберігається в базі даних системи, що забезпечує їх швидку заміну та підвищує надійність системи на етапі експлуатації.*

Ключові слова: *динамічна інформаційна система, синтез структури, варіативні алгоритми, надійність.*

STRUCTURAL SYNTHESIS OF DYNAMIC INFORMATION SYSTEMS

Svetlana Antoschuk, Dmitry Maevsky

Odessa National Polytechnic University
1, Shevchenko boulevard, 65044, Odessa, Ukraine
e-mail: asg@ics.opu.ua, Dmitry.A.Maevsky@gmail.com

Abstract: *Theoretical bases are considered and the method of synthesis of the dynamic information systems structure is created with a capacity for adaptation at the changes of a subject domain. A capacity for adaptation is provided due to the selection of a variation part of algorithms of the system that changing at the changing the subject domain. The variation part of algorithms is kept in the database of the system that provides their rapid replacement and promotes a fail safety on the stage of exploitation.*

Keywords: *dynamic information system, synthesis of structure, variant algorithms, reliability.*

ВСТУП

Інформаційні системи, предметна область (ПрО) яких зазнає змін на етапі експлуатації є дуже розповсюдженими. Будемо називати такі системи динамічними інформаційними системами (ДІС). Типовими представниками ДІС є облікові інформаційні системи, ПрО яких є господарче законодавство України. За даними Державного управління статистики, загальна кількість підприємств, що використовують облікові системи сягає більш ніж дев'ятисот тисяч. Водночас, як показали дослідження, закони та підзаконні акти в цій сфері змінюються в середньому один раз на десять днів [1]. Така частота змін, зважаючи на дуже велику кількість ДІС, породжує проблему забезпечення їх безперервної роботи під час експлуатації. Простої ДІС, що виникають при їх оновленні, негативно впливають на їх надійність, можуть призвести до призупинення виробничих процесів та значних матеріальних втрат. Збереження надійності на цьому етапі можливо лише за рахунок зниження часу простоїв при оновленнях.

Сучасні підходи до структурної побудови усіх ІС передбачають інтеграцію усіх програмних модулів у один великий за обсягом файл [2, 3]. Таку, побудовану за традиційним підходом структуру ІС назвемо статичною структурою.

Треба відмітити, що саме статична структура побудови сприяє виникненню негативного фактору, що збільшує час простоїв ДІС при оновленнях. Дійсно, усі алгоритми, що реалізують функціональні можливості ДІС, входять до складу їх ПЗ та не невід'ємними від нього. При оновленнях, що торкаються, зазвичай

не усієї множини алгоритмів, тобто частки ПЗ, зміни потребує усе воно разом. Це призводить до виникнення таких негативних факторів. По-перше, обсяг інформаційного пакету, що потрібен для оновлення такої ДІС невиправдано збільшується. По-друге, збільшується, відповідно, й час, що потрібен для передачі пакету оновлень до користувачів. У період призупинення для оновлення ДІС частково або повністю перестає виконувати свої функції, що також сприяє зменшенню їх надійності та виникненню ризиків аварій та техногенних катастроф, особливо у галузях безперервного циклу роботи (енергетика, транспорт, банківські системи). По-третє, передача великого обсягу інформації до великої кількості користувачів обмежена, пропускною здатністю існуючих каналів зв'язку. Для з'ясування істотності проблем, що породжуються цим фактором, розглянемо ОІС «Бухгалтерський облік для України», яку розроблено фірмою «1С». Усі програмні модулі та екранні форми цієї ОІС розміщено в одному компаунд-файлі розміром близько десяти мегабайт [4]. Ця ОІС широко використовується більшістю підприємств та організацій України (більше, ніж дев'ятисот тисяч). Тобто, при оновленнях цієї ОІС до серверу з файлом оновлень розміром у десять мегабайт намагатимуться отримати доступ кілька сотень тисяч абонентів одночасно. Зважаючи на обмежену пропускну здатність каналів зв'язку, виникає проблема зниження надійності ДІС за рахунок збільшення часу на їх обслуговування ДІС.

Необхідність передачі значного за обсягом пакету оновлень до великої кількості користувачів фактично еквівалентна розподіленій атаці типу «відмова в обслуговуванні» (DDoS – атаці) на сервер провайдера, де розміщені оновлення. Як відомо, атаки такого типу можуть призвести повної відмови в обслуговуванні усіх клієнтів, які користуються послугами цього провайдера, що призводить до дуже значних фінансових втрат. Так, за дослідженнями компанії Forrester Research, проведеними у 2009 році, якщо на сервері розміщені WEB-сервіси електронної комерції, то призупинення роботи такого сервера коштує від 190 до 650 тисяч доларів за годину.

Таким чином, на етапі експлуатації ДІС виникають такі протиріччя:

– протиріччя між потребою зменшення простоїв ДІС на етапі експлуатації та необхідністю тривалого призупинення їх роботи для виконання оновлень при змінах ПрО;

– протиріччя між необхідністю одночасної передачі великих за обсягом інформаційних пакетів для оновлення ДІС до великої кількості користувачів, та обмеженою пропускною здатністю каналів зв'язку, що призводить до зниження надійності ДІС на етапі експлуатації за рахунок збільшення часу простоїв при оновленнях.

Розв'язати ці протиріччя можна за рахунок спеціально побудованої структури ДІС, яка дозволяє адаптування системи до змінних вимог предметної області. Тому задача теоретичного обґрунтування та створення методів побудови такої структури ДІС є актуальною.

1. ОСНОВНА КОНЦЕПЦІЯ СТРУКТУРНОЇ ПОБУДОВИ ДІС

Основна концепція побудови структури ДІС, яка дозволяє швидку адаптацію при зміні вимог ПрО, полягає в виділенні із всієї множини алгоритмів обробки інформації в ДІС її варіативної підмножини, саме яка й зазнає змін при змінах ПрО. Основною особливістю запропонованої структури є зберігання варіативної підмножини алгоритмів не в загальному програмному коді, а відокремлено від нього, в інформаційній базі системи у вигляді не трансльованого або частково трансльованого коду (так званого «р-коду»). Такий код займає значно менше місця, ніж відповідний йому трансльований двійковий код, що вирішує проблему його швидкої передачі до великої кількості користувачів, що, як показали дослідження, до двох разів зменшує коефіцієнт простоїв ДІС при оновленні.

На самому вищому рівні будь-яку ІС можна розглядати як сукупність двох складових: програмного забезпечення та інформаційної бази. Кожна з цих складових не існує сама по собі. Для їх об'єднання в систему між ними повинні існувати певні взаємозв'язки – певні частини ПЗ певним образом співвідносяться із певними об'єктами ІБ. Тоді модель структури будь-якої ІС можна представити як трійку:

$$S = \{A, I, L\}, \qquad (1)$$

де A – множина усіх алгоритмів ДІС, I – множина усіх структурних одиниць інформаційної бази (ІБ), а L – множина усіх зв'язків між алгоритмами та структурними одиницями ІБ.

Але у випадку ДІС, при змінах ПрО, зміни торкаються тільки певної підмножини A_v варіативних алгоритмів, що змінюються: $A_v \in A$. Проведений аналіз показав, що потужність множини A_v складає близько п'яти відсотків від потужності A. Так, наприклад, в обліковій ДІС

«АгроКомплекс» [5], міститься загалом 7513 програмних модулів, що реалізують алгоритми ДІС. З них варіативна частина складає 360 алгоритмів, тобто лише 4,8 відсотки.

Різницю

$$A_c = A \setminus A_v,\ \{a_c \in A \mid a_c \notin A_v\}$$

будемо називати множиною сталих алгоритмів, що не змінюються при змінах ПрО.

У свою чергу, множина I для ДІС може бути представлена як $I = T \bigcup R$, де T – множина усіх інформативних таблиць ІБ, а R – множина накопичувальних регістрів. Тому модель (1) можна представити як:

$$S = \{\{A_c, A_v\}, \{T, R\}, L\}. \quad (2)$$

Згідно з цією моделлю, при змінах ПрО ДІС, зміні підлягають усі алгоритми множини $A = A_c \bigcup A_v$, у той час, як реально змінюється не більше п'яти відсотків алгоритмів. Тому включення варіативної підмножини A_v до A призводить до протиріччя між необхідністю одночасної передачі великих обсягів інформації для оновлення ДІС до великої кількості користувачів та обмеженою пропускною здатністю каналів зв'язку. Це, у свою чергу, викликає зниження надійності ДІС за рахунок збільшення часу їх простоїв під час оновлення.

Для розв'язання цього протиріччя зробимо перегрупування множин A та I таким чином:

$$A' = A \setminus A_v.$$

Тут нова множина A' утворена як різниця між загальною множиною усіх алгоритмів ДІС та множиною варіативних алгоритмів. Фактично, цією операцією ми відокремлюємо варіативну частину від множини усіх алгоритмів, отримуючи окрему сталу їх частину:

$$A' = A_c.$$

Залишок від різниці $A \setminus A'$ являє собою варіативну частину алгоритмів, яку включимо до множини I, утворюючи нову розширену множину I':

$$I' = I \bigcup (A \setminus A').$$

Тоді структуру (2) можна переписати як:

$$S = \{\{A_c\}, \{T, R, A_v\}, L\}. \quad (3)$$

Структура (3) відрізняється від традиційної структури (2) тим, що варіативні алгоритми переносяться до складу інформаційної бази ДІС. Відокремлення множини сталих алгоритмів за моделлю (3) дозволяє звести процес оновлення ДІС лише до оновлення тієї варіативної підмножини A_v, що змінилася. При чому відмітимо, що тепер при оновленнях зміні підлягають навіть не усі варіативні алгоритми, а саме та їх частка, що змінилася. Таке оновлення потребує передачі по каналах зв'язку значно меншого обсягу даних (декілька десятків кілобайт), що призводить до значного зменшення часу простоїв ДІС при оновленнях.

Розглянемо гнучку структуру ДІС (рис. 1), що відповідає моделі (3) та дозволяє виконувати адаптацію до змін ПрО [6].

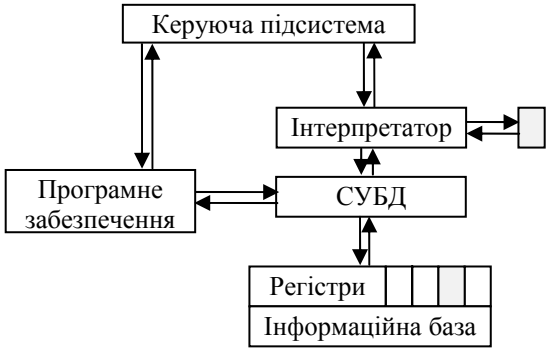


Рис. 1 – Структура ДІС з відокремленим програмним кодом

Робота ДІС із змінною структурою відбувається таким чином [7, 8]. Система отримує запит користувача на виконання дії, що потребує роботи певного алгоритму. Цей запит, як і в попередньому випадку традиційної побудови ДІС, аналізується керуючою підсистемою, яка визначає, до якої підмножини алгоритмів відноситься цей алгоритм – до варіативної або сталої.

У випадку, коли алгоритм відноситься до варіативної множини, керуюча підсистема видає відповідний запит до СУБД, яка знаходить в інформаційній базі потрібний програмний код варіативного алгоритму.

Але далі процес змінюється. До складу ДІС, яку побудовано за новою структурою із можливістю адаптування до змін ПрО входить нова складова – інтерпретуюча підсистема. Знайдений в ІБ ДІС код варіативного алгоритму передається на виконання до цієї інтерпретуючої підсистеми. Як результат його виконання може змінюватись стан інформаційної бази ДІС або формуватись звіти для користувача.

У випадку, коли дія системи повинна виконуватись статичним алгоритмом, робота ДІС відбувається так же, як і завжди. В цьому випадку система просто звертається до потрібного програмного модуля, що входить до складу ПЗ ДІС, а результат його виконання змінює (при потребі) стан ІБ, або видається користувачеві.

Необхідність застосування інтерпретатора та пов'язане з цим можливе зниження швидкодії системи не може вважатись вадою

запропонованого підходу. Проведені численні дослідження довели, що зниження швидкодії за рахунок інтерпретації не є суттєвим [9, 10, 11].

Керуюча підсистема отримує запит користувача на виконання певної дії та визначає, до якої множини алгоритмів належить алгоритм, який повинен виконувати цю дію. Якщо обрано алгоритм із множини варіативних, керуюча підсистема видає відповідну інформацію до СУБД, яка знаходить в інформаційній базі потрібний програмний код та передає його на виконання до інтерпретуючої підсистеми. На рис. 1 програмні модулі варіативної підмножини показано сірим кольором. Інтерпретатор виконує цей програмний код. Як результат його виконання може змінюватись стан інформаційної бази ДІС або формуватись звіти для користувача. У випадку вибору алгоритму із множини статичних, робота ДІС виконується традиційним чином – керування передається відповідній ділянці програмному коду, що є складовою ПЗ ДІС.

2. МЕТОД СТРУКТУРНОГО СИНТЕЗУ ДІС

Для побудови гнучкої структури ДІС із здатністю до адаптації розроблено відповідний метод, який полягає в виділенні варіативної частини алгоритмів на підставі аналізу ПрО, та реалізується послідовністю взаємопов'язаних етапів, кожен з котрих розписано на рівні операцій та елементарних дій (рис. 2).

Метод полягає в послідовному виконанні таких етапів: структуризація системи; моделювання управління підсистемами; декомпозиція підсистем на структурні одиниці; виділення структурних одиниць, що змінюються; розподіл структурних одиниць.

Етап 1 починається із аналізу технічного завдання на розробку ДІС та її ПрО. Визначається множина елементів ПрО та провадиться вибір тієї підмножини, що відображається в ДІС, що проектується. Необхідність цього етапу пов'язана з тим, що ПрО ДІС є більш ширшою, аніж ДІС, що відображає цю ПрО. ДІС здатна відобразити лише формалізовані взаємозв'язки між об'єктами її ПрО.

Після вибору підмножини об'єктів ПрО, що буде відображатися у ДІС, провадиться їх систематизація (кластерізація) за ідентифікаційною ознакою кластерів. Наприклад, в облікових ДІС такими кластерами об'єктів є кластери «Бухгалтерський облік», «Податковий облік», «Оперативний облік». Це виділення кластерів й є змістом операції 1.1 на рис. 2. На підставі аналізу ПрО та вимог ТЗ, обрані на цьому етапі кластери, й будуть визначати структуру майбутньої ДІС.

В операції 1.2 виконується виявлення взаємозв'язків між обраними на попередньому етапі кластерами та формування підсистем майбутньої ДІС. В нашому випадку такими підсистемами виступатимуть обрані кластери ПрО – «Бухгалтерський облік», «Податковий облік», «Оперативний облік».

На цьому етапі слід мати на увазі, що кластери ПрО та відповідні їм підсистеми ДІС не існують самі по собі, а мають між собою тісні взаємозв'язки. В нашому випадку облікової ДІС такі взаємозв'язки існують між підсистемами «Бухгалтерський облік» та «Податковий облік». Їх існування відображає взаємозв'язки між відповідними розділами обліку в ПрО, що відображено в відповідних нормативних актах України. На підставі опису ПрО формується матриця взаємозв'язків між кластерами ПрО та відповідним ним підсистемам ДІС. В наступній операції 1.3, на підставі визначеної множини підсистем формується матриця зв'язків між цими підсистемами. Для формування матриці проводиться визначення окремих об'єктів таких підсистем та визначається (за ознакою «0» чи «1») наявність чи відсутність зв'язків між об'єктами визначених підсистем. Наприклад, об'єкт підсистеми «Бухгалтерський облік» – «Господарська операція» та підсистеми «Податковий облік» – «Податкові зобов'язання» та «Податковий кредит» мають однозначні зв'язки – кожна операція, що проведена по бухгалтерському обліку, має бути відображена й в обліку податковому.

На етапі 2 з'ясовуються взаємні зв'язки між підсистемами, які будуть визначати передачу управління та даних в спроектованій ДІС. Етап полягає в побудові діаграм потоків даних (ДПД) між підсистемами й моделюванні передачі даних на підставі створених діаграм. При моделюванні визначаються типи та структури даних, що циркулюють між підсистемами майбутньої ДІС. Слід відмітити, що під даними на цьому етапі маються на увазі не тільки елементарні типи, таки, як число, рядок та інші. Між окремими підсистемами можливий обмін більш складно агрегованими даними – колекції та структури різнотипних значень. На етапі 3 виконується подальша деталізація структури. Тепер деталізується внутрішня структура окремих підсистем. Метою цього етапу є визначення складу підсистеми та її структури. Склад підсистеми визначає набір програмних модулів, що виконують основні алгоритми по обробці даних.

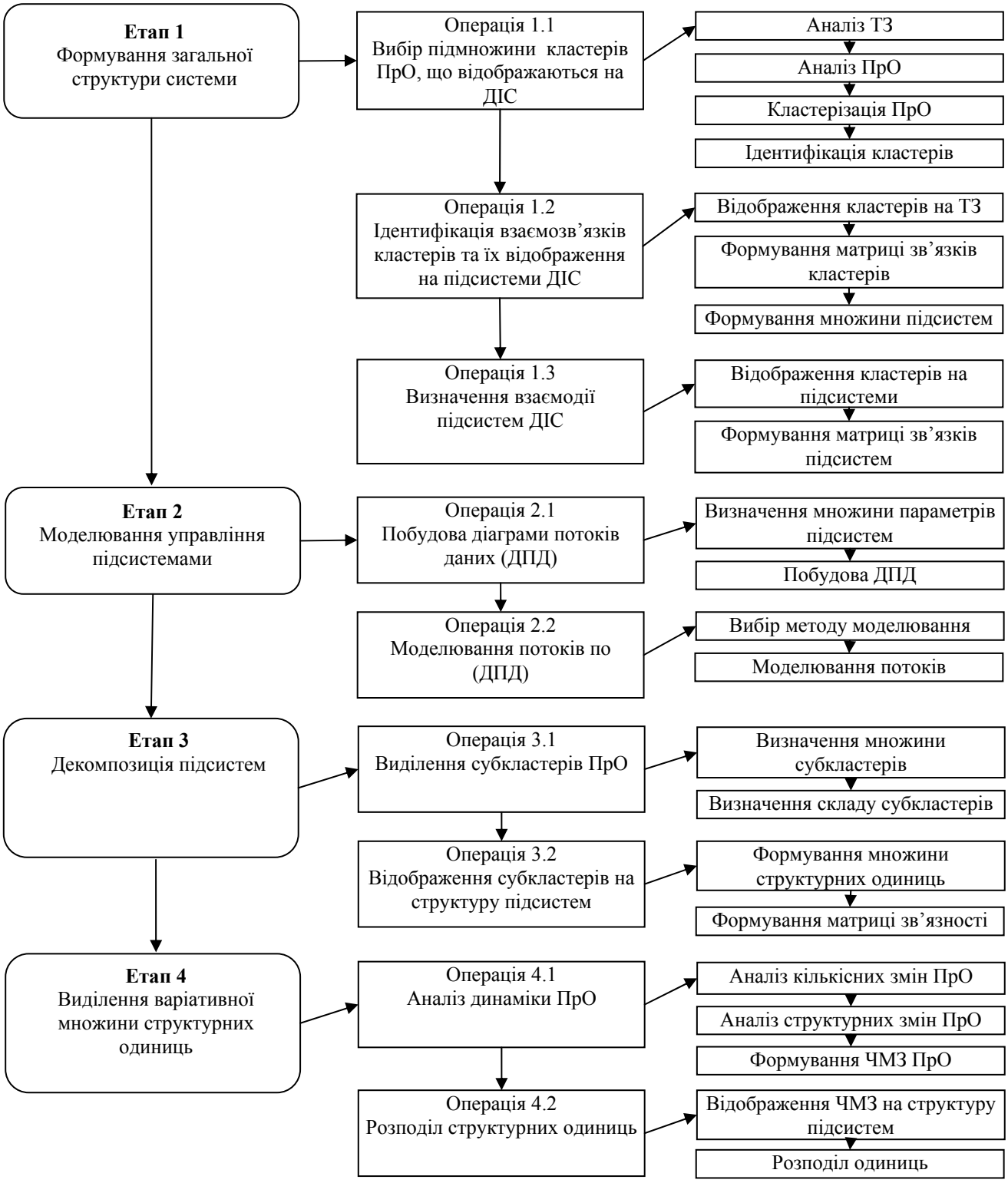


Рис. 2 – Етапи структурного синтезу ДІС

Структуру підсистеми визначають взаємні зв'язки між цими модулями. В операціях 3.1 та 3.2 такі модулі позначено як «субкластери». На етапі 3.2 з'ясовується ступінь зв'язності (СЗ) між окремими модулями підсистеми. Існує сім типів зв'язності [7]:

1. Зв'язність по збігу (СЗ = 0). Відсутні явно виражені внутрішні зв'язки.

2. Логічна зв'язність (СЗ=1). Частини модуля об'єднані за принципом функціональної подібності.

3. Часова зв'язність (СЗ=3). Частини модуля не пов'язані, але потрібні в один і той же період роботи системи.

4. Процедурна зв'язність (СЗ=5). Частини модуля пов'язані порядком виконуваних дій, що

реалізовують деякий сценарій поведінки.

5. Комунікативна зв'язність (СЗ=7). Частини модуля пов'язані за даними (працюють з однією і тією ж структурою даних).

6. Інформаційна (послідовна) зв'язність (СЗ=9). Вихідні дані однієї частини використовуються як вхідні в іншій частині модуля.

7. Функціональна зв'язність (СЗ=10). Частини модуля разом реалізують одну функцію.

Для визначення рівня зв'язності модуля (показника СС) використовується алгоритм, запропонований в [8]. Він містить наступні кроки:

1. Якщо модуль – одинична проблемно-орієнтована функція, то СС=10; кінець алгоритму. Інакше перейти до пункту 2.

2. Якщо дії усередині модуля пов'язані між собою, то перейти до пункту 3. Якщо дії усередині модуля незалежні, то перейти до пункту 6.

3. Якщо дії усередині модуля пов'язані даними, то перейти до пункту 5.

4. Якщо дії усередині модуля пов'язані потоком управління, перейти до пункту 5.

5. Якщо порядок дій усередині модуля важливий, то СС = 9, інакше СС = 7. Кінець алгоритму.

6. Якщо порядок дій усередині модуля важливий, то СС = 5, інакше СС = 3. Кінець алгоритму.

7. Якщо дії усередині модуля належать до однієї категорії, то СС = 1. Якщо дії усередині модуля не належать до однієї категорії, то СС = 0. Кінець алгоритму.

На етапі 4 виконується розподіл множини модулів кожної підсистеми на варіативну та сталу підмножини. Для цього в операції 4.1 виконується частотний аналіз змін ПрО й будується частотна матриця-стовпець змін, яка для кожного модулю показує відносну частоту його змін при змінах ПрО. Відносна частота обчислюється як відношення кількості змін заданого модулю до загальної кількості змін ПрО.

В операції 4.2. власне й виконується розподіл на варіативну та сталу підмножину. До варіативної підмножини виносяться алгоритми, що мають найменший ступінь зв'язності та найбільшу частоту змін.

Слід відмітити, що більшість етапів запропонованого методу не можуть бути формалізовані. Тому побудова автоматичної системи структурного синтезу ДІС на сьогодні є невирішеною проблемою.

3. ПРИКЛАД РЕАЛІЗАЦІЇ МЕТОДУ СТРУКТУРНОГО СИНТЕЗУ ДІС

Запропонований метод структурного синтезу реалізовано при створенні облікової ДІС «АгроКомплекс» для автоматизації усіх видів обліку на сільськогосподарських підприємствах. Облік у сільському господарстві має дуже складну ПрО, опис якої наводиться як в законодавчих актах України, так і великій кількості підзаконних інструкцій та постанов різних відомств. Це робить автоматизацію обліку у сільському господарстві складною та трудомісткою задачею.

Метою створення здатної до адаптації структури ДІС «АгроКомплекс» була необхідність забезпечити зниження часу простоїв системи при таких змінах ПрО:

– зміна алгоритмів проведення існуючих та створення нових господарських операцій;

– зміна обраної підприємством системи оподаткування;

– зміни статей податкового обліку при проведенні господарських операцій;

– можливість вибору, зміни або створення власних алгоритмів надання пільг при стягненні податкових коштів;

– можливість вибору та зміни існуючих або створення нових правил розмежування доступу до об'єктів та даних інформаційної бази ДІС.

Практична експлуатація ДІС «АгроКомплекс» в дев'яти господарствах Одеської області показала, що побудова ДІС за запропонованою структурою дозволила скоротити час простоїв на оновлення через мережу Інтернет. Так, наприклад, впровадження облікової ДІС «АгроКомплекс» із гнучкою структурою, в якій алгоритми обробки, що змінюються винесено до складу бази даних дозволило в 1,87 рази скоротити час й коефіцієнт простоїв при оновленні через мережу Інтернет та зменшити обсяг інформаційного пакету оновлень з 8,1 Мбайт до 196 Кбайт.

Впровадження ДІС «АгроКомплекс» із гнучкою структурою в сільськогосподарському підприємстві «Агролідер Т-7» (м. Татарбунари Одеської області) із підключенням до мережі Інтернет через звичайну телефонну лінію зв'язку дозволило скоротити коефіцієнт простоїв на оновлення через в 3,2 рази при тому ж зменшенні обсягу інформаційного пакету. Для порівняння показників використано аналогічну за призначенням серійну інформаційну систему «1С:Предприятие 8. Управление сельскохозяйственным предприятием» фірми 1С [12].

4. ВИСНОВКИ

В статті обґрунтовано проблеми та протиріччя, що виникають на етапі експлуатації динамічних інформаційних систем при змінах їх предметної області. Показано, що зменшення часу простоїв при оновленнях ДІС можливе тільки за рахунок скорочення обсягу інформаційного пакету, який передається від розробника ДІС до користувачів. Розроблено концепцію, за якою частина програмного коду ДІС зберігається як текст в базі даних системи, що дозволяє змінювати лише окремі ділянки програмного забезпечення ДІС. Створено метод побудови ДІС із гнучкою, здатною до адаптування при змінах ПрО структурою, використання якої призводить до зменшення часу простоїв та забезпечення, за рахунок цього, надійності при їх експлуатації; на базі створеної концепції розроблено методи побудови ДІС із гнучкою структурою для забезпечення надійності при їх оновленнях на етапі експлуатації.

Їх використання дозволило розв'язати протиріччя між потребою зменшення простоїв ДІС на етапі експлуатації та необхідністю тривалого призупинення їх роботи для виконання оновлень при змінах ПрО, а також протиріччя між необхідністю одночасної передачі великих за обсягом інформаційних пакетів для оновлення ДІС до великої кількості користувачів, та обмеженою пропускною здатністю каналів зв'язку.

Практичне впровадження ДІС із запропонованою структурою довело можливість збільшення надійності ДІС на етапі експлуатації за рахунок зниження часу простоїв при оновленнях.

5. СПИСОК ЛІТЕРАТУРИ

[1] D. A. Maevsky, H. J. Maevskaya, V. N. Antoschuk, Adaptation of functioning of the accounting information systems to the requirements normatively-legal acts, *Electrical machine-building and electrical equipment*, 72 (2009), pp. 153-160. (in Ukrainian)

[2] V. N. Petrov, *Information Systems*, SPb: Piter, 2003, 688 p. (in Russian)

[3] V. A. Matiash, A. V. Nikandrov, V. A. Putilov, A. E. Fedorov, V. V. Filchakov, *Structural Analysis at Software of the Real-Time Systems Development*, Apatity, KF PetrSU, 1997, 78 p. (in Russian)

[4] M. G. Radchenko, E. J. Hrustaliova, *Architecture and Work with Data «1C:Enterprise 8.2»*, Moscow, 1C Publishing, 2010, 268 p. (in Russian)

[5] D. A. Maevsky, T. J. Tintulova, V. N. Antoshchuk, The information system agricultural «Complex» for a book-keeping and operative account in agriculture, *Agrarian announcer of Black sea. Engineering sciences*, 48 (2009), pp. 151-156. (in Ukrainian)

[6] D. A. Maevsky, V. N. Antoschuk, Methodology of construction of the information systems is with the variable algorithm of treatment of information, *Works of the Luhansk branch of the International academy of Automation*, 19 (2009), pp. 111-112. (in Russian)

[7] V. Chistiakov, Modern facilities of development: comparison of productivity, *Technology «Client-server»*, 3 (2001), pp. 18-48. (in Russian)

[8] V. Chistiakov, Who is most rapid today? http://www.rsdn.ru/article/devtools/perftest.xml . (in Russian)

[9] M. Guillaume, The speed, size and dependability of programming languages, http://blog.gmarceau. qc.ca/2009/05/speed-size-and-dependability-of. html. – 2009. (in English)

[10] S. Amjan, R. K. Reddy, B. Manda, C. Prakashini, K. Deepthi, An empirical validation of object-oriented design metrics for fault prediction, *Journal of Computer Science*, (7) 4 (2008), pp. 571-577. (in English)

[11] Bases of planning of the programming systems. http://www-csd.univer.kharkov.ua/content/files/cat16/opps.pdf. (in Russian)

[12] 1C:Enterprize 8. Management by an agricultural enterprise, http://solutions.1c.ru/catalog/agr-enterprise. (in Russian)

Антощук Світлана Григорівна, *директор ІКС ОНПУ, завідуюча кафедрою інформаційних систем, д.т.н., професор. Після закінчення в 1981 році Одеського політехнічного інституту (зараз – ОНПУ) працювала в ньому. Пройшла шлях від інженера та асистента до професора. В 2005 році одержала наукову ступінь доктора технічних наук та вчене звання професора кафедри інформаційних систем. Нагороджена нагрудними знаками МОНУ «Відмінник освіти України» та «За наукові досягнення» та грамотами Міністерства освіти і науки України. Автор більш 150 наукових та науково-методичних робіт*

Сфера наукових інтересів – автоматизовані управляючі системи з обробкою візуальної

інформації; інформаційні системи; системи штучного інтелекту.

Маєвський Дмитро Андрійович, *завідуючий кафедрою теоретичних основ та загальної електротехніки ОНПУ, к.т.н., доцент. В 1973 поступив, а в 1978 році з відзнакою закінчив Одеський політехнічний інститут (зараз – ОНПУ). Працював на кафедрі теоретичних основ електротехніки (нині – теоретичних основ та загальної електротехніки) на посадах асистента, старшого викладача, доцента. В 1986 році захистив кандидатську дисертацію, присвячену розробці методів проектування друкованих плат обчислювальної техніки надвисокої швидкодії. В 1991 році отримав звання доцента. З 2005 року очолює кафедру.*

Сфера наукових інтересів – надійність програмного забезпечення, перехідні процеси в полоскових лініях зв'язку. Опублікував 46 наукових праць та 20 методичних вказівок.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

STRUCTURAL SYNTHESIS OF DYNAMIC INFORMATION SYSTEMS

Svetlana Antoschuk, Dmitry Maevsky

Odessa National Polytechnic University
1, Shevchenko boulevard, 65044, Odessa, Ukraine
e-mail: asg@ics.opu.ua, Dmitry.A.Maevsky@gmail.com

Abstract: *Theoretical bases are considered and the method of synthesis of the dynamic information systems structure is created with a capacity for adaptation at the changes of a subject domain. A capacity for adaptation is provided due to the selection of a variation part of algorithms of the system that changing at the changing the subject domain. The variation part of algorithms is kept in the database of the system that provides their rapid replacement and promotes a fail safety on the stage of exploitation.*

Keywords: *dynamic information system, synthesis of structure, variant algorithms, reliability.*

1. INTRODUCTION

The article is dedicated to the description of the method of synthesis of the structure of dynamic information systems (DIS). Dynamic information systems mean information systems subject domain (SD) of which changes at the stage of exploitation. At present time DIS are most widely used systems. At the same time, there is a necessity of frequent alteration of DIS which leads to temporary interruption of its work. Timeouts of DIS which happen during the alterations reduce their reliability, may lead to interruption of industrial processes and significant material losses. Retention of the reliability is only possible through the timeout reduction during the alterations.

Retention of the reliability may be made possible by means of a specially designed DIS structure which allows adaptation of the system to the varying requirements of the SD.

2. MAIN CONCEPT

Basic concept of designing a DIS structure which allows fast adaptation for changed requirements of SD is in the selection of a variable subset of algorithms, which change during the alterations of SD, from the multitude of algorithms for data processing. It is suggested to save the program codes of variable algorithms in the system database (DB) in the form of source of partly compiled code («p-code»).

During DIS updates it is enough to alter one or several variable algorithms in the DB. Main program DIS modules will not be changed which allows sending smaller update data packets to the users.

3. DIS STRUCTURAL SYNTHESIS METHOD

Operation of DIS with the suggested structure is illustrated in the fig. 1.

User's query is analyzed by the managing subsystem for determining the subset of the processing algorithms.

If the algorithm is variable, it is extracted from the DIS DB and is handed over to the interpreter. Constant algorithms are carried out immediately.

To design the structure of DIS which allows adaptation for changed requirements of SD a method of structural synthesis was developed (fig.2).

Its objective is to form a variable subset of algorithms.

Variable algorithms refer to DIS algorithms which have the least connectivity with others and are the most changed.

4. EXAMPLE OF STRUCTURAL SYNTHESIS METHOD

Suggested method of structural synthesis is realized in the account DIS «Agrocomplex» for automation of all kinds of accounting at the agricultural enterprises.

Practical exploitation of DIS «Agrocomplex» at 9 enterprises of Odessa region has showed that designing a DIS with the suggested structure

allowed reducing the coefficient of timeouts from 1,8 up to 3,2 times as well as diminishing the size of the updating data packet from 8,1 MB up to 196 KB.

5. CONCLUSIONS

A concept and method of DIS design with a flexible and adaptable at the changes of SD structure is presented in the article. It allowed reducing of timeouts and increasing of these systems reliability.

This approach made it possible to solve the controversy between the necessity of system program updating and reducing the timeouts of DIS. Reducing the size of data packets and dicreasing of requirements for the network equipment are the advantages of the suggested system.

6. REFERENCES

[1] D. A. Maevsky, H. J. Maevskaya, V. N. Antoschuk, Adaptation of functioning of the accounting information systems to the requirements normatively-legal acts, *Electrical machine-building and electrical equipment*, 72 (2009), pp. 153-160. (in Ukrainian)

[2] V. N. Petrov, *Information Systems*, SPb: Piter, 2003, 688 p. (in Russian)

[3] V. A. Matiash, A. V. Nikandrov, V. A. Putilov, A. E. Fedorov, V. V. Filchakov, *Structural Analysis at Software of the Real-Time Systems Development*, Apatity, KF PetrSU, 1997, 78 p. (in Russian)

[4] M. G. Radchenko, E. J. Hrustaliova, *Architecture and Work with Data «1C:Enterprise 8.2»*, Moscow, 1C Publishing, 2010, 268 p. (in Russian)

[5] D. A. Maevsky, T. J. Tintulova, V. N. Antoshchuk, The information system agricultural «Complex» for a book-keeping and operative account in agriculture, *Agrarian announcer of Black sea. Engineering sciences*, 48 (2009), pp. 151-156. (in Ukrainian)

[6] D. A. Maevsky, V. N. Antoschuk, Methodology of construction of the information systems is with the variable algorithm of treatment of information, *Works of the Luhansk branch of the International academy of Automation*, 19 (2009), pp. 111-112. (in Russian)

[7] V. Chistiakov, Modern facilities of development: comparison of productivity, *Technology «Client-server»*, 3 (2001), pp. 18-48. (in Russian)

[8] V. Chistiakov, Who is most rapid today? http://www.rsdn.ru/article/devtools/perftest.xml . (in Russian)

[9] M. Guillaume, The speed, size and dependability of programming languages, http://blog.gmarceau. qc.ca/2009/05/speed-size-and-dependability-of. html. – 2009. (in English)

[10] S. Amjan, R. K. Reddy, B. Manda, C. Prakashini, K. Deepthi, An empirical validation of object-oriented design metrics for fault prediction, *Journal of Computer Science*, (7) 4 (2008), pp. 571-577. (in English)

[11] Bases of planning of the programming systems. http://www-csd.univer.kharkov.ua/content/files/cat16/opps.pdf. (in Russian)

[12] 1C:Enterprize 8. Management by an agricultural enterprise, http://solutions.1c.ru/catalog/agr-enterprise. (in Russian)

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ANFIS AND NEURAL NETWORK BASED FACIAL EXPRESSION RECOGNITION USING CURVELET FEATURES

S.P.Khandait [1], R.C.Thool [2], P.D.Khandait [3]

[1] Dept. of Info.Tech., K.D.K.C.E., RSTMNU, Nagpur, India, prapti_khandait@yahoo.co.in
[2] Dept. of Info.Tech., SGGSIET, SRTMU, Nanded, Maharashtra, India
[3] Electronics Engineering Department, KDKCE, RSTMNU, Nagpur, Maharashtra, India

Abstract: *Curvelet transform is a promising tool for multi-resolution analysis on images. This paper explains a new approach for facial expression recognition based on curvelet features extracted using curvelet transform. Curvelet transform is applied on the database images and curvelet coefficients are obtained for selected scale for image analysis. Facial curvelet features are compressed using singular value decomposition (SVD) approach. Back propagation neural network (BPNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) are used as classifiers for classifying expressions into one of the seven categories like angry, disgust, fear, happy, neutral, sad and surprise. Experimentation is carried out on JAFFE database. The experimental results show that the novel approach is a better option for extracting feature values and classifying facial expressions.*

Keywords: *Facial expression recognition, Curvelet transform, SVD, BPNN, ANFIS.*

1. INTRODUCTION

The area of Human Computer Interaction (HCI) plays an important role in resolving the absences of neutral sympathy in interaction between human being and machine (computer). HCI will be much more effective and useful if computer can predict about emotional state of human being and hence mood of a person from supplied images on the basis of facial expressions and will be considered as boon for vision community. An Intelligent Biometrics systems aims at localizing and detecting human faces from supplied images so that further recognition of persons and their facial expression recognition will be easy. For classifying facial expressions into different categories, it is necessary to extract important facial features which contribute in identifying proper expressions. Recognition and classification of human facial expression by computer is an important issue to develop automatic facial expression recognition system in vision community. In recent years, much research has been done on machine recognition of human facial expressions [1-5]. In last few years, use of computers for Facial expression and emotion recognition and its related information use in HCI has gained significant research interest which in turn given rise to a number of automatic methods to recognize facial expressions in images or video [6-10].Multiresolution analysis techniques like wavelets are generally used in computer vision areas including facial expression recognition and face recognition [11][12][13].Many results have been achieved with wavelet transform in signal reconstruction, image analysis, facial expression recognition, etc, but wavelet transform has significant limitations in representing image edges as reported recently in [14][15]. This is due to a fact that wavelet transform can only reflect the point singularity and specialty, but difficult to express characteristics for curves and edges. In image processing, the basis of wavelet applied is isotropous, due to which it is unable to express directions of image edges accurately and sparse representation of images. Therefore, it is difficult to represent the important characteristics of facial contour and curve features using the wavelet. In order to overcome these limitations of wavelet transform, a new multi-scale analysis tool-Curvelet transform, is proposed recently [14]. Curvelet transform has good time-frequency localization characteristics, and can describe gradually to any details for an object and it can characterize the facial local information effectively.Though there are several basic results on image denoising,face recognition and palm print recognition using the curvelet features [16][17][18][19]. To best of our knowledge, there is little research on facial expression recognition with curvelet transform uptil now [20]. This paper focuses on the use of curvelet

transform features compressed using SVD(Singular Value Decomposition) for expression recognition using ANFIS (Adaptive Neuro-Fuzzy System) and Back propagation Neural network (BPNN) classifier and presents comparative analysis of two classifiers for curvelet-SVD features.

2. CURVELET TRANSFORM

The first generation of curvelet transform was proposed by Candes and Donoho [21] in 1999. In order to make the implementations simpler, faster and less redundant, the second generation of curvelet transform is proposed in 2006 [14]. It's an effective analytical method for multiresolution, band pass and directions that are considered as three important characteristics the «optimal» image representation should have from the perspective of biological point of view. Therefore, the curvelet transform has better representation capability than wavelet transform for image edges. Curvelet tiling in the frequency domain (left) and spatial domain (right) is as shown in Fig. 1. The scale and angle segmentations in curvelet transform are shown in Fig.2.There are two different digital implementations for curvelets; one is based on the USFFT (Unequispaced Fast Fourier Transform) and the other is based on wrapping idea [14]. In this paper, we select the one with USFFT for implementation simplicity and the algorithm is described as follows [14].

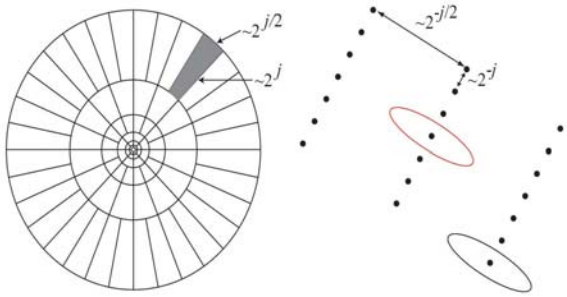


Fig.1 – Curvelet tiling in the frequency domain (left) and spatial domain (right) [14]

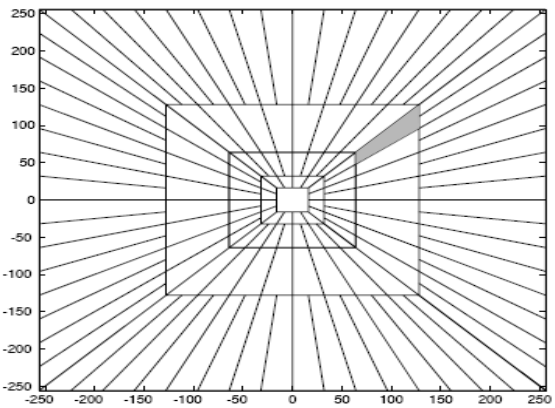


Fig. 2 – The scale and angle segmentation of Curvelet transform [14]

Suppose that we have a two dimensional discrete function

$$f(t_1, t_2) \qquad with \quad 0 \le t_1, t_2 < n.$$

1. Apply the 2D FFT and obtain Fourier samples

$$\hat{f}[n_1, n_2] = \sum_{t_1, t_2 = 0}^{n-1} f(t_1, t_2) e^{-i2\pi(n_1 t_1 + n_2 t_2)/n}, \quad -n/2 \le n_1, n_2 < n/2 \qquad (1)$$

2. For each scale/angle pair (j, l) resample (or interpolate) $\hat{f}[n_1, n_2]$ to obtain sampled values

$$\hat{f}[n_1, n_2 - n_1 \tan\theta_l], \quad for\ (n_1, n_2) \in p_j \qquad (2)$$

where p_j is defined in [14]

3. Multiply the interpolated object $\hat{f}$ with the parabolic window $\tilde{U}_j[n_1, n_2]$ defined in [14] and obtain

$$\hat{f}_{j,l}[n_1, n_2] = \hat{f}[n_1, n_2 - n_1 \tan\theta_l]\, \tilde{U}_j[n_1, n_2] \qquad (3)$$

4. Apply the inverse 2D FFT to each $\hat{f}_{j,l}$, hence collecting the discrete coefficients $c^D(j, l, k)$

3. FACIAL FEATURE EXTRACTION

3.1 SINGULAR VALUE DECOMPOSITION

SVD[22] is an effective algebraic feature extraction method used to compress the image features for reducing dimension. Here image is decomposed into a singular value matrix containing only a few non-zero values.

If matrix $A \in R^{m \times n}$, then there exist two orthogonal matrices:

$$U = [u_1, u_2, \ldots, u_m] \in R^{m \times m}, \quad V = [v_1, v_2, \ldots, v_n] \in R^{n \times n} \qquad (4)$$

Which makes,

$$A = U \sum_{\lambda} V^T \qquad (5)$$

where,

$$\sum_{\lambda} = diag[\sigma_1, \sigma_2, \ldots, \sigma_p], p = min(m, n),$$
$$\sigma_1 \ge \sigma_2 \ge \cdots \sigma_i \ldots \ge \sigma_p \ge 0, \sigma_i (i = 1, 2, \ldots, p)$$

are all non-zero singular values of matrix A. They are the square root of the Eigen values of the AA^H or

$A^H A$ and

$$\sigma_i = \sqrt{\lambda_i} \tag{6}$$

3.2 FEATURE EXTRACTION USING CURVELET AND ITS COMPRESSION USING SVD

SVD is used along with Curvelet [23] for extracting and compressing image facial features successfully. As mentioned in [14][24], wavelets are only suitable for detecting singular points in an image and fail to represent curved discontinuities along edges. On the contrary, the curvelet transform can represent the curved changes. In image processing, the edges in an image represent significant information which can be used for representation and recognition. One significant advantage of curvelet over wavelet is that curvelet includes the detailed information on edges. Curvelet transform is successfully used for face recognition problems but very little contributions arises as for as use of curvelet transform for expression recognition is concerned. Following steps are used to extract facial features-

i) The curvelet transform is performed for each facial expression image using algorithm mentioned in section II.
ii) The coarse, detail and fine coefficients are obtained.
iii) SVD is applied to compress the coefficients.

The reconstructed images with each layer are shown in Fig. 3. From Fig.3, we can find that the detail layers represent the image sketch quite well, and include much richer information than that in the low and high frequency parts. In order to analyze the curvelet coefficients, we need a face image's expression region with resolution $n \times n$ dyadic. The images are then decomposed using the curvelet transform with scale s given as –

$$s = \log_2(n) - 3 \tag{7}$$

The inner layer is a coefficient matrix corresponding to the low frequency of 32x32, and the external layer is coefficient matrix for high frequency of nxn, and the middle parts are detailed layers. For experimentation cropped face image is then resized to size 256×256 pixels. Real valued Curvelet decomposition is achieved in 5 scales. They are low frequency, first detailed layer with 64 directions, second detailed layer with 32 directions, and high frequency respectively. The size of matrix we get after curvelet decomposition is small compared to the original image size. Even then this size is still very high to use it as feature matrix therefore feature compression is achieved using SVD by using same procedure mentioned in section III. A, selecting first 32 diagonal curvelet coefficients.

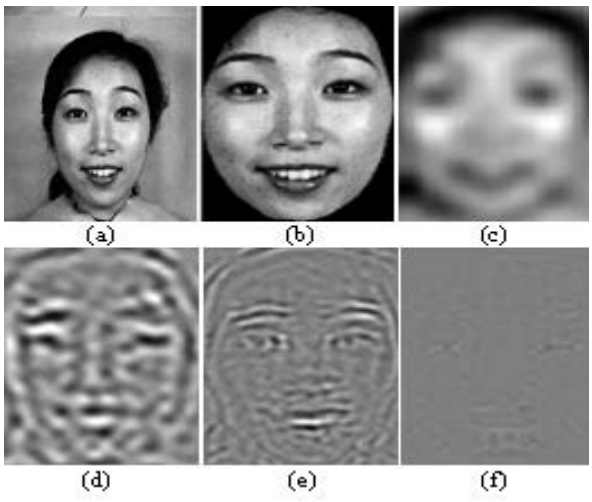


Fig. 3 – (a)-(f) original image, cropped face, low frequency, detail 1, detail 2 and high frequency parts respectively

4. EXPRESSION CLASSIFICATION AND RECOGNITION

4.1 BACK PROPAGATION NEURAL NETWORK FOR EXPRESSION RECOGNITION

The back propagation algorithm is one of the simplest and most general methods for supervised training of multilayer neural networks [25]. Fig. 4 shows the architecture of BPNN.

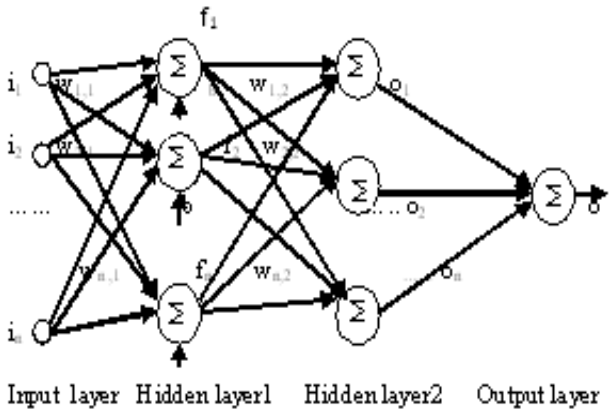


Fig. 4 – Architecture of BPNN

4.2 ANFIS FOR EXPRESSION RECOGNITION

ANFIS proposed by Jang [26] are a class of adaptive networks that are functionally equivalent to fuzzy inference systems. It represent Sugeno

Tsukamoto fuzzy models and uses hybrid learning algorithm. Fig.5 shows architecture of ANFIS model.Thirty-two feature values obtained using curvelet-SVD method are given as an input to the ANFIS model. Initial FIS Structure is generated using grid partitioning method. Here, feature vector data is partitioned into eight groups with four inputs each. Eight ANFIS models are constructed which takes four inputs each. Three Gaussian bell membership functions are associated with each input in order to model the variation of input values (small, medium and large), so the input space is partitioned into fuzzy subspaces. In all 81 rules are generated for an ANFIS model. The premise part of a rule describes a fuzzy subspace, while the consequent part specifies the output with in this fuzzy subspace.

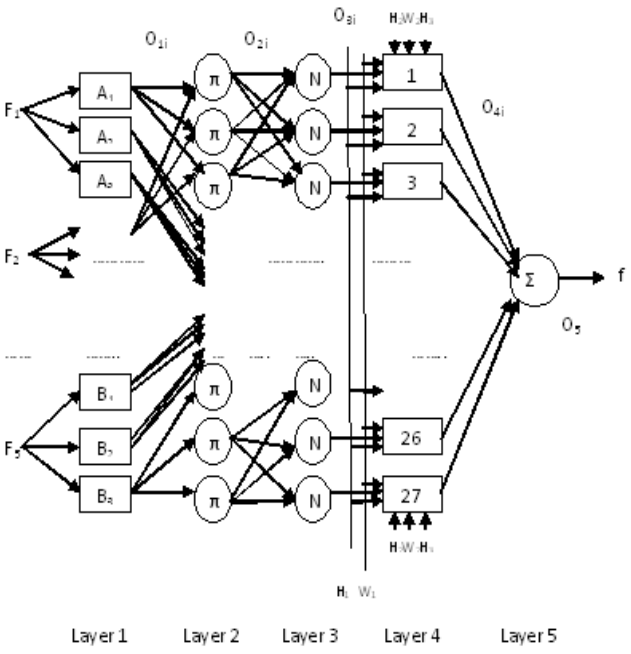


Fig. 5 – ANFIS architecture for 4 inputs, 3 membership functions

Output of layer 5 of each model is given as input to maximum occurrence finder (Fig. 6) which finds the maximum occurrence value of particular expression. For example, output of any five models is angry, one model give neutral and remaining two models give fear expression as output then maximum occurrence finder give angry expression as output due to maximum occurrence of angry expression. Fig. 7 shows membership function before and after training ANFIS model for first four input values of SVD-Curvelet coefficients. Fig. 8 shows performance plot of Curvelet-SVD+ANFIS model. Here network learns gradually and reaches towards the goal.

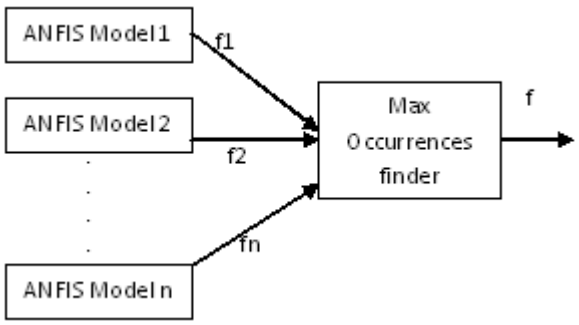


Fig. 6 – ANFIS model o/p predictor for expression recognition

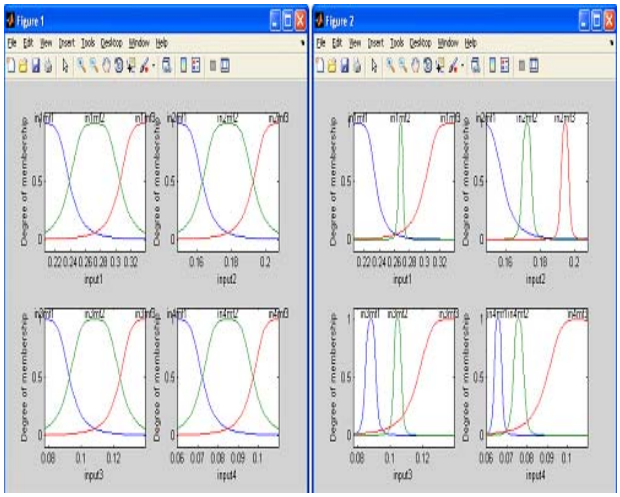


Fig. 7 – Membership function before training and after training ANFIS model

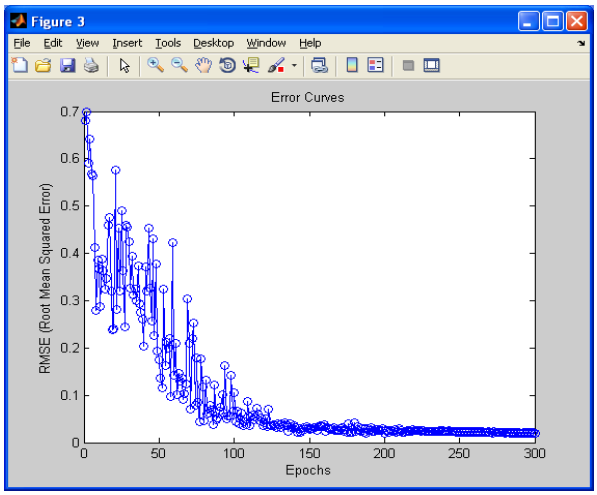


Fig. 8 – Performance plot of Curvelet-SVD+ANFIS model

5. EXPERIMENTAL RESULTS AND ANALYSIS

Experimentation is carried out using JAFFE database images [27],which are preprocessed using the approach mentioned in [28].Wavelet transform is applied on preprocessed images. Thirty-two

feature values obtained using Curvelet-SVD are given to BPNN model. BPNN model with different number of neurons for two hidden layers along with variation in transfer function is tried. Neural Network Model with two hidden layers having 15 and 32 neurons each and an output layer achieved better efficiency. Transfer function used at hidden layer1 is tansig, hidden layer 2 is tansig and at output layer is linear. During training phase, BPNN model is trained for different number of samples (159, 146, 134, and 114) and achieved 100% classification accuracy. For above experimentation selected parameters are given below – BPNN (32-15-32-1), Number of epochs =3000, mse = 0.0001.

During testing phase, trained model is tested for different number of samples (30, 43, 55, and 75) and achieved average recognition efficiency from 93.33% to 80.00% [Table 2, Fig. 11]. Table 1 shows Confusion Matrix for testing dataset with Curvelet-SVD+BPNN based Facial Expression recognition (75 samples). Fig. 9 and 10 shows GUI for recognition of facial expression using Curvelet-SVD+BPNN and ANFIS approach respectively.

Similar to BPNN, ANFIS model is also trained and tested for same number of samples and achieved testing accuracy in the range 90.00% to 80.00% (Table 2, Fig. 11).

Fig.9 – GUI for recognition of facial Expression using Curvelet-SVD+BPNN approach

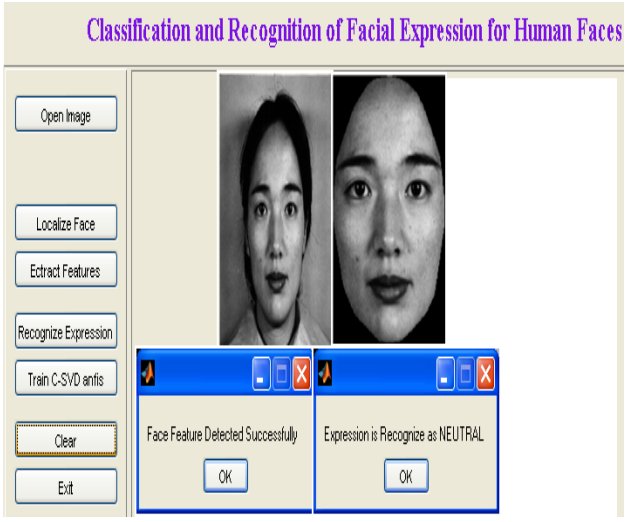


Fig. 10 – GUI for recognition of facial Expression using Curvelet-SVD+ANFIS approach

Table 1. Confusion Matrix for testset with Curvelet-SVD+BPNN based Facial Expression recognition (75 samples)

Expression	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise	Accuracy Rate
Angry	8	0	1	0	0	1	0	80.00
Disgust	0	9	0	0	0	1	0	90.00
Fear	0	1	8	0	0	1	0	80.00
Happy	0	0	0	9	1	0	0	90.00
Neutral	0	0	1	0	13	1	0	86.67
Sad	0	1	0	0	0	9	0	90.00
Surprise	0	0	0	0	0	3	7	70.00
Average Recognition Accuracy								80.00

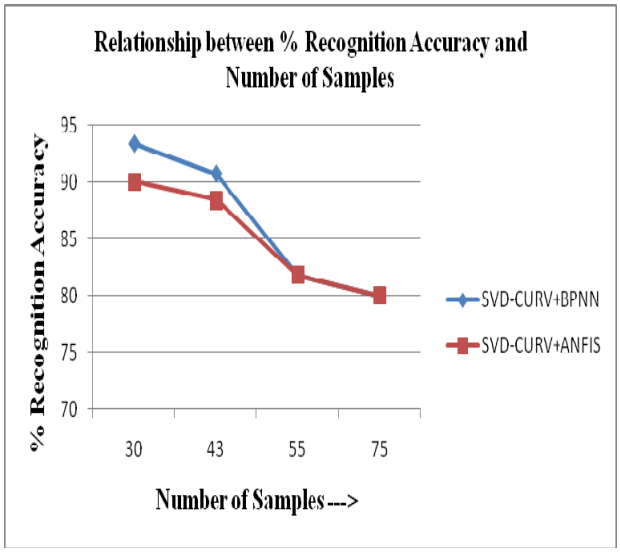


Fig. 11 – Relationship between Number of Samples and % Recognition accuracy for Curvelet-SVD +ANFIS and Curvelet-SVD+BPNN

Table 2. Recognition efficiency for Curvelet-SVD+BPNN and Curvelet-SVD+ANFIS approach

Number of Samples	% Average Recognition Efficiency	
	Curvelet-SVD+ANFIS	Curvelet-SVD+BPNN
30	90.00	93.33
43	88.37	90.69
55	81.81	81.81
75	80.00	80.00

6. CONCLUSION

Appearance Facial features are extracted using Curvelet transform and are compressed using SVD for efficient memory utilization. Extracted feature vector values are given as input to BPNN classifier and ANFIS classifier. Highest % Recognition accuracy (93.33 %) was achieved using curvelet-SVD+BPNN approach for 30 testing samples taken from JAFFE database. Lowest % Recognition accuracy (80.00 %) was achieved for 75 testing samples taken from JAFFE database. From Table 2 and Fig. 11 one can conclude that recognition efficiency of SVD-Curvelet+BPNN is slightly better than SVD-Curvelet+ANFIS. Thus SVD-Curvelet-BPNN outperforms over techniques nentioned in the literature.

7. REFERENCES

[1] Daw-tung lin, Facial expression classification using PCA and hierarchical radial basis function network, *Journal of Information Science and Engineering*, 22 (2006), pp. 1033-1046.

[2] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, Jian-Cheng Huang, A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA, *International Journal of Information Technology*, (11) 11 (2005).

[3] P. S. Aleksic and A. K. Katsaggelos, Automatic facial expression recognition using facial animation parameters and multi-stream HMMs, *6th European Workshop on Image Analysis for Multimedia Interactive Services*, Montreux, Switzerland, 2005.

[4] Limin Ma, David Chelberg and Mehmet Clelenk, Spatio-temporal modeling of facial expressions using Gabor-wavelets and hierarchical hidden Markov models, *in the Proc.of ICIP'2005*, (2005), pp. 57-60.

[5] M. Pantic and Ioannis Patras, Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences, *IEEE transactions on Systems, Man, and Cybernetics. Part B: cybernetics*, (36) 2 (2006).

[6] Ruicong Zhi, Qiuqi Ruan, A comparative study on region-based moments for facial expression recognition, *in Congress on Image and Signal Processing*, 2 (2008), pp. 600-604.

[7] Irene Kotsia, Ioannis Pitas, Facial expression recognition in image sequences using geometric deformation features and support vector machines, *IEEE Transactions on Image Processing*, (16) 1 (2007), pp. 172-187.

[8] Kakumanu P., Nikolaos, G. Bourbakis, A local-global graph approach for facial expression recognition, *ICTAI*, (2006), pp. 685-692.

[9] Aleksic P. S., Aggelos K. Katsaggelos, Automatic facial expression recognition using facial animation parameters and multistream HMMs, *IEEE Transactions on Information Forensics and Security*, (1) 1 (2006), pp. 3-11.

[10] Spiros Ioannou, George Caridakis, Kostas Karpouzis, and Stefanos Kollias, Robust feature detection for facial expression recognition, *EURASIP Journal on Image and Video Processing*, (2007), Article ID 29081.

[11] Xue Weimin, Facial expression recognition based on Gabor filter and SVM, *Chinese Journal of Electronics*, (15) 4 (2006).

[12] Praseeda Lekshmi V., M. Sasikumar, Analysis of facial expression using Gabor and SVM, *International Journal of Recent Trends in Engineering*, (1) 2 (2009).

[13] Fan Chen, Kazunori Kotani, Facial Expression Recognition by SVM-based two-stage classifier on Gabor features, *MVA* 2007 *IAPR Conference on Machine Vision Applications*, (May 16-18, 2007), Tokyo, Japan.

[14] Candes E. J., Demanet L., Donoho D. L., Fast discrete curvelet transforms[R], *California: California Institute of Technology, Applied and Computational Mathematics*, (2006), pp.1-43.

[15] Tanaya Mandal and Q. M. Jonathan Wu, *Face Recognition Using Curvelet Based PCA*, 978-1-4244-2175-6/08/©2008 IEEE.

[16] J. L. Starck, E. J. Candes, D. L. Donoho, The curvelet transform for image denoising, *IEEE Trans. on Image Processing*, (11) 6 (2002), pp. 670-684.

[17] Z. Jiulong, Z. Zhiyu, H. Wei, L.Yanjun and W. Yinghui, Face recognition based on curvefaces, *Proceedings of the Third International Conference on Natural Computation*, Vol. 2, (2007), pp. 627-631.

[18] T. Mandal, A. Majumdar and J. Hu, Face recognition via curvelet based feature extraction, *International Conference on Image Analysis and Recognition*, (2007), pp. 806-817.

[19] A. Majumdar and A. Bhattacharya, Face recognition by multiresolution curvelet transform on bit quantized facial images, *IEEE*

International Conference on Computational Intelligence and Multimedia Applications, (2007), pp. 209-213.

[20] Juxiang Zhou, Yunqiong Wanga, Tianwei Xu, Wanquan Liu, A novel facial expression recognition based on the curvelet features, *IEEE 2010 Fourth Pacific-Rim Symposium on Image and Video Technology*, (2010), pp. 82-87.

[21] Candes E. J., Donoho D. L., Curvelet surprisingly effective non adaptive representation for objects with edges, curve and surface fitting, Saint-Malo, TN: Vanderbilt University Press, 1999, pp. 105-120.

[22] B. Chanda and D. Dutta Majumder, *Digital Image Processing and Analysis*, Prentice-Hall India, New Delhi, 2002.

[23] He Jia, Zhang Xiang Feng, Facial feature extraction and recognition based on curvelet transform and SVD, *In Proc. of IEEE ICACIA-09*, (2009), pp.104-107.

[24] Yan Jingwen, Qu Xiaobo, *Analysis and Application of Super Wavelet*, National Defence Industry Press, 2008.

[25] Duda R.O., Hart P.E., Stork D.G., *Pattern Classification*, Second edition, Wiley student edition, 2006.

[26] J.S.R Jang, ANFIS: adaptive network based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics*, (23), 3 (1993), pp. 665-685.

[27] JAFFE dataset, http://www.kasrl.org/jaffe.html.

[28] S. P. Khandait, R. C. Thool, P. D. Khandait, Automatic facial feature extraction and expression recognition based on neural network, *(IJACSA) International Journal of Advanced Computer Science and Applications*, (2) 1 (2011), pp. 113-118.

***S.P. Khandait**, is a research scholar and an Associate Professor in the department of Information Technology, KDKCE, RSTMNU, Nagpur, Maharashtra, INDIA. She is a Life Member of Indian Society for Technical Education (LM 23789),Associate Member of I.E.(India) (AM 81913/4), Member of I.E.T.E.,India (M 198753). Presently she is pursuing her PhD in Computer Science and Engineering. Her research research interest is Image processing, computer vision and pattern analysis.*

***Dr. R.C. Thool** is Professor in department of Information Technology, SGGSIET, SRTMU, Nanded, Maharashtra, INDIA He did his M.E. and Ph.D. from SRTMU, Nanded. He is member of IEEE. He published more than 20 papers in national and international conferences and journals. His area of interest includes computer vision, robot vision, and image processing and pattern recognition.*

***P. D. Khandait** is an Associate Professor in the department of Electronics Engineering, KDKCE, RSTMNU, Nagpur, Maharashtra, INDIA. He is Life Member of Indian Society for Technical Education (LM 23790), M.I.E. (India) (M-115954), M.I.E.T.E., India (M 198754). His area of research interest includes Signal and Image processing, Biomedical Engineering, soft computing.*

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

МЕТОД ПОБУДОВИ МОДЕЛЕЙ ПОВЕДІНКИ СКЛАДНИХ СИСТЕМ НЕМАРКОВСЬКОГО ТИПУ У ВИГЛЯДІ ГРАФА СТАНІВ І ПЕРЕХОДІВ

Богдан Волочій, Леонід Озірковський, Ігор Кулик

Кафедра теоретичної радіотехніки та радіовимірювання,
Інститут телекомунікацій, радіоелектроніки та електронної техніки
Національний університет «Львівська політехніка»
вул. С.Бандери, 12, Львів, 79013,
e-mail: bvolochiy@ukr.net, lozirkovsky@polynet.lviv.ua, kulyk.iw@gmail.com

Резюме: *Об'єктом розгляду є процес побудови моделей поведінки дискретно-неперервних стохастичних систем немарковського типу за допомогою методу фаз Ерланга. На базі цього методу та з використанням удосконаленої технології моделювання дискретно-неперервних стохастичних систем марковського типу [1] розроблено метод побудови моделей складних систем немарковського типу у вигляді графа станів і переходів. Розроблений метод проілюстровано прикладом аналізу системи масового обслуговування.*

Ключові слова: *математична модель, поведінка системи, метод фаз Ерланга, марковська модель.*

THE METHOD OF BUILDING OF BEHAVIOR MODEL OF NON-MARKOV COMPLEX SYSTEMS AS A GRAPH OF STATES AND TRANSITIONS

Bohdan Y. Volochiy, Leonid D. Ozirkovskyy, Ihor W. Kulyk

Department of Theoretical Radio Engineering and Radiomeasurement,
Institute of Telecommunications, Radioelectronics and Electronic Engineering
National University «Lviv Polytechnic»
12, Stepana Bandery Street, Lviv, 79013, Ukraine
e-mail: bvolochiy@ukr.net, lozirkovsky@polynet.lviv.ua, kulyk.iw@gmail.com

Abstract: *The goal of the paper is the construction of behavior models of discrete-continuous stochastic systems of non-Markov type by the method of Erlang's phases. The improved method of formalized construction of behavior models as a graph of states and transitions is developed. Since this method is a part of technology modeling discrete-continuous stochastic systems, now the process of construction of behavior systems of non-Markov type is automated.*

Keywords: *mathematical model, the behavior of the system, the method of phase of Erlang, Markov models.*

ВСТУП

Для побудови моделей поведінки складних технічних систем у вигляді дискретно-неперервної стохастичної системи (ДНСС) в [1] запропоновано технологію формування графа станів та переходів, ступінь формалізації якої дозволив автоматизувати цей процес. На основі цієї технології розроблено програмний модуль ASNA-1, призначений для визначення показників надійності (ймовірність безвідмовної роботи та середнє значення тривалості безвідмовної роботи) відмовостійких систем [2]. Для об'єкта дослідження з довільною кількістю станів програмний модуль ASNA-1 здійснює безпомилкову побудову графа станів та переходів, що є важливим при розв'язанні проектних задач методом багатоваріантного аналізу. На основі графа, який представляє ДНСС [3] формується математична модель об'єкта дослідження у вигляді системи диференціальних рівнянь Колмогорова-Чепмена. Така модель накладає умову на об'єкт дослідження, яка полягає в тому, що тривалості

всіх процесів в ньому вважаються розподіленими за експоненційним законом розподілу.

Однак, для багатьох об'єктів дослідження така умова знижує ступінь адекватності моделі і вносить помилку в результат. Так в моделях відмовостійких систем з технічним обслуговуванням для тривалостей відновлення і в моделях систем масового обслуговування для тривалостей обслуговування заявок рекомендується використовувати закон розподілу Ерланга [7]. Дискретно-неперервні стохастичні системи, в яких існують процеси, тривалості яких описуються законами розподілу відмінними від експоненційного називають немарковськими. Для побудови моделей таких систем використовують метод фаз Ерланга (ФЕ) [3-6], який дозволяє немарковський дискретно-неперервний процес, тривалості якого розподілені за законом Ерланга апроксимувати марковським процесом. Згідно цього методу необхідно граф станів і переходів трансформувати, шляхом заміни відповідних станів ланцюжками фіктивних станів. Однак, використання цього методу без засобів автоматизації для об'єктів дослідження з великою кількістю станів є трудомісткою задачею.

В статті показано метод побудови моделей поведінки складних систем немарковського типу, в основу якого покладені: метод побудови моделей у вигляді графа, описаний в монографії [1] і метод фаз Ерланга.

1. УДОСКОНАЛЕННЯ ТЕХНОЛОГІЇ ПОБУДОВИ МОДЕЛЕЙ ПОВЕДІНКИ ДИСКРЕТНО-НЕПЕРЕРВНИХ СТОХАСТИЧНИХ СИСТЕМ

В роботах [8, 9, 10] проведено детальний розгляд особливостей використання методу ФЕ та запропоновано методику формалізованої побудови моделей поведінки систем немарковського типу на базі цього методу. Щоб зробити цю методику придатною для використання, в практиці моделювання поведінки складних технічних систем, в технологію моделювання, представлену в монографії [1], внесено ряд доповнень, які стосуються формування вербальної та побудови структурно-автоматної моделей. Внесені доповнення дозволяють здійснювати автоматизовану побудову моделі поведінки системи у вигляді графа станів та переходів, з використанням методу ФЕ, яку забезпечує (реалізує) програмний модуль ASNA-1.

Суть цих доповнень полягає в наступному:

1) При формуванні вербальної моделі необхідно здійснити опис всіх подій, які можуть відбуватися в досліджуваній системі та провести їх класифікацію на базові та супутні. Для кожної базової події необхідно описати всі ситуації, в яких вона може відбуватися.

Для систем немарковського типу, з використанням методу ФЕ, із процесів, які протікають в системі необхідно виділити процеси, тривалість яких не можна представляти експоненційним законом розподілу. Для тривалостей цих процесів визначається математична модель реального закону розподілу ймовірності (густина розподілу f(t), математичне очікування m_c , дисперсія σ^2).

Для використання цих даних при побудові моделей поведінки стохастичної системи за допомогою методу ФЕ, реальні закони розподілу необхідно апроксимувати за допомогою закону розподілу Ерланга або композиції кількох законів розподілу Ерланга. Здійснюється визначення параметрів апроксимуючого закону розподілу або композиції законів розподілу Ерланга (мова йде про порядок закону розподілу або так званий параметр форми n та параметр масштабу α, за допомогою яких апроксимується реальний закон розподілу).

2) Побудова структурно-автоматної моделі (САМ) здійснюється на основі вербальної моделі, яка задає вхідні дані у вигляді переліку базових подій та опису ситуацій (поданих умовами та обставинами), за яких ці події відбуваються.

З обраних базових подій визначаються ті події, які спричиняють процеси, тривалість яких не можна представляти експоненційним законом розподілу.

Для автоматизації процесу формування додаткових ланцюжків фіктивних станів згідно методу ФЕ, необхідно внести деякі зміни в методику формування структурно-автоматної моделі.

Кількість додаткових ланцюжків визначається кількістю законів розподілу Ерланга, які використовуються при апроксимації немарковських процесів в ДНСС, а кількість фіктивних станів в цих додаткових ланцюжках визначається порядком апроксимуючого закону розподілу Ерланга.

Формування вектора станів полягає у виборі компонент, які визначають стан системи (об'єкта дослідження) в кожен момент часу. Кількість компонент в описі поточного стану системи повинна відповідати кількості параметрів, зміну яких визначає поведінка об'єкта дослідження.

При використанні методу ФЕ необхідно призначити додаткові компоненти вектора стану, які визначають перебування ДНСС у фіктивних

станах додаткових ланцюжків. Кількість додаткових компонент вектора стану визначається кількістю процесів, тривалість яких розподілена не експоненційно. Поточне значення додаткової компоненти відображає на якій фазі додаткового процесу перебуває ДНСС. Початкове значення цієї компоненти відповідає кількості фаз додаткового процесу (кількості станів додаткового ланцюжка, які для об'єкта дослідження є фіктивними).

При формуванні множини формальних параметрів задаються значення інтенсивності потоків подій та константи, які визначають структуру об'єкта дослідження.

В побудову дерева правил модифікації компонент вектора стану внесено ряд доповнень. Згідно технології моделювання поданій в [1], у відповідній табличній формі, подається формалізований опис поведінки системи. За допомогою логічних функцій та математичних операцій над компонентами вектора станів проводиться опис умов та обставин, за яких відбуваються базові події.

Для базових подій, які спричиняють появу немарковських процесів в ДНСС необхідно сформувати додаткові умови та обставини, які будуть визначати перебування у фіктивних станах додаткових ланцюжків.

При розгляді базової події, яка спричиняє появу в системі немарковського процесу, необхідно реальний закон розподілу його тривалостей апроксимувати за допомогою закону розподілу Ерланга n-го порядку. Тоді відповідні стани ДНСС необхідно замінити ланцюжком фіктивних станів. Кількість станів цього ланцюжка повинна бути рівною порядку закону розподілу Ерланга, який використовують для апроксимації. Для цього у вектор стану вноситься нова компонента V_k, початкове значення якої визначає кількість станів додаткового ланцюжка n.

Вигляд додаткового ланцюжка з значенням додаткової компоненти V_k для кожного стану зображено на рис. 1.

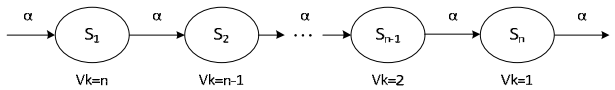


Рис. 1 – Додатковий ланцюжок фіктивних станів дискретно-неперервної стохастичної системи

Для цієї базової події крім умов та обставин, що описують ситуації в яких відбувається подія, необхідно додати дві ознаки, які будуть визначати перебування ДНСС у фіктивних станах $S_1 \ldots S_n$ додаткового ланцюжка і відповідно визначати процедуру його формування в структурі графа станів і переходів.

Перша ознака відображає перебування ДНСС в станах від S_1 до S_{n-1}. Ця ознака записується наступним чином $V_k>=2$, а правило модифікації компоненти вектора стану має вигляд $V_k:=V_k-1$.

Друга ознака відображає завершення перебування ДНСС в останньому стані додаткового ланцюжка. Формалізований запис цієї ознаки є таким $V_k=1$. При формуванні правила модифікації компонент вектора стану необхідно повернути компоненті V_k початкове значення і змінити відповідні компоненти, які відображають перехід в наступний реальний стан об'єкта дослідження.

Розроблена структурно-автоматна модель об'єкта дослідження, з додатковою компонентою вектора стану та відповідними ознаками стану, вводиться в програмний модуль ASNA-1, який здійснює автоматизовану побудову графа станів та переходів об'єкта дослідження та розв'язує систему диференційних рівнянь Колмогорова-Чепмена.

Розв'язання системи рівнянь дозволяє отримати розподіл ймовірностей перебування ДНСС у всіх станах S_k в довільний момент часу t, тобто $P_k(t)$.

Якщо в досліджуваній системі існують немарковські процеси, які розпочинаються не з початкового стану, то при заміні цих станів додатковим ланцюжком не вдається отримати вибраний для апроксимації закон розподілу Ерланга. Ця проблема розглядається в статті [8]. В цій статті показано, що використовуючи метод фаз Ерланга при побудові моделей складних систем необхідно перевіряти точність апроксимації, яку забезпечує апроксимуючий закон розподілу. Для цього на основі розподілу ймовірностей перебування системи в станах, проводиться розрахунок апроксимуючих еквівалентних інтенсивностей, які повинні описувати відповідний процес (відновлення, обслуговування тощо).

Після цього здійснюється формування виразу густини розподілу ймовірності для тривалості апроксимуючого процесу та розраховуються параметри апроксимуючого закону розподілу (математичне очікування m_c, дисперсія σ^2).

Перевірка точності апроксимації здійснюється шляхом порівняння параметрів реального закону розподілу, які визначаються при побудові вербальної моделі та розрахованих параметрів апроксимуючого закону розподілу.

Проводиться оцінка відповідності форми функцій густини розподілу ймовірності реального та апроксимуючого законів розподілу. Для цього оцінюється різниця між

математичними очікуваннями m_c та дисперсіями σ^2 обох законів розподілу.

Якщо математичне очікування апроксимуючого закону розподілу mca більше від значення математичного очікування реального закону розподілу mcт, то значення інтенсивності переходів між додатковими станами необхідно збільшувати, а у протилежному випадку зменшувати.

Після того, як буде досягнута необхідна точність апроксимації реального закону розподілу, отриманий результат розв'язання системи диференціальних рівнянь Колмогорова-Чепмена вважається правильним.

Тепер можна приступити до формування показників ефективності досліджуваної системи. З отриманого розподілу ймовірностей перебування ДНСС у станах формують необхідні показники ефективності об'єкта дослідження. Наприклад, шляхом сумування ймовірностей перебування системи у відповідних станах.

2. ПРИКЛАД ЗАСТОСУВАННЯ УДОСКОНАЛЕНОЇ ТЕХНОЛОГІЇ МОДЕЛЮВАННЯ

Постановка задачі дослідження. Розглядається система обробки інформації, моделлю якої є система масового обслуговування (СМО) з обмеженою чергою та одноканальним, однофазним і ненадійним обслуговуванням рис. 2 [1].

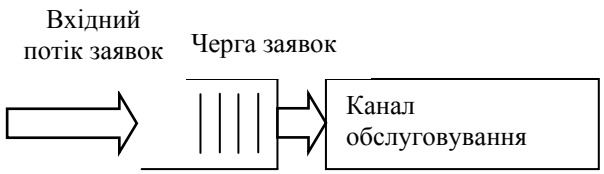


Рис. 2 – Система масового обслуговування з обмеженою чергою та одноканальним, однофазним і ненадійним обслуговуванням

Досліджувана система працює наступним чином: заявка, яка надходить при відсутності черги і незайнятому та працездатному каналі обслуговування поступає до нього на обслуговування. Якщо канал обслуговування зайнятий або несправний і в черзі є вільне місце, заявка яка надходить стає в чергу. Якщо канал обслуговування зайнятий або несправний і в черзі немає вільного місця, заявка яка надходить втрачається.

Часовий інтервал між заявками, для розглянутої СМО, розподілений за законом розподілу Ерланга 3-го порядку з густиною розподілу $f_{er}(t)$:

$$f_{er}(t) = \alpha \cdot e^{-\alpha \cdot t} \cdot \frac{(\alpha \cdot t)^{n-1}}{(n-1)!} \qquad (1)$$

де α – параметр масштабу; n – параметр форми.

Канал обслуговування може виходити з ладу, причому втрата працездатності може статися коли канал є вільним, а може статися коли канал зайнятий обслуговуванням заявки. Заявка, яка обслуговується в момент появи порушення працездатності каналу, повертається в чергу, якщо в ній є вільне місце. Якщо вільного місця немає, вона втрачається. Порушення працездатності каналу обслуговування виявляється засобами контролю і після цього починається ремонт каналу обслуговування, який покладено на ремонтний орган. Кількість ремонтів каналу обслуговування не обмежено, причому ремонт завжди є успішним. Функція контролю стану працездатності каналу обслуговування виконується бездоганно, тобто ймовірність виявлення порушення працездатності рівна одиниці.

Показниками ефективності системи обробки інформації вибрано «ймовірність втрати заявки» та «ймовірність перебування заявки в каналі обслуговування».

Необхідно оцінити різницю між значеннями показників ефективності системи обробки інформації, якщо для проведення розрахунків в якості її моделі використано СМО M|M|1|2 і модель з вищим ступенем адекватності E3|M|1|2.

Розробка структурно-автоматних моделей. Для проведення дослідження здійснена розробка двох САМ. Перша САМ розроблена за умови, що ймовірнісною моделлю для інтервалів часу між вхідними заявками і для тривалостей обслуговування заявок є експоненційний закон розподілу. Друга САМ розроблена за умови, що ймовірнісною моделлю для інтервалів часу між вхідними заявками є закон розподілу Ерланга 3-го порядку, а для тривалостей обслуговування заявок – експоненційний закон розподілу.

Перелік базових подій. Для такої СМО характерними є такі базові події:

Подія 1. «Прихід заявки на обслуговування».

Подія 2. «Завершення обслуговування заявки».

Подія 3. «Втрата працездатності каналу обслуговування».

Подія 4. «Завершення ремонту каналу обслуговування».

Параметри моделі СМО, які відображені в структурно-автоматній моделі:

KF – кількість станів додаткового ланцюжка; для закону розподілу Ерланга 3-го порядку

KF=3;

λ – інтенсивність відмов каналу обслуговування;

μ – інтенсивність відновлення каналу обслуговування;

α – інтенсивність надходження заявки;

β – інтенсивність обробки заявки;

m – максимальна кількість заявок в черзі.

Компоненти вектора стану:

V1 – вказує на поточну кількість заявок в черзі; початкове значення V1=0; максимальне значення компоненти V1=m;

V2 – описує стан каналу обслуговування; 1 – канал справний і вільний; 2 – канал справний і зайнятий обслуговуванням заявки; 0 – канал несправний;

V3 – індикатор втрати заявок;

V4 – додаткова компонента, яка визначає перебування системи у фіктивних станах додаткового ланцюжка; початкове значення V4=KF.

Опис ситуацій, в яких відбуваються базові події:

Подія 1. «Прихід заявки на обслуговування».

Подія «Прихід заявки на обслуговування» закінчує інтервал часу «очікування чергової заявки». Для тривалостей цього інтервалу є справедливим закон розподілу Ерланга. Таким чином стани, в яких відбувається надходження в систему заявок, необхідно замінити додатковим ланцюжком, який складається з трьох фіктивних станів з інтенсивністю переходів між цими станами α.

Згідно технології моделювання в опис ситуацій, в яких відбувається подія 1, необхідно внести ознаки, які визначають процедуру формування додаткових ланцюжків графа станів і переходів.

Подія 1 є незалежною і відбувається без умови при трьох обставинах. Однак при представленні обставин треба врахувати, що реальна подія 1 відбудеться в останньому стані додаткового ланцюжка. А переходи між іншими (попередніми) станами ланцюжка ініціює також подія 1, але її слід вважати фіктивною. Тому у формалізоване представлення кожної обставини необхідно подати ознаку того, в який стан додаткового ланцюжка формується перехід (останній чи ні). Тобто обставини мають бути представленими двома підпунктами. Це важливо тому, що для кожної обставини необхідно задати різні правила модифікації компонент вектора станів.

Обставина 1а. Канал обслуговування є працездатним і вільним (V2=1), а черга заявок є порожньою (V1=0). Ознака перебування не в останньому стані додаткового ланцюжка (V4>=2). Для обставини 1а правило модифікації компонент вектора стану має визначати перехід в наступний стан додаткового ланцюжка (V4:=V4-1).

Обставина 1б. Канал обслуговування є працездатним і вільним (V2=1), а черга заявок є порожньою (V1=0). Ознака перебування в останньому стані додаткового ланцюжка (V4=1). Тепер заявка, що надходить, поступає в канал обслуговування. Тому для обставини 1б правило модифікації компонент вектора стану має повернути компоненті V4 початкове значення (V4:=KF) і змінити стан каналу обслуговування (V2:=2).

Обставина 2а. Канал обслуговування є працездатним і зайнятий обслуговуванням заявки (V2=2) або несправний (V2=0). В черзі є заявки, проте вона не заповнена до кінця і ще залишається місце для заявок, які прибувають (V1<m). Ознака перебування не в останньому стані додаткового ланцюжка (V4>=2). Для обставини 2а правило модифікації компонент вектора стану має визначати тільки перехід в наступний стан додаткового ланцюжка (V4:=V4-1).

Обставина 2б. Канал обслуговування є працездатним і зайнятий обслуговуванням заявки (V2=2) або несправний (V2=0). Ознака перебування в останньому стані додаткового ланцюжка V4=1. В черзі є заявки, проте вона не заповнена до кінця і ще залишається місце для заявок, які прибувають (V1<m). Для обставини 2б правило модифікації компонент вектора стану має повернути компоненті V4 початкове значення (V4:=KF) і змінити поточну кількість заявок в черзі (V1:=V1+1).

Обставина 3а. Канал обслуговування є працездатним і зайнятий обслуговуванням заявки (V2=2) або несправний (V2=0), а черга заявок є заповненою (V1=m). Ознака перебування не в останньому стані додаткового ланцюжка (V4>=2). Для обставини 3а правило модифікації компонент вектора стану має визначати тільки перехід в наступний стан додаткового ланцюжка (V4:=V4-1).

Обставина 3б. Канал обслуговування є працездатним і зайнятий обслуговуванням заявки (V2=2) або несправний (V2=0), а черга заявок є заповненою (V1=m). В такому випадку новоприбула заявка втрачається, а правило модифікації компонент вектора стану для обставини 3б має повернути компоненті V4 початкове значення (V4:=KF) і змінити компоненту, яка сигналізує про втрату заявки (V3:=1).

Інтенсивність події 1 визначає параметр α.

Подія 2. «Завершення обслуговування

заявки». Подія 2 відбувається за умови, що в каналі обслуговування є заявка (V2=2) і процес в додатковому ланцюжку є завершеним (V4=KF) та за наступних двох обставин:

Обставина 1. Черга заявок є порожньою (V1=0). В такому випадку обслужена заявка звільнює канал обслуговування. Правило модифікації компонент вектора станів записується наступним чином (V2:=1).

Обставина 2. Черга заявок не є порожньою (V1>0). По завершенні обслуговування заявка покидає канал обслуговування, а в нього відразу надходить наступна заявка, яка знаходилася в черзі. Правило модифікації компонент вектора стану записується наступним чином (V1:=V1-1).

Інтенсивність події 2 визначає параметр β.

Подія 3. «Втрата працездатності каналу обслуговування». Подія 3 може відбуватися за умови, що канал обслуговування є працездатним ((V2=1) або (V2=2)) і додатковий процес, який моделює появу події 1, завершено (V4=KF). Для події 3 існує три обставини:

Обставина 1. Канал обслуговування є справним і вільним (V2=1) та черга заявок відсутня (V1=0). Для обставини 1 правило модифікації компонент вектора стану має визначати перехід каналу обслуговування в непрацездатний стан (V2:=0).

Обставина 2. Канал обслуговування зайнятий обслуговуванням заявки (V2=2), а черга є порожньою (V1=0) або не є заповненою до кінця (V1<m). При втраті каналом працездатності заявка, яка в цей час обслуговувалась повертається в чергу, а канал переходить у непрацездатний стан. Тому правило модифікації компонент вектора стану має вигляд (V2:=0); (V1:=V1+1).

Обставина 3. Канал обслуговування зайнятий обслуговуванням заявки (V2=2), а черга є повністю заповненою (V1=m). В такому випадку при втраті каналом працездатності заявка, яка в цей час обслуговувалась втрачається, а канал переходить в непрацездатний стан. Правило модифікації компонент вектора станів має вигляд (V2:=0); (V3:=1).

Інтенсивність події 3 визначає параметр λ.

Подія 4. «Завершення ремонту каналу обслуговування». Ця подія відбувається при умові, що канал обслуговування є непрацездатним (V2=0) і додатковий процес, який моделює появу події 1, завершено (V4=KF). Для події 4 існує дві обставини:

Обставина 1. Заявки в черзі відсутні (V1=0). Для обставини 1 правило модифікації компонент вектора стану має визначати (задавати) перехід каналу обслуговування в працездатний стан (V2:=1).

Обставина 2. В черзі є заявки (V1>0). В цьому випадку канал обслуговування відразу отримує заявку, яка знаходилась в черзі. Для обставини 2 правило модифікації компонент вектора стану має визначати (задавати) перехід каналу обслуговування в стан зайнятий обслуговуванням заявки (V2:=2), а також зменшення черги заявок (V1:=V1-1).

Інтенсивність події 4 визначає параметр μ.

Побудова графа станів та переходів. За допомогою програмного модуля ASNA-1, на основі розроблених САМ, отримано 2 моделі СМО у вигляді графа станів та переходів, зображених на рисунках 3 і 4.

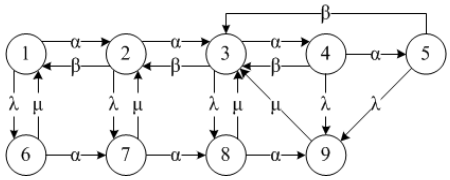


Рис. 3 – Модель СМО типу M|M|1|2 у вигляді графа станів та переходів, як ДНСС марковського типу

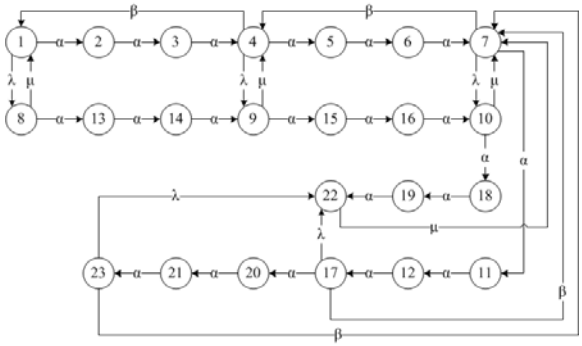


Рис. 4 – Модель СМО типу E3|M|1|2 у вигляді еквівалентного графа станів та переходів, як ДНСС марковського типу з використанням методу ФЕ

Формування показників ефективності досліджуваної системи. З використанням значень ймовірностей перебування в станах, отриманих в результаті розв'язання системи диференціальних рівнянь Колмогорова-Чепмена, здійснюється обрахунок необхідних показників ефективності, а саме «ймовірність втрати заявки» та «ймовірність перебування заявки в каналі обслуговування». В даному випадку показники ефективності формуються шляхом підсумовування ймовірностей перебування системи в певних станах.

Ймовірність втрати заявки розраховується, як сума ймовірностей перебування ДНСС у станах, де зафіксована втрата заявки. Для ДНСС марковського типу, граф станів якої зображено на рис. 3, це стани 5 і 6; для ДНСС немарковського типу, граф станів якої зображено

на рис. 4, це стани 22 і 23. Тому формули для розрахунку ймовірності втрати заявки мають такий вигляд:

$$P_{втр}(t) = P_5(t) + P_9(t) \quad (2)$$

$$P_{втр}(t) = P_{22}(t) + P_{23}(t) \quad (3)$$

Значення ймовірностей втрати заявки для СМО M|M|1|2 становить $P_{втр}=0,125$ і для E3|M|1|2 становить $P_{втр}=0,255$.

Ймовірність перебування заявки в каналі обслуговування розраховується, як сума ймовірностей перебування ДНСС у працездатних станах, в яких канал обслуговування зайнятий обслуговуванням заявки. Для ДНСС марковського типу рис. 3 це стани 2-5; для ДНСС немарковського типу це стани 4-7, 11, 12, 17, 20, 21, 23. Тому формули для розрахунку ймовірності перебування заявки в каналі обслуговування мають такий вигляд:

$$P_{обс}(t) = P_2(t) + P_3(t) + P_4(t) + P_5(t) \quad (4)$$

$$P_{обс}(t) = \sum_{i=4}^{7} P_i(t) + P_{11}(t) + P_{12}(t) + P_{17}(t) + P_{20}(t) + P_{21}(t) + P_{23}(t) \quad (5)$$

Значення ймовірності перебування заявки в каналі обслуговування для СМО M|M|1|2 становить $P_{обс}=0,750$ і для E3|M|1|2 становить $P_{обс}=0,962$.

Отримана різниця між значеннями показників ефективності підтверджує доцільність ускладнення моделі об'єкта дослідження.

3. ВИСНОВКИ

Удосконалено технологію моделювання поведінки складних систем, математичне представлення яких відповідає ДНСС немарковського типу. Це удосконалення дозволяє здійснювати автоматизовану побудову графа станів та переходів для ДНСС немарковського типу з використанням методу фаз Ерланга за допомогою програмного модуля ASNA-1. Удосконалена технологія підвищує ступінь адекватності моделей складних систем за рахунок врахування реальних законів розподілу для тривалостей процедур, які з необхідною точністю можна апроксимувати за допомогою закону розподілу Ерланга.

Здійснено порівняння показників ефективності системи обробки інформації представленої марковською та немарковською ДНСС. Різниця між результатами підтверджує актуальність врахування реального закону розподілу при моделюванні поведінки складної системи.

4. СПИСОК ЛІТЕРАТУРИ

[1] Volochiy B. Y., *Technology Modeling Algorithms Behavior of Information Systems*, Lviv: Publishing National University «Lviv Polytechnic», 2004, 220 p. (in Ukrainian)

[2] B. A. Mandzij, B. Y. Volochiy, L. D. Ozirkovskyy, M. M. Zmysnyy, I. W. Kulyk, Definition of options of strategy for disaster recovery fault-tolerant systems based on plurality of structure, *Bulletin «Lviv Polytechnic». Radio and telecommunications*, 705 (2011), pp. 216-224. (in Ukrainian)

[3] Kochegarov V. A, Frolov H. A., *Designing of System Information Distribution: Markov and non-Markov models*, Moscow, Radio and Communication, 1991, 216 p. (in Russian)

[4] Kenig D., Shtoyan D., *Methods of Queuing Theory*, Moscow, Radio and Communication, 1981, 128 p. (in Russian)

[5] Raynshke K., Ushakov I. A., *Evaluation of Reliability of Systems Using Graph*, Moscow, Radio and Communication, 1988, 208 p. (in Russian)

[6] Cox D., Smith W., *The Theory of Queues*, Translated from English by V. V. Rykova and J. B. Rozhdestvenskyy, edited by A. D. Solovyev, Springer-Verlag, 1966, 221 p. (in Russian)

[7] Bayhelt F., Franken P., *The Reliability and Maintenance. Mathematical approach: first with it*, Moscow, Radio and Communication, 1988, 392 p. (in Russian)

[8] Volochiy B. Y., Ozirkovskyy L. D., Kulyk I. W., Formalization of designing of models of discrete-continuous stochastic systems by Erlang phase method, *Information extraction and processing. National Academy of Sciences of Ukraine. Interbranch collection of scientific papers*, (36) 112 (2012), Lviv. (in Ukrainian)

[9] Mandzij B. A., Volochiy B. Y., Ozirkovskyy L. D., Kulyk I. W., *Automating of building behavior models of non-Markov systems, with using the method of Erlang phase*, *International Symposium «Reliability and Quality-2012»*, Penza, Russia, 2012. (in Russian)

[10] Volochiy B., Osirkovskyy L., Kulyk I. Automation of building of behavior models of the non-Markov discrete-continuous stochastic systems by the method of Erlang phases, *XIth International Conference «Modern problems of radio engeneereng, Telecommunications, and computer science»*, Lviv-Slavske, Ukraine, (February

2012). (in Ukrainian)

***Волочій Богдан Юрійович**, професор, доктор технічних наук, професор кафедри теоретичної радіотехніки і радіовимірювань Національного Університету «Львівська політехніка».*

Автор понад 150 публікацій, зокрема 2 монографій, 2 навчальних посібників, 4 винаходів. Стаж педагогічної роботи у вищій школі – понад 35 років.

Наукові інтереси: теорія і практика системотехнічного проектування радіоелектронних інформаційних систем.

***Озірковський Леонід Деонісійович**, доцент, кандидат технічних наук, декан Інституту телекомунікацій, радіоелектроніки та електронної техніки, доцент кафедри теоретичної радіотехніки та радіовимірювань.*

Стаж педагогічної роботи у вищій школі – понад 10 років.

Опубліковано понад 50 статей та 25 науково-методичних розробок, в тому числі 2 навчальних посібники.

Напрям наукових досліджень – розроблення методів та засобів моделювання функціональної та надійнісної поведінки програмно-апаратних інформаційних систем.

***Кулик Ігор Володимирович**, аспірант кафедри теоретичної радіотехніки і радіовимірювань Національного Університету «Львівська політехніка». Автор 3 статей.*

Напрям наукових досліджень – розробка математичних моделей технічного обслуговування об'єктів інформаційної мережі зв'язку.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

THE METHOD OF BUILDING OF BEHAVIOR MODEL OF NON-MARKOV COMPLEX SYSTEMS AS A GRAPH OF STATES AND TRANSITIONS

Bohdan Y. Volochiy, Leonid D. Ozirkovskyy, Ihor W. Kulyk

Department of Theoretical Radio Engineering and Radiomeasurement,
Institute of Telecommunications, Radioelectronics and Electronic Engineering
National University «Lviv Polytechnic»
12, Stepana Bandery Street, Lviv, 79013, Ukraine
e-mail: bvolochiy@ukr.net, lozirkovsky@polynet.lviv.ua, kulyk.iw@gmail.com

Abstract: *The goal of the paper is the construction of behavior models of discrete-continuous stochastic systems of non-Markov type by the method of Erlang's phases. The improved method of formalized construction of behavior models as a graph of states and transitions is developed. Since this method is a part of technology modeling discrete-continuous stochastic systems, now the process of construction of behavior systems of non-Markov type is automated.*

Keywords: *mathematical model, the behavior of the system, the method of phase of Erlang, Markov models.*

1. INTRODUCTION

To construct the behavior models of complex technical systems, such as a discrete-continuous stochastic systems (DCSS), in [1] proposed a method of forming the graph of states and transitions, the degree of formalization whose allowed to automate this process. This method is incorporated into the technology of modeling the behavior of complex systems [1], based on which developed a software ASNA-1. The software ASNA-1, intended for designing fault-tolerant systems, provides the error-free building of the graph of states and transitions and calculates indices of reliability (reliability as a function of time R(t) and MTBF) fault-tolerant systems [2]. Based on the graph, the mathematical model of the research object is formed as a system of differential equations of the Kolmogorov-Chapman. This model imposes a condition on the object of research, which consists in that the duration of all processes in it are distributed according to exponential distribution law.

But, for many research objects, that condition reduces a degree of adequacy models and made an error in the result. For example, in models of fault tolerant systems with maintenance for the duration of recovery and in models of queuing systems for the duration of maintenance applications is recommended Erlang distribution law [7]. To build such models use the method of phase of Erlang (PE) [3-6], which allows non-markov discrete-continuous process, the duration whose are distributed by law Erlang approximate by Markov processes. According to this method a graph of states and transitions need transform by replacing the corresponding states by chain of fictitious states. But using this method for objects of research with a large number of states without automation tools is time-consuming task.

2. THE IMPROVEMENT OF THE TECHNOLOGY OF BUILDING THE BEHAVIOR MODELS OF DISCRETE-CONTINUOUS STOCHASTIC SYSTEMS

In [8, 9, 10], has been performed the detailed consideration of features of using the method PE and proposed the method of formalized constructing of behavior models of systems non-markov type. This methodology allowed improved the method of constructing of behavior models as a graph of states and transitions. Such as the method is a component the technology modeling [1], the process of building of behavior model systems non-markov type is automated.

The essence of this improvement is follow:

1) In forming verbal model system for the duration of non-markov processes determined the mathematical model of the real law of probability distribution (density distribution f(t), the mathematical expectation mc, variance σ2) and determined the parameters approximating the distribution or composition distribution laws Erlang.

2) In the method of construction of structure-

automatic model of the research object be included the following additions.

In forming the state vector need designate additional components of the vector state that define DCSS stay in fictitious states additional chains. The number of additional components of the vector state defined by the number of processes whose duration is non-exponentially distributed.

When building a event tree the basic events that cause the appearance non-markov processes DCSS, except the conditions that describe situations in which the event occurs, must be create two additional indicators, that will determine the staying system in n fictitious states additional chains in Fig. 1.

The first indicator reflects staying DNSS in fictitious states from S_1 to S_{n-1}. This indicator is written as follows $V_k> = 2$, and a rule modification of component state vector has the form $V_k: = V_{k-1}$.

The second indicator reflects the completion staying DCSS in the last position S_n of additional chain. Formalized recording of this sign is as $V_k = 1$. In forming the rules of modification component of the vector states need to set the initial value component V_k and change the appropriate components that reflect the transition to the next real state of the object of study.

Using the method of PE in most cases can not accurately reproduce the selected order Erlang distribution law. Therefore, after solving the differential equations necessary to verify the accuracy of approximation of the real distribution law. Checking the accuracy of the approximation, made by comparing the actual parameters of the real distribution law to be determined by the construction the verbal models and calculated parameters of the approximating distribution law.

3. CONCLUSIONS

The technology of modeling the behavior of complex systems, mathematical representation which corresponds DCSS non-markov type is improved. This improvement allows exercise the automated build of the graph of states and transitions for DCSS non-markov type using the method of phase of Erlang using the software ASNA-1. The improved technology increase the adequacy of models of complex systems by taking into account the real distribution laws for the duration of procedures with sufficient accuracy can be approximated by the distribution Erlang.

4. REFERENCES

[1] Volochiy B. Y., *Technology Modeling Algorithms Behavior of Information Systems*, Lviv: Publishing National University «Lviv Polytechnic», 2004, 220 p. (in Ukrainian)

[2] B. A. Mandzij, B. Y. Volochiy, L. D. Ozirkovskyy, M. M. Zmysnyy, I. W. Kulyk, Definition of options of strategy for disaster recovery fault-tolerant systems based on plurality of structure, *Bulletin «Lviv Polytechnic». Radio and telecommunications*, 705 (2011), pp. 216-224. (in Ukrainian)

[3] Kochegarov V. A, Frolov H. A., *Designing of System Information Distribution: Markov and non-Markov models*, Moscow, Radio and Communication, 1991, 216 p. (in Russian)

[4] Kenig D., Shtoyan D., *Methods of Queuing Theory*, Moscow, Radio and Communication, 1981, 128 p. (in Russian)

[5] Raynshke K., Ushakov I. A., *Evaluation of Reliability of Systems Using Graph*, Moscow, Radio and Communication, 1988, 208 p. (in Russian)

[6] Cox D., Smith W., *The Theory of Queues*, Translated from English by V. V. Rykova and J. B. Rozhdestvenskyy, edited by A. D. Solovyev, Springer-Verlag, 1966, 221 p. (in Russian)

[7] Bayhelt F., Franken P., *The Reliability and Maintenance. Mathematical approach: first with it*, Moscow, Radio and Communication, 1988, 392 p. (in Russian)

[8] Volochiy B. Y., Ozirkovskyy L. D., Kulyk I. W., Formalization of designing of models of discrete-continuous stochastic systems by Erlang phase method, *Information extraction and processing. National Academy of Sciences of Ukraine. Interbranch collection of scientific papers*, (36) 112 (2012), Lviv. (in Ukrainian)

[9] Mandzij B. A., Volochiy B. Y., Ozirkovskyy L. D., Kulyk I. W., *Automating of building behavior models of non-Markov systems, with using the method of Erlang phase*, *International Symposium «Reliability and Quality-2012»*, Penza, Russia, 2012. (in Russian)

[10] Volochiy B., Osirkovskyy L., Kulyk I. Automation of building of behavior models of the non-Markov discrete-continuous stochastic systems by the method of Erlang phases, *XIth International Conference «Modern problems of radio engeneereng, Telecommunications, and computer science»*, Lviv-Slavske, Ukraine, (February 2012). (in Ukrainian)

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

РОЗРОБКА СИСТЕМИ КЕРУВАННЯ СЕМАНТИЧНИМИ МЕРЕЖАМИ

Юрій Семчишин, Іван Кульпа, Ігор Колосовський, Олександр Гречніков, Петро Гайда

Національний університет «Львівська політехніка»
вул. Степана Бандери, 12, м. Львів, 79013, Україна
e-mail: 7th@ukr.net, ivan.kulpa@gmail.com, igor.kolosovski@gmail.com, sansofter@gmail.com, fezundo@gmail.com

Резюме: *Робота присвячена семантичним мережам – способу представлення знань у виді набору вузлів-сутностей об'єднаних за допомогою ребер-зв'язків. Здійснено огляд історії розвитку та сучасного стану семантичних мереж, обгрунтовано актуальність проблеми. Описано процес проектування та розробки системи керування семантичними мережами: проектування структур даних, розробку алгоритмів семантичного аналізу текстів українською мовою, вибір алгоритмів візуалізації графових структур даних, реалізацію користувацького інтерфейсу у вигляді Web-додатку. Розроблена система є повністю функціональним засобом побудови та дослідження семантичних мереж, що підтвердив свою ефективність та послужить базовою платформою для подальших досліджень.*

Ключові слова: *система керування семантичними мережами, семантична мережа, семантичний аналіз, візуалізація графів.*

DEVELOPING SEMANTIC NETWORK MANAGEMENT SYSTEM

Yuriy Semchyshyn, Ivan Kulpa, Igor Kolosovskyi, Oleksandr Hrechnikov, Petro Hayda

Lviv National Technical University
12 Stepana Bandery Str., Lviv, 79013, Ukraine
e-mail: 7th@ukr.net, ivan.kulpa@gmail.com, igor.kolosovski@gmail.com, sansofter@gmail.com, fezundo@gmail.com

Abstract: *The paper considers semantic networks. It is a way of representing knowledge as a set of nodes-concepts connected by edges-relations. The history and current state-of-the-art of semantic networks is analyzed. The experience of designing and developing semantic network management system is described in four sections covering the following topics: data structures design, semantic analysis algorithm implementation, graph visualization methods selection and Web-based user interface development. The developed system is fully-operational tool for building and studying semantic networks. The system will be used for the further research.*

Keywords: *Semantic Network Management System, Semantic Network, Semantic Analysis, Graph Visualization.*

ВСТУП

Семантична мережа (Semantic Network) – це спосіб представлення знань у виді набору вузлів, які представляють собою деякі сутності та є об'єднаними за допомогою зв'язків, що описують відношення між ними.

Ідея систематизації засобами семантичних відношень зародилась ще у далекому минулому. Наприклад, біологічна класифікація Карла Лінея (Carl Linné) 1735 року, яка описує принципи поділу живих істот на класи та визначає відношення між цими класами [1].

Одним із новаторів свого часу був Фрідріх Людвіг Готлоб Фреге (Friedrich Ludwig Gottlob Frege) – німецький математик і філософ, який представив логічні формули у вигляді дерев. Фреге відзначив різницю між сенсом і значенням поняття, що задається деяким іменем. Під

значенням він розумів предметну область, яка має відношення до заданого імені, а під сенсом – деякий аспект завдяки якому відбувається перегляд даної предметної області. Пари імен, що мають однакове значення але різний сенс отримали назву «семантичних трикутників», або «трикутників Фреге» [2].

Прототипом семантичних мереж можна назвати екзистенціальні графи Чарльза Сандерса Пірса (Charles Sanders Peirce), які він запропонував в 1909 році [3]. Ці графи були вже більш схожі на сучасні семантичні мережі, в яких за допомогою діаграм зображувалися логічні твердження. Пірс вважав, що таке відображення можна назвати «логікою майбутнього».

В розвитку семантичних мереж важливу роль відіграли роботи відомого німецького психолога Отто Зельца (Otto Selz), який вивчав процеси мислення. Для зображення асоціацій і понять він використовував вузли і дуги, тобто графи. Роботи Зельца мали великий вплив на інших вчених, які використовували його ідеї для моделювання людського розуму.

Значного поштовху розвитку семантичних мереж надала робота одного з видатних лінгвістів XX століття француза Люсєна Теньєра (Lucien Tesniure) під назвою «Основи структурного синтаксису» [4]. В цій книзі йдеться про синтаксичні залежності, зв'язки та їх відображення у вигляді дерева залежностей. Саме Теньєр вперше описав поділ синтаксису на статичний і динамічний а також ввів поняття валентності, що визначає способи поєднання слів з іншими елементами.

Першу діючу семантичну мережу розробив у 1956 році Річард Річенс (Richard Richens) при роботі над проектом машинного перекладу Кембриджського дослідницького мовного центру (Cambridge Language Research Unit). Система Річенса використовувала семантичну мережу в якості проміжної форми представлення при здійсненні машинного перекладу.

Наступну версію системи розробила англійська дослідниця Маргарет Мастерман (Margaret Masterman) в 1961 році; система включала в себе близько 100 концептуальних примітивів, за допомогою яких було описано близько 15000 понять. Система Мастерман призначалася для напівавтоматичного перекладу окремих абзаців урядових документів з англійської мови на французьку та давала можливість здійснювати переклад особам, які не знають французької мови [5]. Дослідження в цьому напрямку продовжував Моріс Вінсент Вілкс (Maurice Vincent Wilkes).

Цікавою є робота Роса Квіліана (Ross Quillian) 1968 року, в якій він описує власну семантичну модель. Основним припущенням Квіліана було те, що значення кожного слова можна представити у вигляді великої кількості асоціацій, кожну з яких, в свою чергу, також можна деталізувати. В результаті утвориться величезна павутина вузлів і відношень між ними [6].

Аж до кінця 1970-х років семантичні мережі не мали великої популярності. Проте на початку 1980-х вони набувають значної поширеності, зокрема в системах штучного інтелекту для пошуку відповідей на поставлені запитання та дослідження процесів навчання. На жаль, апаратні засоби того часу не давали можливості створення масштабних систем на базі семантичних мереж.

А вже на початку 1990-х років на основі теорії семантичних мереж виникає концепція семантичної павутини (Semantic Web), автором якої вважається Тім Бернерс-Лі (Tim Berners-Lee) [7]. Ця концепція має на меті здійснення семантичної розмітки інформаційниї ресурсів мережі Internet. Роботу над стандартизацією семантичної павутини було доручено організації «RDF Interest Group», назва якої 2004 року змінилась на «Semantic Web Interest Group». Першим проектом, заснованим на концепції семантичної павутини був проект під назвою «Dublin Core», в межах якого було розроблено кросс-платформенні стандарти метаданих, які можуть використовуватись для вирішення широкого спектру задач [8].

Комп'ютерне програмне забезпечення, що забезпечує користувачам можливість створення, збереження, оновлення, пошуку інформації та контролю доступу в семантичних мережах можна, за аналогією до загальноприйнятого терміну «система керування базами даних», називати системами керування семантичними мережами (СКСМ).

Сучасні системи керування семантичними мережами покликані вирішити проблемні задачі штучного інтелекту: здійснення семантичного аналізу тексту, структуризація специфічної інформації, створення засобів інтелектуального машинного перекладу тощо; серед найбільш вдалих проектів з реалізації систем керування семантичними мережами варто згадати такі:

- WordNet (http://wordnet.princeton.edu/) – лексична база даних, що з 1985 року розробляється у Прінстонському університеті (Princeton University); станом на 2012 рік база даних вміщує 155287 слів згрупованих у 117659 синсетів (synsets, груп синонімів); основною проблемою WordNet є неможливість візуалізації отриманих результатів;
- LexiPedia (http://lexipedia.com/) – шестимовний (англійська, іспанська,

німецька, французька, італійська) тлумачний словник; містить відомості не лише про синоніми та антоніми, але й про фазиніми (fuzzynyms, семантично пов'язані слова); здійснює візуальне представлення зв'язків у вигляді семантичної мережі за допомогою інтерактивного Flash-додатку; до недоліків LexiPedia можна віднести її закритість та відсутність будь-якої технічної документації;

- MLSN (MultiLingual Semantic Network, http://dcook.org/mlsn/) – п'ятимовне (англійська, японська, німецька, китайська, арабська) відгалуження проекту WordNet, що розробляється з 2005 року; володіє базою даних більшого обсягу, ніж батьківський проект; серед недоліків системи варто відзначити можливість роботи виключно з простими іменниками; станом на першу половину 2012 року нестабільність системи унеможливлює візуалізацію результатів;
- TextAnalyst (http://megaputer.com/textanalyst.php) – пропрієтарна система аналізу тексту з функціями формування семантичної мережі та тематичного древа, а також автоматичного реферування, кластеризації та індексації тексту; реалізовано підтримку зовнішніх словників; закритість системи робить неможливими її дослідження та подальший розвиток;
- UMLS (Unified Medical Language System, http://www.nlm.nih.gov/research/umls/) – медична система, що з 1986 року розробляється Національною медичною бібліотекою США (US National Library of Medicine); система призначена для структуризації біомедичної інформації; очевидно, що система не може мати жодних інших застосувань, окрім як в сфері охорони здоров'я.

Як бачимо, жоден з проектів не може претендувати на роль універсальної системи керування семантичними мережами загального призначення; проте потреба в такій системі є на часі, а в майбутньому актуальність досліджень в галузі семантичних мереж лише підвищуватиметься, особливо з огляду на стрімку популяризацію концепції семантичної павутини.

1. ОРГАНІЗАЦІЯ СЕМАНТИЧНОЇ МЕРЕЖІ

Розробку системи керування семантичними мережами було здійснено на мові програмування C# для програмної платформи Microsoft .NET 4.0 засобами середовища розробки Microsoft Visual Studio 2010. Таку комбінацію засобів розробки було обрано завдяки їх гнучкості та потужності, зокрема з огляду на можливість використання запитів, інтегрованих у мову (Language INtegrated Queries, LINQ), автоматичного вивільнення виділеної пам'яті (Garbage Collector, GC) та вбудованих засобів для роботи з розширюваною мовою розмітки (eXtensible Markup Language, XML). Окрім того, вибраний набір засобів розробки робить можливою легку інтеграцію розробленої системи до складу наступних проектів.

В ході вибору структур даних, призначених для представлення семантичної мережі в пам'яті комп'ютера було розглянуто такі варіанти:

- за допомогою матриці суміжності;
- за допомогою списків інцидентності;
- за допомогою деревовидних структур.

Використання матриці суміжності могло б бути доцільним лише у випадку сильнозв'язних графів невеликої розмірності, у випадку ж слабозв'язних графів великої розмірності, які зазвичай виникають при роботі з семантичними мережами накладні витрати пам'яті стали б неприпустимо високими. Використання списків інцидентності дещо сповільнює навігацію по вершинам графа, проте водночас спрощує внутрішню реалізацію системи керування семантичними мережами та забезпечує вищий рівень її масштабованості (scalability). Використання деревовидних структур забезпечило б дещо вищу швидкість доступу до даних, проте неминуче спричинило б зниження масштабованості програмної системи а також виникнення деяких ускладнень в процесі серіалізації графа в сховище даних.

Відтак після аналізу всіх згаданих способів було вирішено використовувати представлення семантичної мережі в пам'яті комп'ютера за допомогою списків інцидентності.

При проектуванні сховища даних для системи керування семантичними мережами було вирішено відмовитись від використання централізованої бази даних, яка б надмірно ускладнила архітектуру програмної системи та викликала б певні труднощі при її розробці, впровадженні та підтримці. З огляду на структурований характер опрацьовуваної інформації було вирішено реалізувати сховище даних у вигляді локальних файлів у текстовому форматі XML, призначеному для зберігання структурованих даних.

Структуру тестового файлу даних в форматі XML представлено на Рис. 1.

```
<?xml version="1.0" ?>
<!-- Developed by Jewana Team -->
<!-- There are all your analyzed relationships -->
- <SemanticNetwork>
  + <Relationship1>
  - <Relationship2>
      <Node1>Ваня</Node1>
      <Relation>любить</Relation>
      <Node2>Машу</Node2>
    </Relationship2>
  - <Relationship3>
      <Node1>Я</Node1>
      <Relation>йду</Relation>
      <Node2>додому</Node2>
    </Relationship3>
  - <Relationship4>
      <Node1>Колобок</Node1>
      <Relation>втік</Relation>
      <Node2>баби</Node2>
    </Relationship4>
```

Рис. 1 – Структура XML-файлу

Для зберігання та опрацювання даних такого виду в пам'яті комп'ютера було на основі стандартної реалізації хеш-таблиць (hash tables) розроблено набір спеціалізованих структур даних, що забезпечують високу швидкість доступу до інформації про семантичні зв'язки на основі повної назви або фрагменту назви однієї з сутностей чи зв'язку між ними.

2. СЕМАНТИЧНИЙ АНАЛІЗ ТЕКСТІВ

В ході проектування програмного забезпечення для семантичного розпізнавання текстів природньою мовою було здійснено аналіз існуючих систем, виділено їх переваги та недоліки. Огляд існуючих продуктів виявив, що ринок засобів семантичного аналізу на сьогоднішній день перебуває в стані занепаду, особливо це стосується української частини ринку. Наразі не існує жодного комерційного програмного забезпечення для семантичного розпізнавання текстів українською мовою. Іншомовні системи переважно базуються на статистичних алгоритмах та зазвичай не можуть бути використані для побудови семантичних мереж на основі заданого тексту природньою мовою.

Загалом можна виділити три основні алгоритмічні підходи до семантичного аналізу текстів природньою мовою:

1. Спосіб генерації всіх граматично правильних словосполучень з подальшим обходом мережі синтактико-семантичних відношень [9]. Такий спосіб не можна вважати достатньо достовірним: генерування всіх граматично правильних словосполучень на основі введеного тексту не гарантує наявності семантичного зв'язку між обраними генератором сутностями.
2. Статистичний спосіб встановлення зв'язків між сутностями на основі підрахунку частоти їх спільної появи в тексті. Даний спосіб є застосовним лише для текстів значного обсягу, оскільки не дає змоги отримати достовірний результат для сутностей, які згадуються в тексті статистично незначну кількість разів. Водночас спосіб є надзвичайно універсальним та може виявляти асоціативні зв'язки між сутностями навіть у граматично некоректному або особливого стилю тексті.
3. Синтаксичний спосіб встановлення зв'язків передбачає виявлення сутностей, змістовно пов'язаних з іншими сутностями на основі визначення синтаксичних конфігурацій в реченні згідно набору апріорних шаблонів [10]. Оскільки реалізація цього способу для кожної нової мови є складною алгоритмічною задачею, його не можна назвати достатньо універсальним. Разом з тим спосіб володіє високою точністю виявлення зв'язків та дає змогу отримати достовірні результати навіть для текстів вкрай малого розміру.

При розробці програмного забезпечення для семантичного розпізнавання текстів природньою мовою було вирішено використовувати синтаксичний спосіб встановлення зв'язків, який ґрунтується на синтаксичному аналізі, який в свою чергу вимагає попереднього виконання лексичного аналізу (визначення частини мови, до якої належить кожна лексема).

Синтаксичний аналіз – це процес співставлення лінійної послідовності лексем мови з її формальною граматикою. Зазвичай результатом такого співставлення є синтаксичне дерево – структура даних, яка відображає синтаксичну структуру вхідного речення. Приклад синтаксичного розбору простого речення представлено в графічній формі на Рис. 2.

Рис. 2 – Синтаксичний розбір простого речення

Семантичний аналіз текстів природньою мовою є складною проблемою з галузі штучного інтелекту. Вичерпне вирішення цієї проблеми виходить далеко за межі даної роботи. Тим не менше було знайдено часткове рішення, яке дало

задовільні результати на використовуваному корпусі текстів. Запропонований алгоритм складається з таких кроків:

[A1] Розбити речення на лексеми.
[A2] Визначити для кожної лексеми її частину мови.
[A3] Визначити імовірний семантичний зв'язок у реченні.
[A4] Визначити імовірний суб'єкт у реченні.
[A5] Визначити імовірний об'єкт у реченні.
[A6] Опрацювати виявлений семантичний зв'язок між суб'єктом та об'єктом.

Визначення частини мови, до якої належить певне слово (крок A2) є нетривіальною задачею. Для її розв'язання розроблене програмне забезпечення здійснює інтеграцію з текстовим процесором Microsoft Word 2010 та використовує можливості вбудованого тезауруса. Недоліком такого способу є неможливість коректного опрацювання власних назв з некапіталізованою першою літерою.

3. ВІЗУАЛІЗАЦІЯ ГРАФОВИХ СТРУКТУР

Графова структура даних – це множина вершин та ребер, що з'єднують ці вершини. В контексті семантичних мереж графова структура представляє собою онтологію, її вершини – множину сутностей, а ребра – відношення, в яких перебувають ці сутності.

Задача візуалізації структур такого виду почала набувати все більшої актуальності разом з розвитком засобів графічного виводу обчислювальних машин, а саме з кінця XX століття. У зв'язку з невпинним ростом графічних можливостей сучасних комп'ютерів та стрімким ускладненням моделей, що використовуються в різних галузях науки і техніки, ця задача потребує з кожним роком все більш якісного розв'язку.

Серед значної кількості різноманітних алгоритмів візуалізації графів найчастіше використовуються модифікації таких шести способів:

1. Найпростішим алгоритмом візуалізації графових структур є хаотичний (Randomized) алгоритм. Він передбачає розміщення вершин та ребер у довільному порядку та положенні. Очевидними перевагами цього алгоритму є вкрай високі швидкодія та простота програмної реалізації. Проте алгоритм володіє також значною кількістю недоліків: кількість ребер що перетинаються та сума відстаней між контекстно-близькими вершинами є неприпустимо високими, а візуальна картина представлення даних не адекватна структурі взаємозв'язків. Типовий вигляд графа при візуалізації за допомогою хаотичного алгоритму представлено на Рис. 3a.

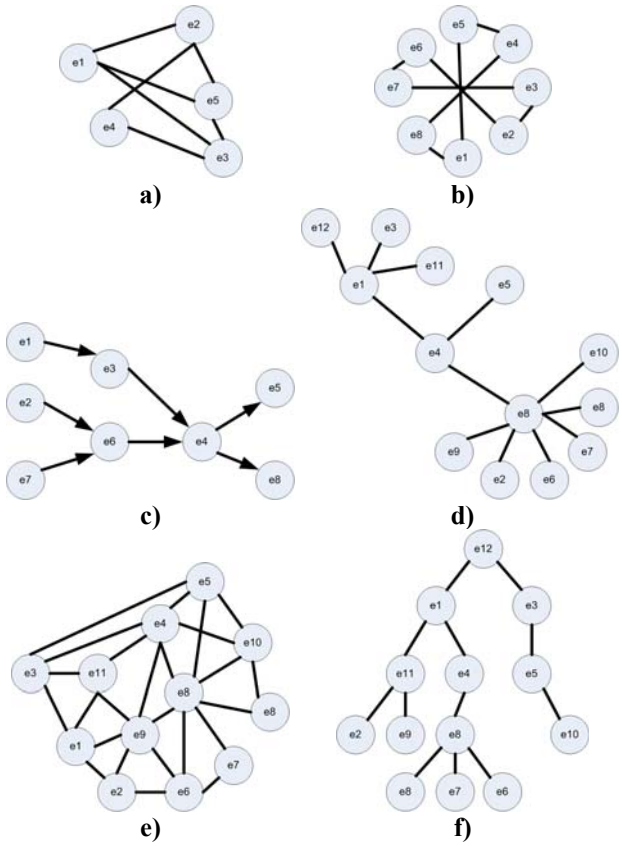


Рис. 3 – Типовий вигляд графа при візуалізації за допомогою алгоритму: a) хаотичного; b) кругового; c) Сугіями; d) Фрухтермана–Райнголда; e) Камади–Кавая; f) променевого дерева

2. Достатньо простим в реалізації є круговий (Circular) алгоритм візуалізації графів. Вершини розміщуються на колі через рівні проміжки в порядку, який дає можливість мінімізувати кількість ребер що перетинаються. Цей алгоритм дає змогу отримати задовільний результат за прийнятний час. Попри це застосовність алгоритму є обмеженою у випадку складних графів великої розмірності, оскільки не відображає належним чином структурні особливості графа. Типовий вигляд графа при візуалізації за допомогою кругового алгоритму представлено на Рис. 3b.
3. Алгоритм Козо Сугіями (Kozo Sugiyama) було описано в роботі [11]. Схема Сугіяма відноситься до методів порівневої візуалізації та призначена перш за все для оптимізації візуальних критеріїв, таких як кількість перетинів між ребрами та їх сумарна довжина. У статті [12] автор описує новий евристичний метод для зображення графів, названий

«Magnetic Spring Model», та, на відміну від попередніх, придатний також для візуалізації орієнтованих графів. Алгоритми Сугіями довели свою ефективність та застосовуються для візуалізації графових структур даних в багатьох програмних системах. Типовий вигляд графа при візуалізації за допомогою алгоритму Сугіями представлено на Рис. 3c.

4. Алгоритм Фрухтермана–Райнґолда (Fruchterman–Reingold) є одним із представників класу «керованих силами» («Force-Directed») алгоритмів візуалізації графів [13]. Цей клас алгоритмів спрямований на оптимізацію візуально-естетичних характеристик (зокрема на мінімізацію кількості ребер що перетинаються) при вирішенні задачі візуалізації графів із хаотично-розміщеними вузлами; такі алгоритми досліджував також Джузеппе ді Батіста (Giuseppe di Battista) [14]. Застосування алгоритму Фрухтермана–Райнґолда є доцільним у випадку великих неорієнтованих графів, оскільки гарантується геометрична близькість пов'язаних вузлів. Проте низька швидкодія алгоритму унеможливлює його застосування для графових структур надто великої розмірності. Типовий вигляд графа при візуалізації за допомогою алгоритму Фрухтермана–Райнґолда представлено на Рис. 3d.
5. Іншим представником класу «керованих силою» алгоритмів є алгоритм Камади–Кавая (Kamada–Kawai). Попри деяку подібність до попереднього алгоритму, основним його принципом є мінімізація енергії шляхом пошуку похідної від рівнянь сили [15]. Алгоритм Камади–Кавая виконує впорядкування вузлів графу досить швидко і тому не накладає обмежень на розмір графів. Попри це, згідно візуально-естетичних критеріїв результат застосування цього алгоритму може потребувати деякого покращення, яке здійснюється, переважно, засобами алгоритму Фрухтермана–Райнґолда. Типовий вигляд графа при візуалізації за допомогою алгоритму Камади–Кавая представлено на Рис. 3e.
6. Алгоритм променевого дерева (Radial Tree) є одним із методів, що візуалізують інформацію у вигляді деревовидних структур [16]. Суть алгоритму полягає у розміщенні самостійної вершини в геометричному центрі області побудови та в наступному послідовному радіальному розміщенні залежних вершин. Типовий вигляд графа при візуалізації за допомогою променевого алгоритму представлено на Рис. 3f.

При розробці підсистеми візуалізації даних системи керування семантичними мережами було здійснено програмну реалізацію трьох алгоритмів візуалізації графових структур, а саме: хаотичного, кругового та «керованого силами» алгоритму Фрухтермана–Райнґолда. Вибір того чи іншого алгоритму візуалізації здійснюється користувачем та не залежить від характеристик семантичної мережі. Всі реалізовані алгоритми підтвердили свою застосовність та ефективність для вирішення задачі візуалізації семантичних мереж.

4. WEB-ІНТЕРФЕЙС СИСТЕМИ

В ході проектування системи керування семантичними мережами було прийнято рішення про реалізацію користувацького інтерфейсу у вигляді Web-додатку. Такий підхід, перш за все, забезпечить зручність використання та кросплатформенність системи, а, крім того, спростить імплементацію розширюваності та балансування навантаження в майбутньому.

Розробку користувацького інтерфейсу було здійснено для платформи Microsoft ASP .NET 4.0 з використанням технології асинхронних JavaScript та XML (Asynchronous Javascript And Xml, AJAX), яка є частиною концепції динамічного HTML (Dynamic HTML, DHTML) та дозволяє фонове виконання запитів до сервера без перезавантаження усієї сторінки [17, 18].

При розробці користувацького інтерфейсу системи було проведено аналіз зручності використання згідно рекомендацій, наведених у [19, 20]. Загальний вигляд користувацького інтерфесу під час роботи з системою представлено на Рис. 4.

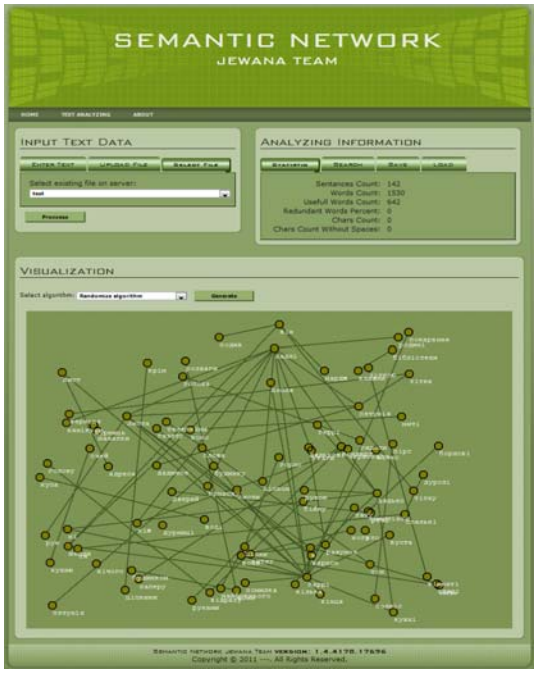


Рис. 4 – Вигляд користувацького інтерфейсу

Після завершення розробки користувацького інтерфейсу було виконано навантажене тестування (Load Test) та стресс-тестування (Stress Test) з наступною оптимізацією швидкодії, що дало змогу підвищити стабільність та надійність Web-додатку.

ВИСНОВКИ

Розроблена система керування семантичними мережами є повністю функціональним засобом побудови та дослідження семантичних мереж, що підтвердив свою ефективність.

Реалізовані алгоритми семантичного аналізу текстів природньою мовою та спеціалізовані структури для збереження даних підтвердили можливість опрацьовування за прийнятний час (десятки секунд) значних обсягів тексту ункраїнською мовою (сотні тисяч знаків), створення на його основі достатньо великих онтологій (тисячі сутностей) та, що найголовніше, швидкого (долі секунди) виконання запитів на отриманих даних. Всі реалізовані алгоритми мають обчислювальну складність лінійно або сублінійно залежну від розміру опрацьовуваного тексту.

Спосіб реалізації користувацького інтерфейсу виправдав себе, а обрані та реалізовані методи візуалізації графових структур зробили можливим представлення даних в інтуїтивно-зрозумілому вигляді.

Попри те, що розроблена система навряд чи може бути без додаткових змін використана для вирішення реальних бізнес-задач контент-аналізу, вона послужить базовою платформою для подальших досліджень авторського колективу, зокрема в напрямку використання гіперграфів для реалізації семантичних мереж з *n*-арними зв'язками, навантажування ребер мітками дати і часу для розв'язання семантичних конфліктів тощо.

СПИСОК ЛІТЕРАТУРИ

[1] Biological classification, *Wikipedia,* http://en.wikipedia.org/wiki/Biological_classification, – 2012-01-01.

[2] Gottlob Frege, *Wikipedia,* http://en.wikipedia.org/wiki/Gottlob_Frege, 2012-01-01.

[3] Existential Graph, *Wikipedia,* http://en.wikipedia.org/wiki/Existential_graph, 2012-01-01.

[4] Tesnière L., *Elements of Structural Syntax,* Second Edition, Klincksieck, 1976, 674 p. (In French)

[5] Yngve V., Yates D. M., Masterman M., von Glasersfeld E., Machine translation researchers: their works and contribution into the development of machine translation, http://www.br.com.ua/referats/ Computers/ 10099.html. – 2012-01-01. (In Ukrainian)

[6] What are Semantic Networks? http://www.cs.bham.ac.uk/research/projects/poplog/computers-and-thought/chap6/node5.html. – 2012-01-01.

[7] Tim Berners-Lee, http://en.wiki pedia.org/ wiki/Tim_Berners-Lee. – 2012-01-01.

[8] Dublin Core, http://en.wiki pedia.org/wiki/ Dublin_Core. – 2012-01-01.

[9] Yermakov A. E., Explication of meaningful elements from text by the means of syntactic analysis-synthesis, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference «Dialogue»*. Nauka, (2003), pp. 136-140. (In Russian)

[10] Kisilyov S. L., Yermakov A. E. and Plyeshko V. V., Search for facts in natural language text based on network descriptions, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference «Dialogue»*, Nauka, (2004), pp. 282-285. (In Russian)

[11] Sugiyama K., Tagawa S., Toda M., Methods for visual understanding of hierarchical systems, *IEEE Transactions on Systems, Man and Cybernetics*, (11) 2 (1981), pp. 109-125.

[12] Sugiyama K. and Misue K., Graph drawing by the magnetic spring model, *Journal of Visual Languages and Computing*, (6) 3 (1995), pp. 217-231.

[13] Fruchterman T. M. J. and Reingold E. M., Graph drawing by force-directed placement, *Software: Practice and Experience*, (21) 11 (1991).

[14] Battista G., Eades P., Tamassia R. and Tollis I. G., Algorithms for drawing graphs: an annotated bibliography, *Computational Geometry: Theory and Applications*, (4) 5 (1994), pp. 235-282.

[15] Kamada T. and Kawai S., An algorithm for drawing general undirected graphs, *Information Processing Letters*, 31 (1989), pp. 7-15.

[16] Di Battista G., Eades P., Tamassia R. and Tollis I. G., *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, 1998, 397 p.

[17] Holzner S. *AJAX Bible*, Wiley, 2007, 695 p.

[18] Woolston D., *Pro AJAX and the .NET 2.0 Platform*., Apress, 2006, 488 p.

[19] Borodayev D. V., *Website as Object of Graphic Design*, Septima, 2006, 288 p. (In Russian)

[20] Nielsen J. and Pernice K., *Eyetracking Web Usability*, New Riders Press, 2009, 456 p.

Юрій Семчишин, *кандидат технічних наук, старший викладач кафедри програмного забезпечення Національного університету «Львівська політехніка». Народився 1984 року в місті Івано-Франківську, 2005 року отримав диплом бакалавра комп'ютерних наук з відзнакою, а 2006 року – диплом магістра програмного забезпечення автоматизованих систем з відзнакою, захистив кандидатську дисертацію в 2010 році. Наукові інтереси зосереджені навколо вирішення різноманітних задач штучного інтелекту.*

Іван Кульпа, *бакалавр програмної інженерії, студент магістратури кафедри програмного забезпечення Національного університету «Львівська політехніка». Народився 1990 року в місті Івано-Франківську, захистив бакалаврську роботу в 2011 році.*

Цікавиться вирішенням задач візуалізації даних.

Ігор Колосовський, *бакалавр програмної інженерії, студент магістратури кафедри програмного забезпечення Національного університету «Львівська політехніка». Народився 1990 року в місті Львові, захистив бакалаврську роботу в 2011 році.*

Цікавиться опрацюванням текстів природньою мовою.

Олександр Гречніков, *бакалавр програмної інженерії, студент магістратури кафедри програмного забезпечення Національного університету «Львівська політехніка». Народився 1990 року в місті Львові, захистив бакалаврську роботу в 2011 році.*

Цікавиться розробкою Web-додатків.

Петро Гайда, *бакалавр програмної інженерії, студент магістратури кафедри програмного забезпечення Національного університету «Львівська політехніка». Народився 1990 року в місті Львові, захистив бакалаврську роботу в 2011 році.*

Цікавиться розробкою високооптимізованих структур даних.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

DEVELOPING SEMANTIC NETWORK MANAGEMENT SYSTEM

Yuriy Semchyshyn, Ivan Kulpa, Igor Kolosovskyi, Oleksandr Hrechnikov, Petro Hayda

Lviv National Technical University
12 Stepana Bandery Str., Lviv, 79013, Ukraine
e-mail: 7th@ukr.net, ivan.kulpa@gmail.com, igor.kolosovski@gmail.com, sansofter@gmail.com, fezundo@gmail.com

Abstract: *The paper considers semantic networks. It is a way of representing knowledge as a set of nodes-concepts connected by edges-relations. The history and current state-of-the-art of semantic networks is analyzed. The experience of designing and developing semantic network management system is described in four sections covering the following topics: data structures design, semantic analysis algorithm implementation, graph visualization methods selection and Web-based user interface development. The developed system is fully-operational tool for building and studying semantic networks. The system will be used for the further research.*

Keywords: *Semantic Network Management System, Semantic Network, Semantic Analysis, Graph Visualization.*

1. INTRODUCTION

Semantic Network is a way of representing knowledge as a set of nodes (some concepts), which are connected by edges (corresponding relations).

The idea of using semantic relations to classify entities have been known since the first half of XVIII century, but only in the late XX century did semantic networks become applicable for solving such problems of artificial intelligence as machine translation and information retrieval [1 – 8].

Amongst the modern semantic network management systems the most sophisticated ones are the following five:

- WordNet (http://wordnet.princeton.edu);
- LexiPedia (http://lexipedia.com);
- MLSN (http://dcook.org/mlsn/);
- TextAnalyst (http://megaputer.com/textanalyst.php);
- UMLS (http://www.nlm.nih.gov/research/umls/).

None of the mentioned projects can be considered as general-purpose semantic network management system. Development of such system is a topical task which will become even more urgent with further popularization of the Semantic Web conception.

2. DATA STRUCTURES

The semantic network management system was developed in accordance with principles of the object-oriented programming paradigm in the C# programming language for the Microsoft .NET 4.0 software platform using the Microsoft Visual Studio 2010 development environment.

The following three ways of storing semantic network in computer memory were carefully evaluated:

- using Adjacency Matrix;
- using Incidence Lists;
- using Tree-Like Structures.

Usage of incidence lists appeared to be the most efficient way; hence, a set of specialized hash-based speed-optimized data structures were designed.

Considering structural nature of data being processed, it was settled to use a set of local files of specific structure as data storage (Fig. 1).

3. SEMANTIC ANALYSIS

There are three possible approaches to semantic analysis of texts in natural language:

1. Combinatorial Approach [9].
2. Statistic Approach.
3. Syntactic Approach [10].

It was decided to implement syntactic approach to semantic analysis which implies lexical analysis of source sentence and recognizes some heuristic patterns so as to build syntactic tree (Fig. 2).

In order to define which part of speech the particular word is, the developed semantic network management system integrates with the Microsoft Word 2010 text processor to utilize its thesaurus.

4. GRAPH VISUALIZATION

The most frequently used graph visualization algorithms are modifications of the following six methods:

1. Randomized Algorithm (Fig. 3a).
2. Circular Algorithm (Fig. 3b).
3. Sugiyama Algorithm (Fig. 3c) [11, 12].
4. Fruchterman–Reingold Algorithm (Fig. 3d) [13, 14].
5. Kamada–Kawai Algorithm (Fig. 3e) [15].
6. Radial Tree Algorithm (Fig. 3f) [16].

In the developed semantic network management system randomized, circular and Fruchterman–Reingold algorithms are available for user's choice.

5. WEB INTERFACE

The Web-based user interface of the developed semantic network management system were designed, implemented and tested using the Microsoft ASP .NET web technology in accordance with all the best practices recommended (Fig. 4) [17 – 20].

6. CONCLUSION

The developed semantic network management system is fully-operational and well-tried tool for building and studying semantic networks.

Implemented algorithms for semantic analysis of texts in natural language and corresponding data structures have proven their efficiency. All the implemented algorithms do have computational complexity that depends linearly or sublinearly on the size of input text.

Developed Web-based user interface implements three different graph visualization algorithms and assures intuitive and friendly interaction.

The developed system will be used as a platform for further research specifically in the directions of using hypergraphs for representing semantic networks with *n*-ary relations, weighting the edges with datetime stamps for resolving semantic conflicts et cetera.

7. REFERENCES

[1] Biological classification, *Wikipedia*, http://en.wikipedia.org/wiki/Biological_classification, – 2012-01-01.

[2] Gottlob Frege, *Wikipedia*, http://en.wikipedia.org/wiki/Gottlob_Frege, 2012-01-01.

[3] Existential Graph, *Wikipedia*, http://en.wikipedia.org/wiki/Existential_graph, 2012-01-01.

[4] Tesnière L., *Elements of Structural Syntax*, Second Edition, Klincksieck, 1976, 674 p. (In French)

[5] Yngve V., Yates D. M., Masterman M., von Glasersfeld E., Machine translation researchers: their works and contribution into the development of machine translation, http://www.br.com.ua/referats/ Computers/ 10099.html. – 2012-01-01. (In Ukrainian)

[6] What are Semantic Networks? http://www.cs.bham.ac.uk/research/projects/poplog/computers-and-thought/chap6/node5.html. – 2012-01-01.

[7] Tim Berners-Lee, http://en.wiki pedia.org/wiki/Tim_Berners-Lee. – 2012-01-01.

[8] Dublin Core, http://en.wiki pedia.org/wiki/Dublin_Core. – 2012-01-01.

[9] Yermakov A. E., Explication of meaningful elements from text by the means of syntactic analysis-synthesis, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference «Dialogue»*. Nauka, (2003), pp. 136-140. (In Russian)

[10] Kisilyov S. L., Yermakov A. E. and Plyeshko V. V., Search for facts in natural language text based on network descriptions, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference «Dialogue»*, Nauka, (2004), pp. 282-285. (In Russian)

[11] Sugiyama K., Tagawa S., Toda M., Methods for visual understanding of hierarchical systems, *IEEE Transactions on Systems, Man and Cybernetics*, (11) 2 (1981), pp. 109-125.

[12] Sugiyama K. and Misue K., Graph drawing by the magnetic spring model, *Journal of Visual Languages and Computing*, (6) 3 (1995), pp. 217-231.

[13] Fruchterman T. M. J. and Reingold E. M., Graph drawing by force-directed placement, *Software: Practice and Experience*, (21) 11 (1991).

[14] Battista G., Eades P., Tamassia R. and Tollis I. G., Algorithms for drawing graphs: an annotated bibliography, *Computational Geometry: Theory and Applications*, (4) 5 (1994), pp. 235-282.

[15] Kamada T. and Kawai S., An algorithm for drawing general undirected graphs, *Information Processing Letters*, 31 (1989), pp. 7-15.

[16] Di Battista G., Eades P., Tamassia R. and Tollis I. G., *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, 1998, 397 p.

[17] Holzner S. *AJAX Bible*, Wiley, 2007, 695 p.

[18] Woolston D., *Pro AJAX and the .NET 2.0 Platform*., Apress, 2006, 488 p.

[19] Borodayev D. V., *Website as Object of Graphic Design*, Septima, 2006, 288 p. (In Russian)

[20] Nielsen J. and Pernice K., *Eyetracking Web Usability*, New Riders Press, 2009, 456 p.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

СПІНОВА МОДЕЛЬ ПОВНОГО ОДНОРОЗРЯДНОГО СУМАТОРА НА ЕЛЕМЕНТАХ ПЕРЕСА

Віталій Дейбук [1], Іван Юрійчук [2], Роман Юрійчук [1]

[1] Кафедра комп'ютерних систем та мереж, факультет комп'ютерних наук,
[2] Кафедра фізики напівпровідників і наноструктур, фізичний факультет,
Чернівецький національний університет імені Юрія Федьковича,
вул. Коцюбинського 2, 58012, Чернівці, Україна,
v.deibuk@chnu.edu.ua, ivmykyur@gmail.com

Резюме: *Побудована чотириспінова модель однорозрядного суматора для твердотільного ЯМР квантового комп'ютера та проаналізована коректність його роботи в залежності від параметрів системи. Встановлено, що для чистих та суперпозиційних станів, суматор на елементах Переса коректно спрацьовує за чотири π-імпульси. Отримано оптимальні значення параметрів обмінної взаємодії на основі аналізу функції чіткості та визначена максимальна відстань між кубітами для коректної роботи суматора.*

Ключові слова: *квантовий біт, квантовий комп'ютер, логічний елемент Переса, спін, модель Ізінга, функція чіткості, повний однорозрядний суматор.*

SPIN MODEL OF FULL SUMMATOR ON PERES GATES

Vitaliy G. Deibuk [1], Ivan M. Yuriychuk [2], Roman I. Yuriychuk [1]

[1] Computer systems and networks department
[2] Physics of semiconductors and nanostructures department
Chernivtsi National University, 2 Kotsubins'kogo Str., 58012, Chernivtsi, Ukraine
e-mail: v.deibuk@chnu.edu.ua, ivmykyur@gmail.com

Abstract: *Four-spin model of single-digit summator for solid state nuclear magnetic resonance (NMR) quantum computer is proposed. Correctness of summator's operation is analyzed in dependence of system parameters. It is found that the summator on Peres gates works correctly in time of four π-pulses on pure numerical and superposition states. Optimal values for exchange interactions and maximum distance between q-bits necessary for correct summator's operation are obtained from the analysis of fidelity function.*

Keywords: *quantum bit, quantum computer, Peres logical gate, spin, Ising model, fidelity function, full single-digit summator.*

ВСТУП

Розвиток нових технологій з кожним днем веде до підвищення швидкодії, мініатюризації та складності комп'ютерної техніки. Це в свою чергу обумовлює збільшення кількості транзисторів в мікросхемах (чіпах), що викликає зростання споживаної потужності. При цьому мільйони вентилів, які виконують логічні операції в традиційних комп'ютерах, є незворотними. Тобто, кожного разу виконання логічної операції приводить до втрати або стирання деякої частини вхідної інформації, яка розсіюється у вигляді теплової енергії. Як було показано Ландауером [1], для незворотної логіки кожен біт втраченої інформації генерує $kT\ln 2$ Дж теплової енергії, де k – стала Больцмана, T – абсолютна температура. При кімнатних температурах на гігагерцових частотах сучасних процесорів, що містять сотні тисяч транзисторів, розсіювана енергія наближається до кількох Вт, що неминуче веде до експоненційної

некоректності як у розрахунках, так і до зниження часового ресурсу мікросхем. Вирішення цих проблем лежить у площині використання нових революційних технологій, які спроможні кардинально зменшити як споживану потужність, так і розсіяння теплової енергії в комп'ютерних системах. Вдалою альтернативою в цьому питанні можна вважати використання зворотної логіки, яка останнім часом досить швидко розвивається [2, 3], оскільки знаходить застосування у різноманітних областях, таких як квантовий комп'ютинг, нанотехнології, біоінформатика, оптичний комп'ютинг тощо, де важливою умовою є екстремально низьке розсіяння тепла. Можливість використання зворотних логічних операцій, які не знищують вхідну інформацію, теоретично не веде до розсіяння енергії в системах, що їх реалізують. Така логіка дозволяє відтворити вхідні стани за вихідними. Зворотний комп'ютинг використовує як термодинамічну, так і логічну зворотність і використовує адіабатичні системи, що відновлюють свою енергію, випромінюючи дуже мало тепла. Зворотними є кола (вентилі), в яких вектор вхідних станів завжди можна відновити з вектора вихідних станів. Логічні функції, які відповідають зворотним вентилям, повинні бути бієктивними, тобто реалізувати взаємно-однозначне відображення між вхідним та вихідним векторами. Розробці цифрових зворотних пристроїв присвячена велика кількість досліджень [2,3], однак в основному вони стосуються схемотехнічного аспекту. В даній роботі побудована спінова модель однорозрядного суматора на елементах Переса для твердотільного квантового комп'ютера на ядерному магнітному резонансі (ЯМР) та проаналізована коректність його роботи в залежності від параметрів системи.

2. ПОВНИЙ ОДНОРОЗРЯДНИЙ ЗВОРОТНИЙ СУМАТОР

Суматори є одними з основних блоків, які входять до складу більшості обчислювальних пристроїв. Описані вище необхідні зміни у логіці квантових обчислень вимагають відповідних змін у реалізації суматорів як на логічному, так і на фізичному рівні. В якості базисних логічних елементів можна вибрати зворотні функціонально повні елементи Тоффолі, Фредкіна, Переса та ін. [1,3].

Розглянемо тривходові зворотні елементи. Елемент Тоффолі (CCNOT – двічі контрольоване НЕ), зображений на рис. 1, виконує функцію:

$$P = A,\ Q = B,\ R = AB \oplus C\ .$$

Цей елемент є універсальним, тобто з його допомогою можна отримати довільну логічну функцію, однак він не зберігає парність $(A \oplus B \oplus C \neq P \oplus Q \oplus R)$.

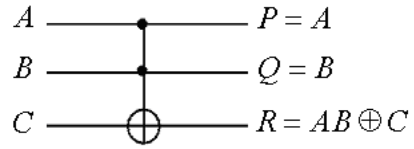


Рис. 1 – Елемент Тоффолі

У попередніх роботах [4,5] нами була проаналізована фізична модель квантового елемента Фредкіна (рис. 2), який є функціонально повним, зворотним логічним елементом, що зберігає парність, тобто вага за Хеммінгом вхідних сигналів зберігається на виході.

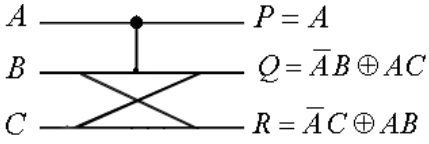


Рис. 2 – Елемент Фредкіна

Елемент Фредкіна (CSWAP – контрольований обмін) виконує функцію:

$$P = A,$$
$$Q = if\,(A) then (C) else (B),$$
$$R = if\,(A) then (B) else (C).$$

В даній роботі ми вибрали в якості базисного елемент Переса. Трикубітовий зворотний елемент Переса, як видно з рис. 3, поєднує функції елементів Фейнмана і Тоффолі:

$$P = A,\ \ Q = A \oplus B,\ \ R = AB \oplus C\ .$$

Хоча він не зберігає парність, однак має найменшу квантову вартість серед усіх трикубітових квантових елементів, є універсальним, а тому знайшов широке використання в комбінаційних зворотних схемах. Зокрема, квантова вартість елементів Тоффолі та Фредкіна становить 5, елемента Переса – 4 і т.д. [1,3].

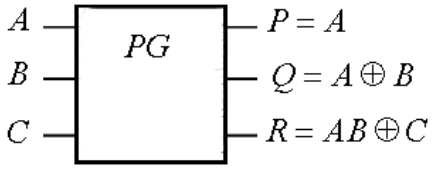


Рис. 3 – Елемент Переса

Дослідимо деякі схеми суматорів, побудовані на зворотних логічних елементах і оцінимо їх ефективність з точки зору використання для квантового комп'ютингу. Порівняно з класичними схемами у квантових є ряд обмежень. По-перше, у квантових схемах не допускаються «цикли», тобто зворотний зв'язок від однієї частини квантової схеми до іншої, схема має бути «ациклічною». По-друге, заборонена операція FANIN (можливість з'єднувати проводи в один, який містить побітове OR входів), яка є незворотною, а, отже, неунітарною. По-третє, в квантових схемах недопустима й обернена операція FANOUT (розгалуження за виходом) через теорему про заборону клонування квантових станів [1]. Крім цього, має місце умова нормування на квантові стани, що відповідає їх імовірнісному характеру.

Абстрагуючись від апаратної частини, побудуємо функціональні схеми суматорів на зворотних логічних елементах, відповідно до вище наведених критеріїв. Проведемо оцінку схем за квантовою вартістю та наявністю зайвих (стокових) виходів, використавши елементи Переса та Фредкіна.

Повний однорозрядний суматор виконує дві функції: додавання:

$$S = A \oplus B \oplus C_{in}$$

та формування перенесення в наступний розряд

$$C_{out} = (A \oplus B)C_{in} \vee AB.$$

Такий суматор можна побудувати на двох елементах Переса (рис. 4). Квантова вартість такої схеми становить 8, апаратна складність – 2.

При побудові квантових мереж велика увага приділяється так званим постійним входам та зайвим виходам (Garbage Inputs/Outputs). Наявність таких входів-виходів суперечить ідеї мінімізації втрат енергії, адже на зайві виходи енергія поступає та розсіюється і необхідно приймати спеціальні заходи щодо її утилізації. Подана схема містить 1 постійний вхід на 2 зайві виходи.

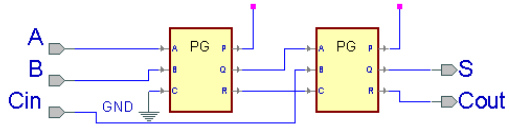


Рис. 4 – Повний однорозрядний суматор на елементах Переса

На базі елемента Фредкіна можна побудувати повний однорозрядний суматор [6] (рис. 5). Наведена схема не містить циклів та розгалужень по виходу, що відповідає описаним вище критеріям до квантових схем. Схема складається з п'яти елементів Фредкіна, має квантову вартість рівну 25, апаратну складність – 5, кількість постійних входів – 4, зайвих виходів– 4.

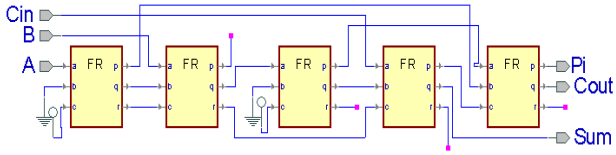


Рис. 5 – Логічна схема повного однорозрядного суматора на елементах Фредкіна [6]

На базі повного однорозрядного суматора можна побудувати довільний *n*-розрядний суматор. Враховуючи наведені вище критерії до квантових мереж, можна зробити висновок, що з точки зору як апаратної складності, так і квантової вартості та кількості зайвих виводів, схема повного однорозрядного суматора на елементах Переса є переважною. Саме тому проаналізуємо її фізичну модель.

3. АНАЛІЗ ФІЗИЧНОЇ МОДЕЛІ СУМАТОРА

Довільна дворівнева квантова система, в принципі, може бути використана для квантових обчислень. В таких обчисленнях для обробки інформації використовуються квантові біти (кубіти). Кубіт (КвБ) є суперпозицією двох основних квантових станів $|0\rangle$ і $|1\rangle$:

$$\Psi = c_0|0\rangle + c_1|1\rangle$$

при умові

$$|c_0|^2 + |c_1|^2 = 1.$$

Тензорний добуток *N*-кубітів задає регістр довжиною *N*: $|x\rangle = |i_{N-1}, \ldots, i_0\rangle$, де $i_j = 0,1$ і квантовий комп'ютер з *N*-кубітами працює в 2^N-вимірному гільбертовому просторі з елементами

$$\Psi = \sum c_x|x\rangle, \qquad \sum |c_x|^2 = 1.$$

Довільна операція з регістрами задається унітарним перетворенням, що визначає відповідний квантовий логічний елемент. Хоча квантові комп'ютери з кількома кубітами вже технічно реалізовані і на них успішно перевірені

власне квантові алгоритми, серйозні комп'ютерні розрахунки вимагають квантових комп'ютерів принаймні зі 100-кубітовими регістрами, на що слід сподіватися у недалекому майбутньому.

Модель твердотільного комп'ютера є достатньо реалістичною і зручною для аналітичного та числового моделювання квантових логічних елементів та пристроїв на їх основі. Основу такого підходу складає квантово-механічна модель Ізінга [7] для ланцюжка взаємодіючих між собою чотирьох спінів ½, котрі перебувають під дією сильного статичного магнітного поля (СМП) і керуючого слабкого поперечного радіочастотного магнітного поля (РЧМП) з круговою поляризацією. Гамільтоніан системи, тобто оператор повної енергії, має вигляд [7]:

$$\hat{H} = \hat{H}_0 + \hat{W}, \tag{1}$$

де

$$\hat{H}_0 = -\hbar\sum_{k=0}^{3}\{\omega_k I_k^z + 2J\left[I_0^z I_1^z + I_1^z I_2^z + I_2^z I_3^z + \right. \\ \left. + \alpha\left(I_0^z I_2^z + I_1^z I_3^z\right) + \alpha' I_0^z I_3^z\right]\}, \tag{2}$$

$$\hat{W} = -\hbar\Omega\sum_{k=0}^{3}\left[I_k^+ \exp(i\omega t) + I_k^- \exp(-i\omega t)\right]. \tag{3}$$

Оператор (1) при врахуванні (2) і (3) визначає поведінку системи взаємодіючих спінів у магнітному полі $B = \left(b\cos\omega t, -b\sin\omega t, B(z)\right)$, де b та ω – амплітуда та частота РЧМП, відповідно, $B(z)$ – залежна від координати z індукція СМП. Окрім того у формулах (2) і (3) I_k^z – оператори проекції k-го спіну на вісь z, $I_k^{\pm} = I_k^x \pm iI_k^y$ – так звані оператори підвищення і пониження, $\omega_k = \gamma B(z_k)$ – Ларморові частоти прецесії спінів, $\Omega = \gamma b$ – частоти осциляцій Рабі, γ – гіромагнітний фактор протона, J – стала обмінної взаємодії між найближчими сусідами, $\alpha = J'/J$ – відносний внесок обмінної взаємодії других сусідів (J'), $\hbar$ – стала Планка, поділена на 2π, α та α' – відносні внески взаємодії сусідів через одного і через два, відповідно [5]. Така модель має 16 станів, які можуть бути занумеровані десятковими числами $n = 1,2,...,16$, кожному з яких поставимо у відповідність його двійкове зображення $(i_3 i_2 i_1 i_0)$, де вміст кожного розряду дорівнює 0 або 1, що відповідає орієнтації спіну за або проти напрямку СМП. Через повноту системи базисних станів $|k\rangle$ довільний стан системи можна записати у вигляді їх лінійної комбінації:

$$\Psi(t) = \sum_{k=1}^{16} C_k(t)\,|k\rangle. \tag{4}$$

Хвильові функції базисних станів $|k\rangle$ є власними функціями оператора H_0

$$H_0\,|k\rangle = E_k\,|k\rangle, \tag{5}$$

тоді власні енергії базисних станів аналізованої системи визначаються формулою [5]:

$$E_n \equiv E_{i_3 i_2 i_1 i_0} = -0.5\hbar\left\{\sum_{k=0}^{3}(-1)^{i_k}\omega_k + \right. \\ + J\left[(-1)^{i_0+i_1} + (-1)^{i_1+i_2} + (-1)^{i_2+i_3} + \right. \\ \left. \left. + \alpha\left((-1)^{i_0+i_2} + (-1)^{i_1+i_3}\right) + \alpha'(-1)^{i_0+i_3}\right]\right\}. \tag{6}$$

Модель містить ряд параметрів, зокрема частоти осциляцій ω_k. Вибравши $B(z_k)$ так, щоб частоти осциляцій спінів дорівнювали $\omega_k = 2\omega_{k-1}$, енергетичний спектр системи представляє собою систему розщеплених ефектом Зеємана 16 еквідистантних рівнів, віддаль $\hbar\omega_0$ між якими визначається величиною статичного магнітного поля. Основним рівнем системи буде рівень (0000), а найвищим по енергії – рівень (1111). Числове моделювання проведено для наступних значень параметрів системи (в одиницях 2π МГц):

$$\omega_0 = 100,\quad \omega_1 = 200,\quad \omega_2 = 400,\quad \omega_3 = 800,$$
$$\alpha = \alpha' = 0{,}1,\quad J = 5,\quad \Omega = 0{,}1.$$

На рис. 6 представлена діаграма енергетичних рівнів системи. Стрілками на рисунку вказані дозволені переходи. Так з основного рівня (0000) можливими є переходи з переворотом одного зі спінів: (0000) → (0001), 1 → 2; (0000) → (0001), 1 → 3; (0000) → (0010), 1 → 5; (0100) → (1000), 1 → 9.

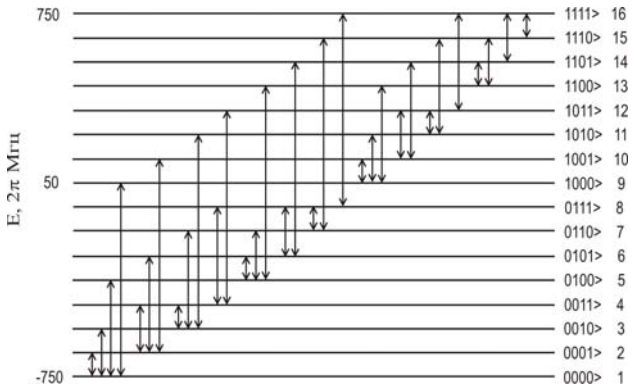


Рис. 6 – Діаграма енергетичних рівнів системи з чотирьох спінів

Динаміка системи описується нестаціонарним (залежним від часу *t*) рівнянням Шредінгера, яке має вигляд:

$$i\hbar\frac{\partial\Psi}{\partial t}=\hat{H}\Psi\,. \tag{7}$$

Беручи до уваги розклад (4), останнє рівняння еквівалентне системі 16-ти рівнянь з шістнадцятьма комплексними невідомими C_m (m = 1…16):

$$i\hbar\dot{C}_m=E_mC_m+\sum_{n=1}^{16}W_{mn}C_n\,, \tag{8}$$

де $W_{mn}(t)$ – матричний елемент оператора збурення (3), обчислений на хвильових функціях базисних станів. Для аналізованої системи матриця $W(t)$ має структуру [5].

Для спрощення (8) перейдемо до системи відліку, пов'язаної зі спінами, які прецесують у СМП, тобто представимо коефіцієнти C_m у вигляді:

$$C_m=D_m\exp\left(-iE_mt\hbar^{-1}\right). \tag{9}$$

Тоді для системи рівнянь (8) можна записати

$$\dot{D}_m=\sum_{n=1}^{16}T_{mn}(t)D_n\,, \tag{10}$$

причому матричні елементи $T_{mn}(t)$ мають наступний вигляд:

$$T_{mn}(t)=-\frac{i}{\hbar}W_{mn}(t)\exp(i\omega_{mn}t)\,, \tag{11}$$

де

$$\omega_{mn}=\frac{E_m-E_n}{\hbar}\,. \tag{12}$$

Система рівнянь (10) є системою диференціальних рівнянь першого порядку з невідомими $D_n(t)$. Програму числового інтегрування системи розроблено і реалізовано в середовищі MatLab з використання процедури ode113.

Розглянемо спочатку роботу повного однорозрядного суматора, що складається з двох елементів Переса (рис. 7), на чистих цифрових станах. Вхідний вектор повного суматора складається з чотирьох бітів, причому третій має нульове значення. В таблиці 1 представлена таблиця істинності першого блоку однорозрядного суматора, що містить перший елемент Переса.

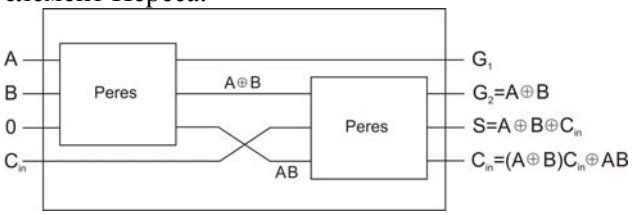


Рис. 7 – Повний суматор на елементах Переса

Таблиця 1. Таблиця істинності першого блоку однорозрядного суматора

A	B	C	C_{IN}	A'	B'	C'	C'_{IN}
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1
0	1	0	0	0	1	0	0
0	1	0	1	0	1	0	1
1	0	0	0	1	1	0	0
1	0	0	1	1	1	0	1
1	1	0	0	1	0	1	0
1	1	0	1	1	0	1	1

Прямі переходи 13(1100) → 11(1010) і 14(1101) → 12(1011) заборонені, оскільки під час них відбуваються перевороти відразу двох спінів. Однак дозволеними є двостадійні переходи через проміжний стан (рис. 8). Для переходу 13(1100) → 11(1010) можливі два варіанти: перехід через проміжний стан 9(1000):

13(1100) → 9(1000) → 11(1010), і перехід через проміжний стан 15(1110): 13(1100) → 15(1110) → 11(1010). Для переходу 14(1101) → 12(1011) можливі два варіанти: перехід через проміжний стан 10(1001): 14(1101) → 10(1001) → 12(1011), і перехід через проміжний стан 16(1110): 14(1101) → 16(1110) → 12(1011).

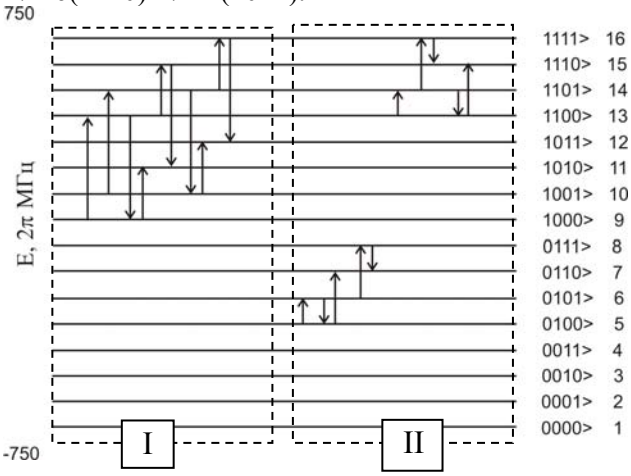


Рис. 8 – Дозволені переходи, необхідні для роботи першого (I) та другого (II) блоку суматора

Зауважимо, що для здійснення будь-якого дозволеного переходу з чистого стану з номером L у чистий стан з номером M необхідно налаштувати РЧМП у резонанс, тобто вибрати $\omega = |\omega_{LM}|$. У цьому випадку систему рівнянь (10) можна проінтегрувати аналітично. Тоді, наприклад, для переходу 13 → 9 отримуємо для амплітуд станів 13 і 9:

$$D_{13}(t) = \cos\frac{\Omega t}{2}, \qquad (13)$$

$$D_9(t) = i\sin\frac{\Omega t}{2}. \qquad (14)$$

Враховуючи, що ймовірності реалізації базисних станів $P_n(t)$ дорівнюють квадратам модулів елементів матриці $D_n(t)$, остаточно отримаємо такі залежності ймовірностей реалізації станів 13 і 9 від часу:

$$P_{13}(t) = \cos^2\frac{\Omega t}{2}, \qquad (15)$$

$$P_9(t) = \sin^2\frac{\Omega t}{2}. \qquad (16)$$

З формул (15) та (16) видно, що протягом так званого π-імпульсу, тривалість якого $t_0 = \pi/\Omega$, ймовірність реалізації початкового чистого стану з номером 13 плавно спадає від 1 до 0, а ймовірність реалізації кінцевого чистого стану з номером 9 плавно зростає від 0 до 1. Імовірність решти станів весь час залишається нульовою. Це справедливо для будь-якої пари чистих станів, переходи між якими дозволені. Таким чином, на чистих станах за умови настроювання РЧМП на кожній стадії в резонанс, однорозрядний суматор на першому етапі чітко, тобто з імовірністю 1, спрацьовує за два π-імпульси.

Вихідний вектор першого блоку однорозрядного суматора подається на другий блок (рис. 7). Отже, вхідним вектором блоку з другим елементом Переса є чотирибітовий вектор, значення першого біту якого не змінюється.

В таблиці 2 наведена таблиця істинності другого блоку однорозрядного суматора. У даному випадку забороненими є прямі переходи 6(0101) → 7(0110) і 14(1101) → 15(1110), тому для їх реалізації необхідні двостадійні переходи через проміжний стан. Для переходу 6(0101) → 7(0110) можливі два варіанти – перехід через проміжний стан 5(0100): 6(0101) → 5(0100) → 7(0110), і перехід через проміжний стан 8(0111): 6(0101) → 8(0111) → 7(0110). Для переходу 14(0101) → 15(0110) можливі два варіанти – перехід через проміжний стан 16(1111): 14(0101) → 16(1111) → 15(0110), і перехід через проміжний стан 13(1100): 14(0101) → 13(1100) → 15(0110). На рис. 8. стрілками зображені дозволені переходи, необхідні для роботи другого блоку однорозрядного суматора.

Таблиця 2. Таблиця істинності другого блоку однорозрядного суматора

A'	B'	C'_{IN}	C_{IN}	G_1	G_2	S	C_{OUT}
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1
0	1	0	0	0	1	0	1
0	1	0	1	0	1	1	0
1	1	0	0	1	1	0	1
1	1	0	1	1	1	1	0
1	0	1	0	1	0	1	0
1	0	1	1	1	0	1	1

В таблиці 3 наведені всі можливі переходи, необхідні для реалізації роботи однорозрядного суматора. Отже, в загальному випадку повний

однорозрядний суматор чітко спрацьовує на цифрових станах за чотири π-імпульси.

Таблиця 3. Переходи, що реалізують повний однорозрядний суматор

1-й блок	2-й блок
1 → 1	1 → 1
2 → 2	2 → 2
5 → 5	5 → 6
6 → 6	6 → 5 → 7 6 → 8 → 7
9 → 13	13 → 14
10 → 14	14 → 16 → 15 14 → 13 → 15
13 → 9 → 11 13 → 15 → 11	11 → 11
14 → 10 →12 14 → 16 → 12	12 → 12

4. ДОСЛІДЖЕННЯ КОРЕКТНОСТІ РОБОТИ СУМАТОРА

Аналізу переходів між чистими цифровими станами недостатньо для встановлення оптимальних параметрів системи. Крім того, з точки зору застосування квантового паралелізму прямий інтерес представляє реалізація однорозрядного суматора на суперпозиційних станах. Для цього розглянемо два суперпозиційн стани. Перший з них є рівномірною суперпозицією базисних станів:

$$\Psi_1(0) = \left(\sqrt{16}\right)^{-1} \sum_{k=1}^{16} |k\rangle . \qquad (17)$$

Другий – це такий суперпозиційний стан, в якому цифрові стани 1 – 8 вільні, а решта – рівномірно заповнені з імовірністю 1/8:

$$\Psi_2(0) = \left(\sqrt{8}\right)^{-1} \sum_{k=9}^{16} |k\rangle . \qquad (18)$$

Подіємо на ці стани першим блоком однорозрядного суматора. Позначимо утворені при цьому очікувані матриці коефіцієнтів розкладу по базисних станах через $D^{(1e)}$ та $D^{(2e)}$. Реальні матриці, утворені дією на вказані вище початкові суперпозиційні стани того ж однорозрядного суматора при менших значеннях α в діапазоні $0 \le \alpha \le 0{,}1$ і незмінних інших параметрах системи, позначимо відповідно через $D^{(1r)}$ та $D^{(2r)}$. Міру збігу реальних матриць з очікуваними визначимо через функцію чіткості *F* [7]:

$$F^{(1)} = \left| D^{(1e)+} D^{(1r)} \right|, \qquad (19)$$

$$F^{(2)} = \left| D^{(2e)+} D^{(2r)} \right|, \qquad (20)$$

де хрестик означає ермітове спряження. В обох випадках матриця-рядок множиться на матрицю-стовпець. Вважатимемо, що при кожному вибраному значенні α однорозрядний суматор працює коректно, якщо функції чіткості (19) і (20) дорівнюють 1.

Розглянемо двостадійні переходи, наприклад переходи 13(1100) → 9(1000) → 11(1010). Графіки функцій чіткості, побудовані на основі розв'язання системи (10) для обох суперпозиційних станів представлені на рис. 9(а). Результати розрахунків для рівномірної суперпозиції зображено суцільною лінією, а для нерівномірної суперпозиції – пунктирною. Як випливає з рис. 9(а), рівномірна суперпозиція є більш «жорсткою», оскільки для неї функція чіткості при α<0,04 має менші значення. Обидві функції чіткості осцилюють в залежності від α з періодом, який для вибраних параметрів задачі приблизно дорівнює 0,02. Найменше значення α, при якому значення функції чіткості дорівнюють 1, становить 0,04. Але вже при α= 0,03 (у другому мінімумі) ці значення досить істотно відрізняються від 1 і дорівнюють відповідно 0,96 для нерівномірної і 0,94 для рівномірної суперпозиції. Висновки, зроблені при розгляді переходів 13(1100) → 9(1000) → 11(1010), залишаються справедливими і для інших переходів, які необхідні для роботи повного однорозрядного суматора (табл. 3). Аналогічні висновки отримуються з аналізу інших, в тому числі двостадійних переходів, необхідних для роботи однорозрядного суматора як для рівномірної, так і нерівномірної суперпозицій.

Взаємодія між ядерними спінами виникає внаслідок обмінної взаємоді електронів сусідніх кубітів, якщо вони розташовані достатньо близько один до одного і хвильові функції електронів яких перекриваються, тому величина обмінної взаємодії між спінами залежить від відстані між спінами.

Якщо спіни знаходяться на відстані *r*, то для обмінної енергії справедливо [8]:

$$4J(r) \cong 1{,}6\frac{e^2}{\varepsilon a_B}\eta^{5/2}\exp(-2\eta), \qquad (21)$$

де $\eta = r/a_B$, ε – діелектрична постійна середовища, a_B – борівський радіус.

Проаналізована залежність функції чіткості спрацювання однорозрядного суматора від відстані (η) між спінами (рис. 9(*b*)) для рівномірної (суцільна крива) і нерівномірної (пунктирна крива) суперпозицій при переходах 13(1100)→9(1000)→11(1010). Максимальне значення параметра η, для якого однорозрядний суматор спрацьовує коректно, дорівнює η = 1,52. Залежність функції чіткості від параметра η для інших переходів виявляє таку ж закономірність. Враховуючи, що для іонів фосфору в кремнії борівський радіус складає 3 нм, для максимальної віддалі між кубітами при якій суматор коректно працює отримуємо величину 4,56 нм.

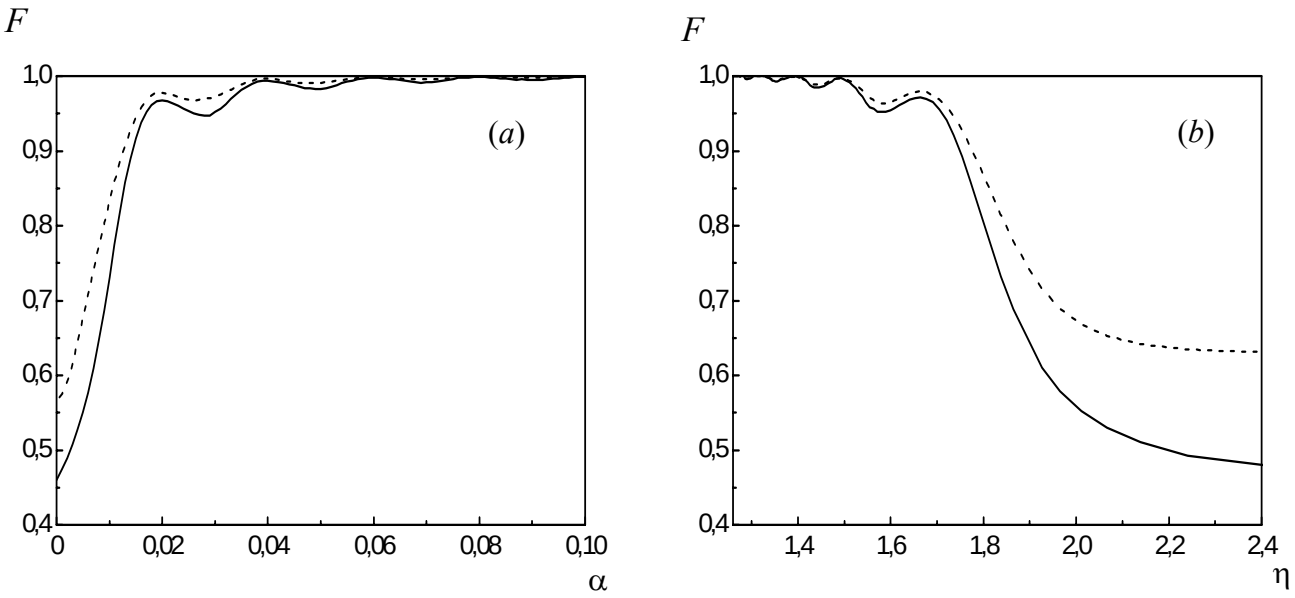


Рис. 9 – Залежність функцій чіткості спрацювання однорозрядного суматора для рівномірної (суцільна крива) і нерівномірної (пунктирна крива) суперпозицій від α (a) та η (b)

5. ВИСНОВКИ

Побудована чотириспінова модель однорозрядного суматора для твердотільного ЯМР квантового комп'ютера та проаналізована коректність його роботи в залежності від параметрів системи. Встановлено, що для чистих та суперпозиційних станів, суматор на елементах Переса коректно спрацьовує за чотири π-імпульси. Встановлено оптимальні значення параметрів обмінної взаємодії на основі аналізу функції чіткості та визначена максимальна відстань між кубітами (r = 4,56 нм), необхідна для коректної роботи суматора.

6. СПИСОК ЛІТЕРАТУРИ

[1] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information,* Cambridge University Press, NY, 2001, 676 p.

[2] A. De Vos, *Reversible Computing: Fundamentals, Quantum Computing, and Applications,* Wiley-VCH, 2010, 252 p.

[3] M. Saeedi, I. L. Markov, Synthesis and optimization of reversible circuits – a survey, *ACM Computing Surveys*, (44) 3 (2012), pp. A1-A34.

[4] G. P. Gorskyi, V. G. Deibuk, Frequency noise influence on functioning of Fredkin quantum gate, *Int. J. Computing*, (9) 2 (2010), pp. 118-126. (in Ukrainian)

[5] G. P. Gorskyi, V. G. Deibuk, Four spins model of universal quantum Fredkin gate, *Informational Technologies and Computer Engineering*, (21) 2 (2011), pp. 56-63. (in Ukrainian)

[6] J. W. Bruce, M. A. Tornton et al., Efficient adder circuits based on the conservative reversible logic gates, *Proc. IEEE Comp. Soc. Ann. Symp. on VLSI*, Pittsburgh, PA (April 25-26, 2002), pp. 83-88.

[7] G. V. Lopez, L. Lara, Numerical simulation of controlled-controlled-not (CCN) quantum gate in a chain of three interacting spins system, *J. Phys. B: At. Opt. Mol. Phys.,* (39) 9 (2006), pp. 3897-3904.

[8] B. Kane, A silicon-based nuclear spin quantum computer, *Nature*, (398) 5 (1998), pp. 133-137.

***Віталій Дейбук**, професор, доктор фізико-математичних наук, професор кафедри «Комп'ютерні системи та мережі» Чернівецького національного університету імені Юрія Федьковича. Стаж науково-педагогічної діяльності у вищій школі 30 років. Автор понад 150 наукових та науково-методичних праць, в тому числі 3 навчальних посібники з грифом Міністерства освіти та науки України.*

Наукові інтереси – комп'ютерне моделювання фізичних властивостей конденсованих систем, проблеми квантової інформатики, дослідження властивостей матеріалів для квантових комп'ютерів.

***Іван Юрійчук**, кандидат фізико-математичних наук, доцент кафедри фізики напівпровідників і наноструктур Чернівецького національного університету імені Юрія Федьковича. Стаж науково-педагогічної діяльності у вищій школі 25 років.*

Автор понад 70 наукових та науково-методичних праць. Наукові інтереси – дослідження енергетичного спектру низькорозмірних напівпровідникових квантових структур, вивчення енергетичної структури системи дефектів в напівпровідниках.

***Роман Юрійчук**, магістр кафедри «Комп'ютерні системи та мережі» Чернівецького національного університету імені Юрія Федьковича. Наукові інтереси – комп'ютерне моделювання елементів квантових комп'ютерів.*

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

SPIN MODEL OF FULL SUMMATOR ON PERES GATES

Vitaliy G. Deibuk [1)], Ivan M. Yuriychuk [2)], Roman I. Yuriychuk [1)]

[1)] Computer systems and networks department
[2)] Physics of semiconductors and nanostructures department
Chernivtsi National University, 2 Kotsubins'kogo Str., 58012, Chernivtsi, Ukraine
e-mail: v.deibuk@chnu.edu.ua, ivmykyur@gmail.com

Abstract: *Four-spin model of single-digit summator for solid state nuclear magnetic resonance (NMR) quantum computer is proposed. Correctness of summator's operation is analyzed in dependence of system parameters. It is found that the summator on Peres gates works correctly in time of four π-pulses on pure numerical and superposition states. Optimal values for exchange interactions and maximum distance between q-bits necessary for correct summator's operation are obtained from the analysis of fidelity function.*

Keywords: *quantum bit, quantum computer, Peres logical gate, spin, Ising model, fidelity function, full single-digit summator.*

1. INTRODUCTION

Development of new technologies leads to higher performance, miniaturization and complexity of computer technology and causes an increase in the number of transistors in chips and therefore growth of power consumption. Millions of gates that perform logical operations in traditional computers are irreversible. That is, every time execution of logical operations leads to the loss of some of the input information, which dissipated as heat. Successful alternative in this question can be considered recently rapidly developing in various fields the use of reverse logic [1]. The ability to use reverse logic operations that do not destroy the incoming data, in theory does not lead to energy dissipation in the system. This logic allows to reproduce input states from output states.

The development of digital reverce devices is a subject of many papers [1], but they mainly relate to circuitry implementation. Adder is one of the main blocks of computing devices. Above mantioned changes in the logic of quantum computing require corresponding changes in the implementation of adders both in logical and physical level. In previous paper [2] we analyzed physical model of quantum Fredkin gate, which is functionally complete, reversible logic element. In this work a spin model for single-digit adder on Peres gates for solid state NMR quantum computer is proposed and correctness of its work depending on system parameters is analyzed.

2. RESULTS AND DISCUSSION

We consider quantum mechanics Ising model for iteracting four spins which are subjected to a strong constant magnetic field (CMF) and a weak transverse radio frequency magnetic field (RFMF). Hamiltonian of the system may by written as [3]:

$$\hat{H} = \hat{H}_0 + \hat{W}\,, \tag{1}$$

where

$$\hat{H}_0 = -\hbar\sum_{k=0}^{3}\Big\{\omega_k I_k^z + 2J\Big[I_0^z I_1^z + I_1^z I_2^z + I_2^z I_3^z + \alpha\left(I_0^z I_2^z + I_1^z I_3^z\right) + \alpha' I_0^z I_3^z\Big]\Big\}, \tag{2}$$

$$\hat{W} = -\hbar\Omega\sum_{k=0}^{3}\left[I_k^+ \exp(i\omega t) + I_k^- \exp(-i\omega t)\right]. \tag{3}$$

Operator (1) sets forth the system of four interacting spins in magnetic field $B = \left(b\cos\omega t, -b\sin\omega t, B(z)\right)$, where b and ω – amplitude and frequency of RFMF, $B(z)$ – z-coordinate dependent CMF induction, I_k^z – projection operator of k-th spin on z-axis, $I_k^{\pm} = I_k^x \pm iI_k^y$ – so-called «descend» and «ascend» operators, $\omega_k = \gamma B(z_k)$ – Larmore's precession frequencies, $\Omega = \gamma b$ – Rabi's frequency, γ – proton gyromagnetic ratio, J – exchange interaction

constant for nearest neighbors, α and α' – relative exchange interaction for second and third neighbors correspondingly [2], $\hbar$ – Plank's constant. The model has 16 states each of which correspond to binary representation $(i_3 i_2 i_1 i_0)$, where i is equal to 0 or 1. An arbitrary state of the system can be written as a linear combination of the basis states $|k\rangle$, which are eigenvectors of the operator $\hat{H}_0$:

$$\Psi(t) = \sum_{k=1}^{16} C_k(t) |k\rangle . \quad (4)$$

Evolution of the system is described by time-dependent Schrödinger equation, which may be represented in the form:

$$\dot{D}_m = \sum_{n=1}^{16} T_{mn}(t) D_n . \quad (10)$$

where time-dependent matrix $T_{mn}(t)$ has the form:

$$T_{mn}(t) = -\frac{i}{\hbar} W_{mn}(t) \exp(i\omega_{mn} t). \quad (5)$$

Here $W_{mn}(t)$ – matrix of the perturbation operator $\hat{W}$ on the basis states |k› and $\omega_{mn} = (E_m - E_n)/\hbar$.

System (10) is a system of first order differential equations with unknown amplitudes $D_n(t)$. Numerical integration of the system is realized in MatLab program environment. For numerical simulation following parameters of the system were chosen (in units 2π MHz): $\omega_0 = 100$, $\omega_1 = 200$, $\omega_2 = 400$, $\omega_3 = 800$, $\alpha = \alpha' = 0{,}1$, $J = 5$, Ω=0,1.

Full single-digit adder consists of two Peres gates. Input register of the adder is a four-bit vector, one bit of which is equal to 0. Output vector from the first Peres gate is directed to the second Peres gate.

We study the action of the adder both on pure numerical states and also on superposition states. Two types of superposition states are considered: a uniform superposition of basis states and a superposition state where basis states 1-8 are free and the rest has probability 1/8.

For pure numerical states during so-called π-pulse probability of initial numerical state continuously decrease from 1 to 0, and probability of final state continuously increases form 0 to 1. Probability of other states remains unchanged. It is true for every pair of pure numerical states Thus, while frequency of RFMF tuned in resonance the adder on pure numerical works coorectly (probability equal to unity) during four π-pulses.

For superposition states we analyzed dependence of fidelity function [3] on distance between qubits (spins). Interaction between nuclear spins is caused by electrons exchange interaction of neighbouring qubits and is a result of electrons wave functions overlapping. Therefore the magnitude of spin interaction depends on distance between nuclear spins. Exchange interaction of well separate spins can be written as [4]:

$$4J(r) \cong 1{,}6 \frac{e^2}{\varepsilon a_B} \eta^{5/2} \exp(-2\eta), \quad (21)$$

where $\eta = r / a_B$, r – spin distance, ε – dielectric constant of the semiconductors, a_B – semiconductor Bohr radius.

From the analisys of fidelity function for single-digit adder action on uniform and non-uniform superposition states it is found maximal value of parameter η=1,52 necessary for correct action of the adder. Taking into account that Bohr radius for phosphorous ions in silicon is 3 nm maximal spin distance is 4,56 nm.

3. SUMMARY

Four-spin model of single-digit adder for solid state NMR quantum computer is proposed. Correctness of adder's operation is analyzed in dependence of system parameters. It is found that on pure numerical states the adder on Preres gates works correctly in times of four π-pulses. From the analysis of fidelity function optimal values of exchange interactions and maximal distance between qubits (r=4,56 nm) necessary for correct adder's operation are obtained.

4. REFERENCES

[1] A. De Vos, *Reversible Computing: Fundamentals, Quantum Computing, and Applications,* Wiley-VCH, 2010, 252 p.

[2] G. P. Gorskyi, V. G. Deibuk, Four spins model of universal quantum Fredkin gate, *Informational Technologies and Computer Engineering*, (21) 2 (2011), pp. 56-63. (in Ukrainian)

[3] G. V. Lopez, L. Lara, Numerical simulation of controlled-controlled-not (CCN) quantum gate in a chain of three interacting spins system, *J. Phys. B: At. Opt. Mol. Phys.,* (39) 9 (2006), pp. 3897-3904.

[4] B. Kane, A silicon-based nuclear spin quantum computer, *Nature*, (398) 5 (1998), pp. 133-137.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

ЕСТЕСТВЕННЫЕ РЕСУРСЫ И ИХ ИСПОЛЬЗОВАНИЕ ДЛЯ ПОВЫШЕНИЯ КОНТРОЛЕПРИГОДНОСТИ ЦИФРОВЫХ КОМПОНЕНТОВ СИСТЕМ КРИТИЧЕСКОГО ПРИМЕНЕНИЯ

Юлия Дрозд [1)], Александр Дрозд [1)], Юлиан Сулима [2)]

[1)] Одесский национальный политехнический университет,
пр. Шевченко, 1, Одесса, 65044, Украина,
drozd@ukr.net
[2)] Одесса, ул. Балковская, 54, mr_lemur@mail.ru

Резюме: *Рассмотрены модели, методы и средства как целевые ресурсы для решения задач проектирования и диагностирования компьютерных систем и их компонентов. Определяется критерий выбора целевых ресурсов, активирующих естественные ресурсы, направленные на повышение эффективности решения задач. Рассмотрена проблема низкой контролепригодности цифровых компонентов систем критического применения, и показаны пути ее решения при выборе целевых ресурсов в соответствии с предложенным критерием. Описан путь устранения противоречия между целевыми ресурсами, направленными на обеспечение контролепригодности, производительности и малой сложности цифровых компонентов. Этот путь основывается на распараллеливании вычислений с использованием последовательных кодов.*

Ключевые слова: *Естественные ресурсы, системы критического применения, цифровые компоненты, контролепригодность, распараллеливание вычислений в последовательных кодах.*

NATURAL RESOURCES AND THEIR USE FOR CHECKABILITY INCREASING THE DIGITAL COMPONENTS OF SAFETY-CRITICAL SYSTEMS

Julia Drozd [1)], Alexander Drozd [1)], Julian Sulima [2)]

[1)] Odessa National Politechnic University,
1, Shevchenko prospect, Odessa, 65044, Ukraine,
Drozd@ukr.net
[2)] Odessa, str. Balkovskaya, 54, mr_lemur@mail.ru

Abstract: *The models, methods and means as target resources for solving tasks for design and testing of computer systems and their components are considered. A criterion for choosing the target resources activating the natural resources for the increasing the efficiency of task solutions is determined. A problem of low checkability of digital components of safety-critical systems is examined and the ways of its solution is showed by choosing the target resources in accordance with the offered criterion. The way of elimination of a contradiction between the target resources aimed to maintain the checkability, productivity and low complexity of digital components is described. This way is based on parallelization of computations with the use of the serial codes.*

Keywords: *Natural resources, safety-critical systems, digital components, checkability, paralleling the calculations in serial codes.*

ВВЕДЕНИЕ

Для решения определенной задачи, например, проектирования, диагностики или выполнения вычислений, необходимо затратить некоторые ресурсы, получившие название целевых [1].

Целевые ресурсы (ЦР) – это привлекаемые для решения задачи модели, методы и средства в самом широком их понимании. Модели и методы

относятся к информационной части ресурсов, а средства – к технологической. Модели – это наши представления. Методы – описания некоторых преобразований. Средства – воплощение метода, включая используемый для этого инструмент. Задачи, как и ЦР можно разделить на две группы: синтеза и анализа, т.е. соответственно создания чего-либо и оценки полученного решения. Кроме ЦР в решении задачи могут использоваться естественные ресурсы (ЕР). Широко известны два их представителя: естественная информационная избыточность (ЕИИ) [2] и естественная структурно-временная избыточность (ЕВИ) [3], используемые в диагностике, т.е. при решении задач анализа. Естественная информационная избыточность проявляется в виде запрещенных значений кода результата, например, двоичного кода полного произведения.

При умножении двухразрядных двоичных чисел вычисляется 4-хразрядное произведение, принимающее значения 0, 1, 2, 3, 4, 6 и 9, т.е. только 7 значений. Остальные 9 значений 4-разрядного двоичного кода являются запрещенными и составляют ЕИИ. Естественная временная избыточность может проявляться, например, в избытке времени при формировании управляющих слов для инерционного объекта управления. Изучение ЕР является основой для их эффективного использования, направленного на упрощение ЦР и повышение качества решения задач.

Важной областью приложения ЕР являются информационные управляющие системы (ИУС) критического применения, обслуживающие объекты повышенного риска в энергетике, на транспорте, в оборонной и космической отрасли [4]. Такие системы проектируются для работы в двух режимах: штатном и критическом. Причем основное время они проводят в штатном режиме, хотя создаются ради критического режима.

Основным требованием к ИУС критического применения является обеспечение их функциональной безопасности, что традиционно решается с использованием отказоустойчивых структур [5]. Для однорежимных коммерческих систем этого достаточно.

В отказоустойчивых решениях двухрежимных ИУС сохраняется проблема низкой контролепригодности цифровых компонентов.

В данной статье анализируются вопросы определения, выявления и использования ЕР (раздел 1), почему контролепригодность является проблемой для цифровых компонентов ИУС критического применения (раздел 2) и предлагаются ЕР для повышения проблемной контролепригодности (раздел 3).

1. ЧТО ТАКОЕ ЕСТЕСТВЕННЫЙ РЕСУРС?

На примере ЕИИ и ЕВИ можно пояснить естественность ресурса тем, что он был получен в процессе решения задачи синтеза как побочный продукт и может быть использован в решении задачи анализа как особенность ЦР в виде подарка, т.е. имеет даровой характер.

Определение ЕР как особенности ЦР требует уточнения естественности в рамках решения только задачи синтеза или анализа. Получается, что ЕР, например, проектирования возникает в результате недомыслия, при котором изначально в проекте решения задачи была упущена особенность ЦР, полезная для получения результата.

Следует отметить, что деление задач на синтез и анализ приобретает все более размытый и условный характер. Оценка решения становится неотъемлемой частью самого решения, т.е. его результата. Диагностика прошла путь от выносной к встроенной, оперативной, постоянно действующей.

Деление на синтез и анализ является следствием структурного подхода (метода), вне которого невозможно решать сложные задачи.

К решению задачи предъявляются требования:

– достижения необходимой степени адекватности результата по отношению к действительности, т.е. достоверности результата;

– получения результата в срок, обеспечивая для этого необходимую производительность;

– вписывания решения в рамки предоставляемых ЦР.

Выполнение этих требований разбивает единую задачу не только на синтез и анализ, но к тому же на значительное множество подзадач со своими целевыми функциями и выделяемыми для их обеспечения ЦР, а также сетевым графиком очередности решения. Тогда особенности ЦР, использованных для решения предшествующих подзадач становятся ЕР для решения следующих подзадач.

Например, особенности обеспечения отказоустойчивости схемотехнического решения могут быть использованы для повышения его контролепригодности.

И задачи, и ЦР для их решения получают развитие и как система элементов (в части организации), и как элемент системы (в части функционирования – взаимодействия элементов).

Прежде всего, они являются элементами такой системы как наш Мир и развиваются в соответствии с его реалиями, т.е. его особенностями, иными словами – ЕР организации.

Примечательно, что особенность познается в сравнении. Поэтому нам доступно анализировать особенности организации Мира, изучая развитие его элементов, сравнивая их на различных этапах развития. Среди таких ЕР Мира выделяется его параллелизм и приближенность. В частности можно проследить этапы развития компьютерных систем и их компонентов по пути повышения уровня схемного и системного параллелизма [6], а также роста значимости обработки приближенных данных, выполняемой, как правило, в форматах с плавающей точкой [7].

Возрастающий уровень параллелизма ЦР расширяет круг решаемых задач в направлении роста их размерности, когда точные модели и методы теряют эффективность и совершенствуются в приближенные, т.е. более адекватные приближенному Миру.

Показательным является распространение метода заготовки результатов, повышающего эффективность с ростом уровня параллелизма в решаемых задачах [1].

Примером может служить современное проектирование цифровых компонентов на FPGA, когда микросхема представляет собой заготовку под множество проектов, а проект – заготовку результатов на LUT (т.е. в таблицах, зашитых в узлах памяти) для множества входных данных.

Все программное обеспечение основывается на заготовленных ветвях алгоритмов и постоянном их выборе, заготовленных подпрограммах и программных модулях.

В цифровой схемотехнике метод заготовки результатов позволяет одновременно снижать и время вычислений, и сложность решения. Этот эффект усиливается с повышением уровня параллелизма ЦР.

Например, обращение к биту памяти предполагает дешифрацию его адреса, сложность которого, может быть оценена количеством выходов дешифратора. Для разрядности адреса 10 сложность дешифратора составляет 2^{10}, т.е. 1024 выхода.

В микросхемах памяти с архитектурой 2,5 *D* адрес памяти разбивается на две части [8].

При дешифрации первой части, например, половины адреса, заготавливается множество из 2^5 результатов, из которых при дешифрации второй половины адреса выбирается один бит. Сложность такого решения складывается из 2^5 выходов двух дешифраторов и такой же по сложности схемы выбора бита, что составляет 96 выходов, т.е. более чем в 10 раз проще дешифрации всего адреса. С увеличением разрядности адреса до 16-ти и 20-ти решение упрощается более чем в 85 и 340 раз, соответственно. Одновременно растет число разрядов второй половины адреса, по которым многократно снижается время вычислений.

Как правило, время вычислений противопоставляется сложности аппаратного решения: для повышения быстродействия увеличивают затраты оборудования, причем каждое следующее повышение быстродействия обходится дороже. В рассмотренном примере, упрощение решения сопутствует повышению быстродействия, что проявляется с ростом схемного параллелизма.

Развитие ЦР происходит их структурированием под одни и те же реалии Мира, т.е. его ЕР. Идет процесс сближения ЦР, их интеграции и взаимного усиления. В конечном счете, в этом состоят их особенности, составляющие ЕР.

Например, для достижения высокой производительности компьютерных систем и их компонентов наращивается множество операционных элементов, и развиваются функции выбора результатов. Те же ЦР и в том же направлении совершенствуются для построения отказоустойчивых структур.

Таким образом, для проявления ЕР следует выбирать ЦР в направлении их естественного развития в этом Мире.

Основной вывод состоит в том, что критерием выбора ЦР для решения подзадач является их непротиворечивость. Сохранение противоречий свидетельствует о недостаточном уровне развития ЦР или упущениях в их подборе.

Весь вопрос заключается в том, чтобы плыть по течению мирового развития, получая в дар ЕР, или против течения, ухудшая результат с привлечением дополнительных ЦР.

Параллелизм и приближенность, конечно, не являются единственными ЕР Мира. Можно ожидать, что подобно тому, как параллелизм проявляет приближенность, обе особенности на определенных этапах их развития будут способствовать проявлению и актуализации других ЕР Мира.

К таким наиболее заметным особенностям Мира можно отнести его динамичность, которая проявляется в ускорении процессов развития.

В этих условиях растет актуальность проблемы скрытых процессов и ограниченных возможностей их контроля, снижающих достоверность получаемых результатов.

2. КОНТРОЛЕПРИГОДНОСТЬ ЦИФРОВЫХ КОМПОНЕНТОВ

В ИУС критического применения, которые занимают передовые позиции в развитии компьютерных и информационных технологий, проблема скрытых процессов находит отражение в контролепригодности цифровых компонентов.

Под контролепригодностью понимают свойство устройства, обуславливающее приспособленность к проведению контроля его технического состояния в процессе изготовления и эксплуатации [9].

Техническое состояние цифровых устройств оценивается с использованием методов и средств тестового диагностирования. Поэтому контролепригодность традиционно относится к этой области, где называется также тестопригодностью и оценивается по управляемости и наблюдаемости устройства [10].

В рабочем диагностировании, направленном на оценку достоверности результата вычислений, проблема ограниченной контролепригодности до сих пор проявилась лишь в теории и практике построения полностью самопроверяемых схем. В рамках этой теории было введено условие самотестируемости цифровых схем, противодействующее накоплению в них скрытых неисправностей в части контроля [11].

В ИУС критического применения возможности рабочего диагностирования существенно ограничены контролепригодностью цифровых компонентов в доступном для постоянного анализа штатном режиме.

Как правило, штатный режим ИУС характеризуется высокой стабильностью поступающих от датчиков сигналов, которые после оцифровывания дают ограниченное множество двоичных кодов. Они являются входными словами цифровых компонентов ИУС, наиболее часто – компараторов, сравнивающих их с пороговыми значениями, обозначающими различные уровни опасности.

В качестве цифровых компонентов ИУС критического применения, как правило, используются однотактные устройства, обеспечивающие высокую производительность, обрабатывая числовые данные в параллельных кодах.

Ограниченное количество входных слов в штатном режиме обуславливает низкую контролепригодность однотактных устройств, что чревато накоплением в их схемах скрытых неисправностей, проявляющихся в критическом режиме ИУС.

3. ОЦЕНКА И ПОВЫШЕНИЕ КОНТРОЛЕПРИГОДНОСТИ

Контролепригодность цифровых компонентов в тестовом и рабочем диагностировании служит разным целям, и потому должна оцениваться с разных позиций. В тестовом диагностировании контролепригодность оценивается с целью определения возможности получения тестовой входной последовательности, вычисляя в каждой точке схемы произведение ее управляемости на наблюдаемость. В рабочем диагностировании контролепригодность важна с позиции учета потенциальной угрозы со стороны скрытых неисправностей. При этом управляемость точки на рабочих входных последовательностях утрачивает свое значение, и контролепригодность полностью определяется только наблюдаемостью точек схемы.

Вместе с тем, понятие наблюдаемости целесообразно расширить с учетом возможности частичного исключения скрытых неисправностей.

Определение 1. Точка схемы называется частично наблюдаемой: 0-наблюдаемой или 1-наблюдаемой, если на множестве входных слов активируется путь от этой точки только при ее значении «0» или «1», соответственно. Если путь активируется при всех ее значениях, то точка называется наблюдаемой, а в противном случае ненаблюдаемой. Путь активируется при передаче изменения значения точки в контрольную точку схемы, подключаемую к схемам контроля цифрового компонента.

Определение 2. Точка схемы называется контролепригодной или частично контролепригодной, если является соответственно наблюдаемой или частично наблюдаемой, и неконтролепригодной в противном случае.

Контролепригодность цифрового компонента может быть оценена по следующей формуле:

$$C = (N_C + 0{,}5\ N_P) / N_T,$$

где N_C, N_P и N_T – количество контролепригодных, частично контролепригодных и всех точек схемы, соответственно.

В ИУС критического применения контролепригодность цифровых компонентов необходимо обеспечивать в комплексе с рядом других требований.

Среди них наиболее часто выделяются требования, относящиеся к производительности, отказоустойчивости и сложности решения. Для эффективного решения подзадач по выполнению этих требований обеспечивающие их ЦР следует выбирать непротиворечивыми. В распространенных решениях обеспечение контролепригодности входит в противоречие с высокой производительностью, достигаемой с использованием однотактных устройств.

Пространственный матричный параллелизм, лежащий в основе структур однотактных устройств, с одной стороны, обеспечивает их производительность путем обработки данных в параллельных кодах, а с другой стороны, ограничивает использование точек схемы

количеством входных слов. Для ИУС критического применения такое ограничение снижает контролепригодность цифровых компонентов в штатном режиме.

Для устранения данного противоречия необходимо совместить параллелизм структур цифровых компонентов, необходимый для обеспечения высокой производительности, с многоразовым использованием точек схемы для каждого входного слова. Иными словами, обработка входного слова должна выполняться в течение многих тактов, т.е. последовательно разряд за разрядом. Следовательно, распараллеливание вычислений и структур цифровых компонентов необходимо выполнять с обработкой данных не в параллельных, а последовательных кодах.

Такие решения наработаны в рамках вертикальной обработки данных [1].

Обработка параллельного кода последовательно разряд за разрядом увеличивает количество тактов выполнения операции в соответствии с разрядностью кода.

Вместе с тем, в той же мере снижается продолжительность такта, что способствует сохранению производительности при переходе от обработки параллельных кодов за один такт к многотактным вычислениям в последовательных кодах.

Возможны также компромиссные решения с последовательно-параллельной организацией вычислений. Она предусматривает разбивку числового данного на части, обрабатываемые последовательно друг за другом и параллельно по их разрядам.

Для обработки параллельных кодов наработаны многочисленные методы ускорения вычислений, например, распространения переноса в сумматорах, которые дополнительно повышают производительность однотактных устройств.

Вместе с тем, методы ускорения вычислений, включая метод заготовки результатов, могут быть использованы и при обработке данных в последовательных кодах.

Перспективными являются также параллельно-последовательные вычисления, в которых данное разделяется на параллельно обрабатываемые потоки разрядов, а также многопоточная обработка данных в последовательных кодах. В этих решениях производительность повышается в соответствии с количеством потоков.

Переход к обработке данных в последовательных кодах позволяет также упростить ЦР, которые в однотактных устройствах используются нерационально. Например, матричный умножитель n-разрядных двоичных кодов, состоящий из $n(n-1)$ операционных элементов (ОЭ), выполняет умножение за $2n-2$ задержек, вносимых этими элементами [12]. Таким образом, каждый из ОЭ задействован в такте на $(2n-2)$-ю часть его продолжительности. Матричный делитель чисел состоит из $(n+1)^2$ последовательно соединенных ОЭ [13], т.е. каждый их них используется только на $100 / (n+1)^2$%. Для $n = 8$ это составляет 1%.

Возможности повышения контролепригодности цифровых компонентов при обработке данных в последовательных кодах могут быть оценены на примере цифрового компаратора для сравнения двух 16-тиразрядных двоичных чисел A и B – входного данного и порога, соответственно.

Компаратор вычисляет результат сравнения $F = 0$ при $A < B$ (штатный режим) и $F = 1$ в противном случае (критический режим).

Цифровой компаратор построен на FPGA в пяти вариантах 1 – 5 с использованием одного, двух, четырех, восьми и 16-ти LUT, обрабатывающих данные соответственно за 16, 8, 4, 2 и 1 тактов.

Оценка контролепригодности цифрового компаратора выполнена в штатном режиме на его программной модели с получением среднего значения на заданном множестве экспериментов.

Множество экспериментов формируется перебором значений порога B (от заданного B_{MIN} до максимального $B_{MAX} = 2^{16} - 1$). С каждым значением порога B входное данное A образует заданное количество N_A слов, изменяясь от его базового значения A_B с шагом 1. Базовое значение обновляется с каждым экспериментом, изменяясь с полным перебором от начального значения A_{MIN} до значения $B - N_A$

В табл. 1 приведены результаты моделирования в процентах, полученные для схем 1 – 5 при $B_{MIN} = 511$, $A_{MIN} = 10$ и различных значениях N_A.

Таблица 1. Контролепригодность компаратора

N_A	1	2	3	4	5
1	100	85	69	44	26
5	100	85	84	56	34
10	100	85	84	64	38
20	100	100	92	68	42
30	100	100	92	72	42
40	100	100	92	72	42
50	100	100	92	76	46
100	100	100	94	80	48
200	100	100	94	84	53
300	100	100	95	84	57
400	100	100	95	84	57
500	100	100	95	88	59

Для сравнения в табл. 2 приведены значения контролепригодности, полученные без учета частично контролепригодных точек.

Таблица 2. Контролепригодность компаратора без учета частично контролепригодных точек

N_A	1	2	3	4	5
1	100	71	46	12	0
5	100	85	61	28	6
10	100	85	76	36	10
20	100	100	84	44	14
30	100	100	84	48	16
40	100	100	84	48	16
50	100	100	84	52	20
100	100	100	84	60	22
200	100	100	92	68	26
300	100	100	92	72	30
400	100	100	93	72	30
500	100	100	93	72	32

Сравнительный анализ данных, приведенных в таблицах, показывает на рост контролепригодности с увеличением количества слов и с понижением номера схемы, т.е. переходом от обработки параллельных кодов к вычислениям в последовательных кодах.

Параллельный компаратор (схема 5) демонстрирует низкую контролепригодность даже при значительном количестве входных слов. К тому же в основном она достигается за счет частично контролепригодных точек.

Схемы 3 и 4 достигают высоких значений контролепригодности 84% и 60% при использовании соответственно 20-ти и ста входных слов. Компаратор, выполняющий операцию сравнения за 8 тактов (схема 2), достигает 100% контролепригодности на множестве из 20-ти входных слов

Контролепригодность последовательного компаратора (схема 1) составляет 100%, начиная с одного входного слова.

Таким образом, обработка данных в последовательных кодах решает проблему низкой контролепригодности современных цифровых компонентов в составе ИУС критического применения. Причем такое решение не противоречит достижению высокой производительности.

4. ВЫВОДЫ

Важную роль в решении задач проектирования и диагностики компьютерных систем и их компонентов играет выбор ЦР, обеспечивающих получение результатов с заданными целевыми функциями. Развитие ЦР – моделей, методов и средств – демонстрирует их структурирование под реалии Мира, среди которых выделяются параллелизм и приближенность. Критерием выбора ЦР является их непротиворечивость при решении подзадач по обеспечению требуемых целевых функций. При этом включаются ЕР, т.е. особенности ЦР, проявляющиеся в интеграции ЦР и их взаимном усилении в направлении обеспечения целевых функций подзадач.

Примером может служить решение проблемы низкой контролепригодности цифровых компонентов ИУС критического применения. Противоречие между контролепригодностью и производительностью современных устройств снимается путем распараллеливания вычислений в последовательных, а не традиционно используемых параллельных кодах.

5. СПИСОК ЛИТЕРАТУРЫ

[1] *On Line Testing of the Safe Instrumentation and Control Systems*, A. Drozd and V. Kharchenko (Eds.), National Aerospace University by N.E. Zhukovsky “KhAI”, 2012, 614 p. (in Russian)

[2] J.G. Savchenko, *Fault Tolerant Digital Devices*, Moscow, Sovetskoye radio, 1977, 176 p. (in Russian)

[3] A.M. Romankevich, V.N. Valuyski, V.A. Ostafin, *Structural-Time Redundancy in Control Circuits*, Kyiv, “High school”, 1979, 160 p. (in Russian)

[4] M.A. Yastrebenetsky (ed.), *NPP I&Cs: Problems of Safety*, Kyiv, Technika, 2004, 472 p. (translated in USA by NPC, 2007).

[5] A.A. Siora, V.A. Krasnobayev, V.S. *Kharchenko, Fault Tolerant Systems with Version-Information Redundancy*, National Aerospace University “KhAI, 2009, 321 p. (in Russian)

[6] E. Tanenbaum, *Architecture of computer*, *4th ed.*, SPb, Piter, 2003, 698 p. (in Russian)

[7] D. Goldberg, What every computer scientist should know about floating-point arithmetic, *ACM Computer Surveys*, (23) 1 (1991), pp. 5-18.

[8] E.P. Ugryumov, *Digital Schemotechnics, 2nd edition*, SPb., BHV-Petersburg, 2004, 800 p. (in Russian)

[9] N.S. Scherbakov, *Reliability of Digital Devices Work*, Moscow, Manufacturing Engineering, 1989, 224 p. (in Russian)

[10] R.J. Bennets, *Design of Checkable Logical Schemes*, Moscow, Radio and Communication, 1995, 180 p. (in Russian)

[11] D.A. Anderson, G. Metze, Design of totally self-checking check circuits for m-out-of-n codes, IEEE Trans. Comput, (C-22) 3 (1977), pp. 263-269.

[12] A.O. Melnyk, *Computer Architecture*, Lutsk: Volyn regional printing house, 2008, 470 p. (in Ukrainian)
[13] K.G. Samofalov, A.M. Romankevich, V.N. Valuisky, Y.S. Kanevsky, M.M. Pinevich, *Applied Theory of Digital Automats*, Kyiv, "High school", 1987, 375 p. (in Russian)

***Юлия Дрозд**, доцент, к.т.н., доцент кафедры информационных систем Одесского национального политехнического университета.*

Научные интересы – вертикальная обработка данных, естественные ресурсы проектирования и диагностирования компьютерных систем.

***Александр Дрозд**, профессор, д.т.н., профессор кафедры компьютерных интеллектуальных систем и сетей Одесского национального политехнического ун-та.*

Научные интересы – проектирование и диагностирование встроенных систем критического применения.

***Юлиан Сулима**, аспирант кафедры компьютерных интеллектуальных систем и сетей Одесского национального политехнического ун-та.*

Научные интересы – контролепригодность цифровых компонентов систем критического применения.

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

NATURAL RESOURCES AND THEIR USE FOR CHECKABILITY INCREASING THE DIGITAL COMPONENTS OF SAFETY-CRITICAL SYSTEMS

Julia Drozd [1], Alexander Drozd [1], Julian Sulima [2]

[1] Odessa National Politechnic University,
1, Shevchenko prospect, Odessa, 65044, Ukraine,
Drozd@ukr.net
[2] Odessa, str. Balkovskaya, 54, mr_lemur@mail.ru

Abstract: *The models, methods and means as target resources for solving tasks for design and testing of computer systems and their components are considered. A criterion for choosing the target resources activating the natural resources for the increasing the efficiency of task solutions is determined. A problem of low checkability of digital components of safety-critical systems is examined and the ways of its solution is showed by choosing the target resources in accordance with the offered criterion. The way of elimination of a contradiction between the target resources aimed to maintain the checkability, productivity and low complexity of digital components is described. This way is based on parallelization of computations with the use of the serial codes.*

Keywords: *Natural resources, safety-critical systems, digital components, checkability, paralleling the calculations in serial codes.*

In order to solve a task (for example, in co-design or testing) the target resources (TR): the models, methods and means are used. The models and methods are related to information part of TR and the means belong to their technology one. The TR can be applied in the tasks of synthesis and analysis [1].

Except the natural resources (NR) can be used in solution of a task. Two representatives of NR are wide known: natural information redundancy and natural structural-time redundancy [2, 3].

These kinds of natural redundancy are formed during process of task solution as indirect product and they can be used for solving a task of analysis as particularities of TR in form of free present. The NR are not limited to natural information and structural time redundancy as they belong to diverse set of TR particularities.

The task is solved taking into account the requirements to productivity, reliability and limited resources. Execution of these requirements divides the task into large set of subtasks with their target functions, the TR appropriated for maintaining the target functions and schedule of execution priority. Particularities of TR used for solving the previous subtasks become NR to solve the following ones.

The tasks and TR for their solving receive development as a system of elements and as an element of system.

First of all they are the elements such system like our Universe and are developed in accordance with its particularities, i.e. its NR.

Among such NR the parallelism and proximity of the Universe are allocated.

In particular we can retrace the stages of development if the computer systems and their digital components by the way of increasing the parallelism level and progress in approximate data processing executed as a rule in floating-point format [4, 5].

Growing level of TR parallelism expands a set of solved tasks in direction of raising their dimensions when exact models and methods lose efficiency and improve in approximate ones, i.e. more adequate to approximate Universe.

We can consider this issue examining the process of expansion of the preserved results method which increases efficiency with raising the parallelism level of solved tasks [1].

The modern co-design of the digital components on FPGA can be considered as an example of this method application. A chip is a blank for a set of circuit projects, and every project is a blank of results reserved on LUT for a set of input data.

The software is based on reserved program modules, branches of the algorithms and on permanent process of their choice.

In digital component design the method of preserved results allows simultaneously to reduce time operation and hardware overhead.

Development of TR passes by their structuring under the same particularities of Universe. This process determines narrowing of the TR, integration and mutual amplification. These particularities of TR are NR of design and testing.

In order to activation of NR it is necessary to choose the TR in direction of their natural development in this Universe. The main issue determines a criterion of TR choice for solving the subtasks: these TR should be consistent. Conservation of contradictions shows low level of the TR development or overlook in their choice.

All content of this question is to drift with the stream of Universe development receiving NR like gift or up stream degrading the result with use of additional TR.

The following particularity of Universe is its dynamics which is demonstrated in acceleration of development processes. In these conditions the actuality of problem of the latent processes and limited opportunities of their supervision reducing reliability of received results.

In safety-critical Instrumentation & Control Systems (I&CS) the problem of latent processes is shown in checkability of the digital components [1].

The I&CS are designed for operation in two modes: normal and emergency one. For most of operating time, the safety-critical I&CS runs in the normal mode. The emergency one, for which the safety-critical I&CS is designed, is a rare event as a rule and at best way never occur [6].

The main requirement to safety-critical I&CS is the ensuring and maintaining their functional safety which is traditionally solved by use development of the fault-tolerant structures [7].

However they do not guarantee calculation of reliable results owing to low checkability of digital components.

The checkability of a digital component in on-line testing can be estimated by amount of circuit points which are observable in a normal mode.

In safety-critical I&CS it is necessary to provide the checkability of the digital components on the limited set of input words in a normal mode in a complex with other requirements such as productivity, fault tolerance and low complexity.

In traditional solutions the maintaining of checkability and low complexity contradicts the high productivity achievable with use of simultaneous digital components.

The contradictions between the checkability and productivity, productivity and complexity are eliminated by paralleling the calculations in series codes.

Data processing in series codes multiple simplifies TR which in simultaneous digital components are used irrationally.

For example, operational elements of a 8-bit iterative array divider of codewords are used in clock cycle only on 1% of time.

Opportunities of increase in checkability of digital components using data processing in series codes are appreciated by the example of the 16-bit digital comparator for comparison of input number A with a threshold B.

The comparator calculates result of comparison $F = 0$ for $A < B$ (a normal mode) and $F = 1$ otherwise (an emergency mode).

The digital comparator is constructed on FPGA in five variants with use of one, two, four, eight and sixteen LUT, processing the data accordingly for 16, 8, 4, 2 and 1 clock cycles.

The estimation of the comparator checkability is executed in a normal mode on program model with reception of average value on set of values of threshold B. For each of them operand A forms the set of the input words located successively.

Simulation of digital comparator demonstrates low checkability in case of parallel code processing and 100% of checkability beginning of one input word at execution of calculations in series codes.

REFERENCES

[1] *On Line Testing of the Safe Instrumentation and Control Systems*, A. Drozd and V. Kharchenko (Eds.), National Aerospace University by N.E. Zhukovsky "KhAI", 2012, 614 p. (in Russian)

[2] J.G. Savchenko, *Fault Tolerant Digital Devices*, Moscow, Sovetskoye radio, 1977, 176 p. (in Russian)

[3] A.M. Romankevich, V.N. Valuyski, V.A. Ostafin, *Structural-Time Redundancy in Control Circuits*, Kyiv, "High school", 1979, 160 p. (in Russian)

[4] E. Tanenbaum, *Architecture of computer*, *4th ed.*, SPb, Piter, 2003, 698 p. (in Russian)

[5] D. Goldberg, What every computer scientist should know about floating-point arithmetic, *ACM Computer Surveys*, (23) 1 (1991), pp. 5-18.

[6] M.A. Yastrebenetsky (ed.), *NPP I&Cs: Problems of Safety*, Kyiv, Technika, 2004, 472 p. (translated in USA by NPC, 2007).

[7] A.A. Siora, V.A. Krasnobayev, V.S. *Kharchenko, Fault Tolerant Systems with Version-Information Redundancy*, National Aerospace University "KhAI, 2009, 321 p. (in Russian)

computing@computingonline.net
www.computingonline.net

ISSN 1727-6209
International Journal of Computing

THE GENERALIZED CONSTRUCTION OF PSEUDONONDETERMINISTIC HASHING

Volodymyr A. Luzhetsky, Yuriy V. Baryshev

Information protection department of Vinnytsia National Technical University,
Khmelnytske shosse, 95, Vinnytsia, Ukraine,
yuriy.baryshev@gmail.com

Abstract: *This article is devoted to the development of hash constructions, which are based on the pseudonondeterministic hash conception. The conception allows to design hash functions with improved infeasibility to cryptanalysis. Both proposed constructions and known ones are generalized as pseudonondeterministic constructions.*

Keywords: *hashing, multipipe, the pseudonondeterminancy,the automaton, the cryptography.*

1. INTRODUCTION

The requirement of the cryptographic algorithms publicity is brought forward by the cryptographic tools market. It is caused by the customer's desire of being able to insure these tools efficiency before they buy them. At once it is known that the private cryptographic algorithm is more difficult to break than the public one [1]. If algorithm infeasibility is proven theoretically, then this publicity doesn't cause problems, but algorithms of this type seldom gain much popularity and spreading within practical implementations, because they are based on operations, that are unnatural for the computing machinery. The publicity of cryptographic algorithms, infeasibility of which isn't theoretically proven, let talk about their practical infeasibility in the case these algorithms weren't broken by known methods. Still this kind of infeasibility couldn't be guaranteed in the future, when new methods of breaking appear. Thereafter the cryptanalytic's knowledge of round transformation, which infeasibility is proven only practically, makes it easier to study and break the hash function. That's why the development of hash functions, which would provide both the publicity of algorithms, those implement it, and round transformation nondeterminancy to the intruder during certain round cryptanalysis is topical.

The goal of the research is improving of hash infeasibility by public hashing approaches development, which would provide round transformation nondeterminancy to a cryptanalytic, and their implementation by the generalized hash construction.

The following tasks are to be solved to achieve the goal:

- known hash approaches analysis;
- the development of the new hash conception, which is based on the nondeterministic automaton;
- the development of the generalized hash construction which is based on the pseudonondeterministic hash conception.

2. ANALYSIS OF MODERN HASH APPROACHES

The majority of modern hash approaches reaches by their roots the Merkle-Damgaard construction [2], which became the classical one. In correspondence with the construction hashing is to be meant process of gaining fixed length hash digit according to the arbitrary length message, which is spitted up to the l data blocks of the same length. The initialization vector h_0 usually is used as a key. Thus the hash iteration has the following view [2]:

$$h_i = f\left(h_{i-1}, m_i\right), \qquad (1)$$

where h_i – the intermediate hash digit, that is obtained after the i-th data block processing;

m_i – the i-th data block ($i = \overline{1, l}$);

$f(\cdot)$ – the reduction function, which provides fixed length of the output value.

The hash digit of the whole message would be h_l [2]. A lot of other hash constructions are known, which differ from the Merkle-Damgaard

construction, but all of them stipulate fixed sequence of actions, that are to be performed for hash digit computing. It allows cryptanalytics to analyze algorithms and find weaknesses in hashing, which appeared at the constructions level.

Several works were performed to avoid these weaknesses within hashing [3-5]. The work [3] is devoted to the hash construction development, that contains two nonstandard hash parameters and could be formalized in this way:

$$h_i = f\left(h_{i-1}, m_i, \#bits, c\right), \quad (2)$$

where $\#bits$ – the quantity of data hashed so far;
c – the certain constant – the cryptographic salt.

According to the work [3] the constant c in the hash construction (2) is new for an each hash process. Thus the construction has certain parameter, which is to be chosen at random. That's why it is more difficult for the cryptanalytic to prepare attack beforehand, when the parameter's value is unknown. Nevertheless, the hash algorithm, because of being known, could be an object of cryptanalytic's research and known methods of breaking or their modifications could be implemented. Moreover, the approach reduces hashing rapidity comparatively with the construction (1), because the additional data is to be processed.

The work [4] contains the brief description of two hash functions Dynamic SHA and Dynamic SHA-2, the peculiarity of which is using of the data driven shift. Also cryptanalysis of these functions is performed at the work [4] and the attack, that allows to break them, is proposed. The attack could succeed through fixed structure of hash algorithms and their publicity.

The work [5] is devoted to the hashing, based on data driven primitives: permutations and substitutions. These primitives have two types of inputs: one – for a data to be hashed, other – for driven data, which value determines the type of the substitution or the permutation is to be performed under the data from the former input at each round. The driven hash primitives, which manipulate input data of a small width (2-3 bits of input/output data and 1-2 bits of driven data), are considered at the work [5]. These primitives are cascaded in order to implement more complicated elements. The data permutation is used for correspondence maintaining between different bits of a data block and for arbitrary dependence gaining between each input data bit and output bits. Thereby authors of the work [5] use several reduction functions, each of them is chosen at the certain round depending on the driven vector value:

$$h_i = f_{v_i}\left(h_{i-1}, m_i\right), \quad (3)$$

where v_i – the driven vector, which value is computed using the following function:

$$v_i = g\left(v_{i-1}, k, m_i\right), \quad (4)$$

where k – the key data.

The peculiarity of the hash construction (3) is composed of forming driven vector using both the key and input data, that allows authors of the work [4] to achieve nonlinearity of functions based on these primitives usage. Notwithstanding the large amount of developed substitution-permutation networks and the generalized methodology of their developing, the work [4] has a weakness – it also could be an object of the cryptanalysis, because the driven vector is formed with the participation of the input data. In the other hand there wouldn't be much correlation between the input and output data of the reduction function without the participation and this is rather unwanted for cryptographic hash functions [1, 2].

3. THE CONCEPTION OF PSEUDONONDETERMINISTIC HASHING

The conception, which is based on the pseudonondeterministic approach [6], was proposed to avoid weaknesses of modern approaches of round transformation hiding from cryptanalytics, which was described supra. The hash algorithm is considered as an actions sequence, that is performed by the automaton, states of which are intermediate hash values $\{h_i\}$, and input signals are data blocks $\{m_i\}$, those are to be hashed. Thereafter the set of all possible data blocks is the alphabet of the automaton, and all possible sets of data blocks – rows of the alphabet.

An automaton is deterministic one if it proceeds the one and only one state from any current state for any input symbol of its alphabet [7]. A deterministic automaton is described by the set $\{\mathbf{S}, \mathbf{A}, \delta, s_0, \mathbf{D}\}$, where $\mathbf{S}$ – the set of the automaton's states; $\mathbf{A}$ – the alphabet of input symbols; δ – the function, that implements the mapping $\mathbf{S} \times A \to \mathbf{S}$, s_0 – the initial state of the automaton ($s_0 \in \mathbf{S}$); $\mathbf{D}$ – the subset of the set $\mathbf{S}$, which is called set of final states [8]. Thus an automaton, that performs deterministic hashing could be described as the set $\{\mathbf{H}, \mathbf{M}, f(\cdot), h_0, h_l\}$, where $\mathbf{H}$ – the set of all intermediate hash values; $\mathbf{M}$ – the set of all data blocks values; h_l – the final

hash state, which is obtained after finishing of the last *l*th data block of the input message processing, $h_l \in \mathbf{H}$.

It is obvious, that the infeasibility of modern key hash algorithms is based on the point, that an intruder doesn't know the initial state h_0 of the deterministic automaton, gaining which the intruder breaks the hash function. The intruder deals with the deterministic automaton, while trying to break hash algorithms, those don't use a key, the main difficulty in the case concerns finding a path of the same length as the original message l $\{h'_1, h'_2, ..., h'_l\}$, (where $h'_i = f(h'_{i-1}, m'_i)$, and $h'_0 = h_0$), which has the same final state $h'_l = h_l$. The situation, that hash breaking task adds up to this one is caused by the determinancy of the automaton – each instant automaton's state (h_i, m_i) causes unambiguous automaton proceeding into the next state h_{i+1}. This feature of hashing makes effective the set of attacks, which use precomputation and are considered in particular at the work [3].

A nondeterministic automaton is the one, the proceeding rule of which isn't obligatory described by a function. The automaton could proceed several different states, when it stands at the certain state s_i and the same input symbol is processed [7]. Let ε be an empty message (zero-length), then a nondeterministic automaton is described by the set $\{\mathbf{S}, \mathbf{A}, \delta', s_0, \mathbf{D}\}$, which is analogical to the one of a deterministic automaton, where δ' is mapping $\mathbf{S} \times (A \cup \{\varepsilon\}) \to \mathbf{S}$ [8]. It is obviously, that deterministic automaton is the special case of the nondeterministic one.

The task of hash breaking could be more difficult if hashing process would be described by the nondeterministic automaton model. In the case it is more difficult to an intruder to find collisions, because he doesn't have the information about automaton proceeding to the next state, while he knows the instant state (h_i, m_i). It is clear that the nondeterministic automaton couldn't be used for the hash digit computation in the pure state (as is), because the same message would have several hash digits in this case, that contradicts the definition and the purpose of the hashing process. That's why the methods of pseudonondeterministic hashing are proposed. The hashing process is described for the intruder by the model similar to the nondeterministic automaton model according to these methods. So the automaton, which implements pseudonondeterministic hashing is describe by the set $\{\mathbf{H}, \mathbf{M}, \mathbf{F}, h_0, h_l\}$, where $\mathbf{F}$ – the set of functions $\{f_{v_i}(\cdot)\}$, which could be performed by the automaton at the each round depending on the driven vector's value v_i. The vector should be hidden from the intruder in the case of keyed hashing, and it should significantly depend on the input data in the case of unkeyed hashing.

4. PSEUDONONDETERMINISTIC HASH CONSTRUCTIONS

The construction (3) could be similar to the Merkle-Damgaard construction for pseudonondeterministic hashing, because it anticipates several reduction functions which are chosen according to the driven vector's value. However it is needed to exclude driven vector depending on its previous value and the input data from the formula (4). Instead, it is proposed to generate driven vector using intermediate hash values to keep the dependence on the input data. So the vector forming function has the following view:

$$v_i = g(h_{i-1}). \tag{5}$$

As both the reduction function and the vector forming function at the construction, which is described by formulas (3) and (5), consider output receiving depending on the intermediate hash value received at the previous iteration, that's why it is impossible for the hashing process based on the construction (3), (5) to achieve pure pseudonondeterminancy, because the same round transformation would be always performed for the certain intermediate hash value. It is proposed to avoid the undesirable dependence by vector forming and the hash value computation depending on different intermediate hash values. In the simplest case the hash construction is the following:

$$\begin{cases} h_i = f(h_{i-1}, m_i); \\ v_i = g(h_{i-2}). \end{cases} \tag{6}$$

The construction (6) is generalized in the following way:

$$\begin{cases} h_i = f_{v_i}(\mathbf{H}_i^*, m_i); \\ v_i = g(\mathbf{H}_i^{**}), \end{cases} \tag{7}$$

where $\mathbf{H}_i^*$, $\mathbf{H}_i^{**} \subset \{h_0, h_1, ..., h_{i-1}\}$ and $\mathbf{H}_i^* \cap \mathbf{H}_i^{**} = \varnothing$.

The construction (7) is inconvenient from the practical point of view, because in the case of keyed

hashing the key should have the length exceeded the output hash value's one. That's why authors find it more perspective to implement pseudonondeterministic hashing for the multipiped case of hashing. The instance of the proposed approach implementation for the double-pipe hashing is the following:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(h_{i-1}^{(1)}, m_i\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(h_{i-1}^{(2)}, m_i\right); \\ v_i^{(1)} = g^{(1)}\left(h_{i-1}^{(2)}\right); \\ v_i^{(2)} = g^{(2)}\left(h_{i-1}^{(1)}\right), \end{cases} \quad (8)$$

where $h_i^{(j)}$ – the intermediate hash value, which is received at the *j*th pipe at the *i*th iteration.

The generalization of the construction (8) for the arbitrary number of pipes q $(q \geq 2)$ is the following:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_i\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_i\right); \\ \ldots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_i\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \ldots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right), \end{cases} \quad (9)$$

where $\mathbf{H}_i^{*(j)}$, $\mathbf{H}_i^{**(j)} \subset \left\{h_0^{(1)}, h_1^{(1)}, \ldots., h_{i-1}^{(1)}, h_0^{(2)}, h_1^{(2)}, \ldots., h_{i-1}^{(q)}\right\}$ and $\mathbf{H}_i^{*(j)} \cap \mathbf{H}_i^{**(j)} = \varnothing$, $j = \overline{1, q}$.

The construction (9), as most known hash constructions [2, 3, 5, 9-12], considers consequent data block processing – one block is processed at the one iteration. The approach simplifies the hash methods implementation, but simultaneously it facilitates to the cryptanalytic the task of the breaking automaton design. That's why it is proposed to use several data blocks at the each iteration. For instance, the following construction is proposed to "bond" the data block to their positions at the message, that would complicate the part of the message replacement task for the cryptanalytic:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_i, m_{i-rand(m_i)}\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_i, m_{i-rand(m_i)}\right); \\ \ldots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_i, m_{i-rand(m_i)}\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \ldots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right), \end{cases} \quad (10)$$

where $rand(\cdot)$ – the function of random number sequence generating.

Using of the construction (10) considers processing of two data blocks at the each iteration, but in the general case all data blocks could be processed in the following way:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_1, m_2, \ldots, m_l\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_1, m_2, \ldots, m_l\right); \\ \ldots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_1, m_2, \ldots, m_l\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \ldots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right). \end{cases} \quad (11)$$

The cryptographic salt (a random number) could be used as the argument of the reduction function $f(\cdot)$ in addition to data blocks and intermediate hash values, which are used at the construction (11), as it was proposed at the work [3]. However it was proposed to use the same salt value at the each iteration at the work [3], so the cryptographic salt is the static one. The static cryptographic salt is generalized by the dynamic one, so it is proposed to input new pseudorandom number at the each iteration [13, 14]:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_1, m_2, \dots, m_l, r_i\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_1, m_2, \dots, m_l, r_i\right); \\ \dots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_1, m_2, \dots, m_l, r_i\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \dots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right); \\ r_i = rand\left(r_{i-1}\right), \end{cases} \quad (12)$$

where r_i – the dynamic cryptographic salt and r_0 is determined before beginning of the hash process in the same way as the static salt is determined.

The dynamic cryptographic salt is the same for each pipe at the construction (12), but it could be different for the each pipe in the general case:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_1, m_2, \dots, m_l, r_i^{(1)}\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_1, m_2, \dots, m_l, r_i^{(2)}\right); \\ \dots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_1, m_2, \dots, m_l, r_i^{(q)}\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \dots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right); \\ r_i^{(1)} = rand^{(1)}\left(r_{i-1}^{(1)}\right); \\ r_i^{(2)} = rand^{(2)}\left(r_{i-1}^{(2)}\right); \\ \dots \\ r_i^{(q)} = rand^{(q)}\left(r_{i-1}^{(q)}\right). \end{cases} \quad (13)$$

The dynamic cryptographic salt could be used several times at the each iteration similar to the data blocks at the construction (11). Consequently the construction (13) is generalized by the following construction:

$$\begin{cases} h_i^{(1)} = f_{v_i^{(1)}}\left(\mathbf{H}_i^{*(1)}, m_1, m_2, \dots, m_l, \mathbf{R}^{(1)}\right); \\ h_i^{(2)} = f_{v_i^{(2)}}\left(\mathbf{H}_i^{*(2)}, m_1, m_2, \dots, m_l, \mathbf{R}^{(2)}\right); \\ \dots \\ h_i^{(q)} = f_{v_i^{(q)}}\left(\mathbf{H}_i^{*(q)}, m_1, m_2, \dots, m_l, \mathbf{R}^{(q)}\right); \\ v_i^{(1)} = g^{(1)}\left(\mathbf{H}_i^{**(1)}\right); \\ v_i^{(2)} = g^{(2)}\left(\mathbf{H}_i^{**(2)}\right); \\ \dots \\ v_i^{(q)} = g^{(q)}\left(\mathbf{H}_i^{**(q)}\right), \end{cases} \quad (14)$$

where $\mathbf{R}^{(j)}$ – the set of pseudorandom numbers, which is used at the *j*th hashing pipe.

Let us use the following parametric description of hashing, which is based on the construction (14) $PNDH_q\left(k; d; z; \gamma; \phi\right)$ (PseudoNonDeterministic Hashing), where the following denotations are used:

q – the quantity of hashing pipes;

k – the quantity of intermediate hash values, which are arguments of the reduction function $f^{(j)}(\cdot)$ at the *j*th $\left(j = \overline{1, q}\right)$ pipe;

d – the quantity of data blocks, which are processed for the intermediate hash values computation at the *j*th pipe $\left(d = \overline{1, l}\right)$;

z – the quantity of pseudorandom numbers, which are used at the each iteration at the one pipe;

γ – the correlation between the intermediate hash value width and the output hash digit width;

ϕ – the quantity of intermediate hash values, which are used to form the driven vector at the *j*th pipe.

The construction (14) was received by the sequential generalization of hash constructions, that's why it generalized all constructions considered supra. Parametric descriptions of the known hash constructions is presented at the table 1 to prove, that the construction (14) is generalized for other constructions and deterministic hashing is particular case of the pseudonondeterministic one.

Table 1. Parametric descriptions of known hash constructions

The hash construction	The parametric description
Merkle-Damgaard [2]	$PNDH_1(1; 1; 0; 1; 0)$
Cascading [2]	$PNDH_q(1; 1; 0; 1; 0)$
HAIFA [3]	$PNDH_1(1; 1; 1; 1; 0)$
Hirose [10]	$PNDH_2(2; 1; 1; 2; 0)$
Double-pipe [11]	$PNDH_2(2; 1; 0; 2; 0)$
Wilde pipe [11]	$PNDH_1(1; 1; 0; 2; 0)$
3C [9]	$PNDH_2(2/1; 0/1; 0; 2; 0)$
Sponge [12]	$PNDH_1(1; 1; 0; \gamma; 0)$

Parameters k and d at the table 1 for the construction 3C are described using slash because of different arguments of pipes reduction functions. For the first pipe they are the data block and the intermediate hash value from the pipe, for the second pipe – intermediate hash values from both pipes.

5. AN EXAMPLE OF PSEUDONONDETERMINISTIC HASH FUNCTION IMPLEMENTATION

Several hash functions were developed to implement hash constructions described above [15]. These hash functions are based on well-known non-linear functions of three arguments x_1, x_2, x_3 (in particular first pair of them is used at the hashing standard SHA-2 [16]):

$$y = (x_1 \wedge x_2) \oplus (\neg x_1 \wedge x_3), \quad (15)$$

$$y = (x_1 \wedge x_2) \oplus (x_1 \wedge x_3) \oplus (x_2 \wedge x_3), \quad (16)$$

$$y = (x_1 \vee x_2) \oplus (\neg x_1 \vee x_3), \quad (17)$$

$$y = (x_1 \vee x_2) \oplus (x_1 \vee x_3) \oplus (x_2 \vee x_3). \quad (18)$$

Each iteration of hashing is supposed to use one of functions (15-18) according to the driven vector value. Also it is proposed to use circular shift operation for each argument before it would be used by one of these non-linear functions.

For instance the hash function was developed, which implements hashing $PNDH_8(5;1;0;1;3)$ with output hash digest length of 256 bits. At the start of each iteration the driven vector value is computed according to the function:

$$v_i^{(j)} = h_{i-1}^{((j+1) \bmod q)} \oplus h_{i-1}^{((j+4) \bmod q)} \oplus h_{i-1}^{((j+7) \bmod q)}. \quad (19)$$

Then the driven vector $v_i^{(j)}$ is divided into seven parts: one part of 2 bits length c and six parts of 5 bits length $s_1, s_2, ..., s_6$. The reduction function is chosen according to the c value:

- "11" – the function (15);
- "10" – the function (16);
- "01" – the function (17);
- "00" – the function (18).

At each iteration reduction function arguments m_i, $h_{i-1}^{(j)}$, $h_{i-1}^{((j+2) \bmod q)}$, $h_{i-1}^{((j+3) \bmod q)}$, $h_{i-1}^{((j+5) \bmod q)}$, $h_{i-1}^{((j+6) \bmod q)}$ are circular shifted rightwards to $s_1, s_2, s_3, s_4, s_5, s_6$ bits respectively. For example if the driven vector value is 0x01234567 it would be divided in binary form into codes: 00 00000 10010 00110 10001 01011 00111. The following reduction function $f_i^{(j)}(\cdot)$ is to be performed at this iteration:

$$\begin{aligned} h_i^{(j)} = {} & \left(m_i \wedge \left|h_{i-1}^{(j)}\right|_{>>>18}\right) \oplus \\ & \oplus \left(\neg m_i \wedge \left|h_{i-1}^{((j+2) \bmod q)}\right|_{>>>6}\right) \oplus \\ & \oplus \left(\left|h_{i-1}^{((j+3) \bmod q)}\right|_{>>>17} \wedge \left|h_{i-1}^{((j+5) \bmod q)}\right|_{>>>11}\right) \oplus \\ & \oplus \left(\left|\neg h_{i-1}^{((j+3) \bmod q)}\right|_{>>>17} \wedge \left|h_{i-1}^{((j+6) \bmod q)}\right|_{>>>7}\right). \end{aligned} \quad (20)$$

The hash function was tested by Known answer tests, which were used by NIST (USA) for SHA-3 competitors [15, 17]. For instance, the results of extremely long message testing are the following:

Repeat = 163840
Text = abcdefghbcdefghicdefghijdefghijkefghijklfgh ijklmghijklmnhijklmno
MD = 611D589256CD92FCE44225A445B6DC1 729F953851F4DAE07D204EDC8FAF40AE9

The short message test results are the following:

Len = 0
Msg = 00
MD = C4377DE5D712E89EFF4AB66FB2D635FE3 43A6B9869C06295468D483199D4AFC8

Len = 1
Msg = 00
MD = D89E7E9FD569456A989333B4B3A3B3AB E882E038E3C3F4FE690808596F16ACDE

Len = 2
Msg = C0
MD = 66B816FF6E5EC2D52A78B31BE896D980 76EBB323A55F878139F333B8ACDF88A2

Len = 3
Msg = C0
MD = 54D4B72538B7A5D72736EE8C53A18141 7E343B34A9B134FC605E5D607CF12447

…

Results of testing showed that the hash function doesn't degenerate while extremely long or short messages are processed.

6. CONCLUSIONS

The performed analysis of known hash approaches showed, that they could be described by the deterministic automaton model, moreover it is the very feature used by cryptanalytics for many attacks performing. That's why the conception of pseudonondeterministic hashing is proposed by

authors. The conception is generalization of the deterministic hashing one. While hash algorithms are developed according to the conception, an intruder is unavailable to perform cryptanalysis of round transformations, because he doesn't know operations, which are performed to process the certain data block. Moreover the approach is proposed, which allows to hide the information about the data block being processed at the certain iteration from an intruder. The set of hash constructions is proposed to implement the pseudonondeterministic hash conception. These constructions provide hash infeasibility improving towards known attacks.

7. REFERENCES

[1] S. Burnett, S. Paine, *RSA Security's Official Guide to Cryptography*, Binom-press, Moscow, 2002, p. 384. (in Russian)

[2] B. Preneel, *Analysis and Design of Cryptographic Hash Functions,* Katholieke Universiteit Leuven, 1993, p. 323. http://homes.esat.kuleuven.be/~preneel/phd_preneel_feb1993.pdf

[3] E. Biham, O. Dunkelman, A framework for iterative hash functions, 2007, p. 9. http://csrc.nist.gov/groups/ST/hash/documents/DUNKELMAN_NIST3.pdf

[4] J-P. Aumasson, O. Dunkelman, S. Indesteege and B. Preneel, *Cryptanalysis of Dynamic SHA(2)*, COmputer Security and Industrial Cryptography publications, 2009, p. 18. https://www.cosic.esat.kuleuven.be/publications/article-1277.pdf

[5] N. A. Moldovyan, A. A. Moldovyan, M. A. Eremeev, *Cryptography: from Primitives to the Algoritms Synhesis,* BHV-Petersburgh, St. Petersburgh, 2004, p. 448. (in Russian)

[6] V. A. Luzhetsky, Y. V. Baryshev, The pseudo-nondeterministic hashing conception, *Systems of Control, Navigation and Communications*, (3) (2010), pp. 94-98. (in Ukrainian)

[7] J. A. Anderson, *Discrete Mathematics with Combinatorics*, Williams Publishing House, Moscow, 2004, p. 960. (in Russian)

[8] A. V. Aho, J. E. Hopcroft, J. D. Ullman, *The Design and Analysis of Computer Algoritms,* Mir, Moscow, 1979, p. 536. (in Russian)

[9] P. Gauravaram, *Cryptographic Hash Functions: Cryptanalysis, Design and Applications,* Thesis submitted in accordance with the regulations for Degree of Doctor of Philosophy, 2009, p. 298, http://eprints.qut.edu.au/16372/1/Praveen_Gauravaram_Thesis.pdf

[10] S. Hirose, *Some Plausible Constructions of Double-Block-Length Hash Functions*, 2006, p. 13. www.iacr.org/archive/fse2006/40470213/40470213.pdf

[11] S. Lucks, Design principles for iterated hash functions, *Cryptology ePrint Archive*, 2004, p. 22. http://eprint.iacr.org/2004/253.pdf

[12] G. Bertoni, J. Daemen, M. Peeters, G. Van Assche, *Sponge Functions*, 2007, p. 22, http://sponge.noekeon.org/SpongeFunctions.pdf

[13] Y. V. Baryshev, Methods of multipipe hash function infeasibility improving against generic attacks, *Computer Science and Engineering – 2010*, Lviv Polytechnic Publishing, Lviv, 2010, pp. 338-339. (in Ukrainian)

[14] Y. V. Baryshev. Pseudonondeterministic hashing mathematical model and cryptographic primitives for its implementation, *Information technologies and computer engineering – 2010,* VNTU, Vinnytsia, pp. 268-269. (in Ukrainian)

[15] Y. V. Baryshev, Methods and software means of multipipe driven hashing, *Methods and tools of coding, protection and compression of information-2011,* VNTU, Vinnytsia, pp. 100-101. (in Ukrainian)

[16] Secure Hash Standard: Federal Information Processing Publication Standard Publication 180-3. – Gaithersburg, 2008, p. 27, http://csrc.nist.gov/publications/fips/fips180-3/fips180-3_final.pdf

[17] Test files and Source Code for Conducting KAT and MCT, NIST, http://csrc.nist.gov/groups/ST/hash/sha-3/documents/KAT1.zip

---

Volodymyr A. Luzhetsky, *Dr. of Science, professor, head of the information protection department of Vinnytsia National Technical University.*

Scientific interests: the cryptography, data compression methods, Fibonacci codes.

Yuriy V. Baryshev*, phD, member of the information protection department of Vinnytsia National Technical University.*

Scientific interests: multipipe hashing and generic attacks, the evaluation of information protection.

Ashok Kumar, Vinay Goyal

EARCT – ENVIRONMENT FOR AUTOMATED RANK BASED CONTINUOUS AGENT TESTING

Testing MAS (Multi-agent System) is a challenging task because these systems are distributed, complex and autonomous in nature. Agents exist in an open environment having their own locus of control and they require context awareness. So due to these agent's characteristics, testing MAS system using existing testing techniques becomes a very tedious job. Agents also pose problems regarding message communication and semantic interoperability, as well as synchronization with other agents existing in the environment. All these features are known to be hard not only to design and to code, but also to test. In this paper we will propose a unique environment EARCT to test MAS keeping in mind the essential software engineering paradigms such as effort consumed, errors revealed etc.

Anton Mykhailiuk, Olena Mykhailiuk, Oleksiy Pylypchuk, Volodymyr Tarasenko

A CREATION OF THE LINGUISTIC ONTOLOGY BASED ON A STRUCTURED ELECTRONIC ENCYCLOPEDIC RESOURCE

Solving the problem of intelligent information retrieval requires the use of additional specialized linguistic resources. One of such type of resources is a linguistic ontology of a subject field. The paper considers an approach to develop software for an automatic creation of an ontological knowledge base by the converting structured encyclopedic resource into appropriate ontology's objects. The procedures of creation of the ontology entities base, the hierarchy of entities and the network of associative links are considered. Qualitative and quantitative analysis of the produced experimental ontology based on Ukrainian part of Wikipedia is investigated.

Marina Polyakova, Victor Krylov, Natalya Volkova

THE METHOD OF WAVELETS CONSTRUCTION BY TRANSFORMATION OF A GRAPH OF POWER FUNCTION FOR EDGE DETECTION

In this paper the method to construct the improved wavelets by transformation of a graph of power function is developed for the edge detection.

Dipak V. Patil, Rajankumar S. Bichkar

INTEGRATED EFFECT OF DATA CLEANING AND SAMPLING ON DECISION TREE LEARNING OF LARGE DATA SETS

The advances and use of technology in all walks of life results in tremendous growth of data available for data mining. Large amount of knowledge available can be utilized to improve decision-making process. The data contains the noise or outlier data to some extent which hampers the classification performance of classifier built on that training data. The learning process on large data set becomes very slow, as it has to be done serially on available large datasets. It has been proved that random data reduction techniques can be used to build optimal decision trees. Thus, we can integrate data cleaning and data sampling techniques to overcome the problems in handling large data sets. In this proposed technique outlier data is first filtered out to get clean data with improved quality and then random sampling technique is applied on this clean data set to get reduced data set. This reduced data set is used to construct optimal decision tree.

Experiments performed on several data sets proved that the proposed technique builds decision trees with enhanced classification accuracy as compared to classification performance on complete data set. Due to use of classification filter a quality of data is improved and sampling reduces the size of the data set. Thus, the proposed method constructs more accurate and optimal sized decision trees and it also avoids problems like overloading of memory and processor with large data sets. In addition, the time required to build a model on clean data is significantly reduced providing significant speedup.

Anton Kabysh, Vladimir Golovko

GENERAL MODEL FOR ORGANIZING INTERACTIONS IN MULTI-AGENT SYSTEMS

In this article we describe a model for finding optimal for learning the behavior of a group of agents in a collaborative multiagent setting. This model contains a set of scalable techniques that organize behavior of a multi-agent system. As a basis we use the framework of coordination graphs which exploits the dependencies between agents to decompose the global payoff function into a sum of local terms. To estimate a quality of interactions between agents we are using the concepts of the influence value learning paradigm. In last section we present the implementation of the considered model via reinforcement learning and experimental results of the use of this paradigm.

Artur Voronych, Yaroslav Nykolaychuk, Volodymyr Hladyuk
FORMATION AND PROCESSING OF ENTROPY-MANIPULATED CORRECTING SIGNAL CODES IN WIRELESS SENSOR NETWORKS

It is described a systematization of theoretical bases and analytical expressions for evaluation of entropy measures. It is showed a characteristic of signal corrective codes of Galois field. It is described the use of entropic manipulation and corrective signal codes in wireless sensor networks. It is founded that entropy approach of applying the corrective codes of Galois field is the most effective and promising for the formation and processing of signals during transmission in wireless sensor networks.

Svetlana Antoschuk, Dmitry Maevsky
STRUCTURAL SYNTHESIS OF DYNAMIC INFORMATION SYSTEMS

Theoretical bases are considered and the method of synthesis of the dynamic information systems structure is created with a capacity for adaptation at the changes of a subject domain. A capacity for adaptation is provided due to the selection of a variation part of algorithms of the system that changing at the changing the subject domain. The variation part of algorithms is kept in the database of the system that provides their rapid replacement and promotes a fail safety on the stage of exploitation.

S.P.Khandait, R.C.Thool, P.D.Khandait
ANFIS AND NEURAL NETWORK BASED FACIAL EXPRESSION RECOGNITION USING CURVELET FEATURES

Curvelet transform is a promising tool for multi-resolution analysis on images. This paper explains a new approach for facial expression recognition based on curvelet features extracted using curvelet transform. Curvelet transform is applied on the database images and curvelet coefficients are obtained for selected scale for image analysis. Facial curvelet features are compressed using singular value decomposition (SVD) approach. Back propagation neural network (BPNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) are used as classifiers for classifying expressions into one of the seven categories like angry, disgust, fear, happy, neutral, sad and surprise. Experimentation is carried out on JAFFE database. The experimental results show that the novel approach is a better option for extracting feature values and classifying facial expressions.

Bohdan Y. Volochiy, Leonid D. Ozirkovskyy, Ihor W. Kulyk
THE METHOD OF BUILDING OF BEHAVIOR MODEL OF NON-MARKOV COMPLEX SYSTEMS AS A GRAPH OF STATES AND TRANSITIONS

The goal of the paper is the construction of behavior models of discrete-continuous stochastic systems of non-Markov type by the method of Erlang's phases. The improved method of formalized construction of behavior models as a graph of states and transitions is developed. Since this method is a part of technology modeling discrete-continuous stochastic systems, now the process of construction of behavior systems of non-Markov type is automated.

Yuriy Semchyshyn, Ivan Kulpa, Igor Kolosovskyi, Oleksandr Hrechnikov, Petro Hayda
DEVELOPING SEMANTIC NETWORK MANAGEMENT SYSTEM

The paper considers semantic networks. It is a way of representing knowledge as a set of nodes-concepts connected by edges-relations. The history and current state-of-the-art of semantic networks is analyzed. The experience of designing and developing semantic network management system is described in four sections covering the following topics: data structures design, semantic analysis algorithm implementation, graph visualization methods selection and Web-based user interface development. The developed system is fully-operational tool for building and studying semantic networks. The system will be used for the further research.

Vitaliy G. Deibuk, Ivan M. Yuriychuk, Roman I. Yuriychuk
SPIN MODEL OF FULL SUMMATOR ON PERES GATES

Four-spin model of single-digit summator for solid state nuclear magnetic resonance (NMR) quantum computer is proposed. Correctness of summator's operation is analyzed in dependence of system parameters. It is found that the summator on Peres gates works correctly in time of four π-pulses on pure numerical and superposition states. Optimal values for exchange interactions and maximum distance between q-bits necessary for correct summator's operation are obtained from the analysis of fidelity function.

Julia Drozd, Alexander Drozd, Julian Sulima

NATURAL RESOURCES AND THEIR USE FOR CHECKABILITY INCREASING THE DIGITAL COMPONENTS OF SAFETY-CRITICAL SYSTEMS

The models, methods and means as target resources for solving tasks for design and testing of computer systems and their components are considered. A criterion for choosing the target resources activating the natural resources for the increasing the efficiency of task solutions is determined. A problem of low checkability of digital components of safety-critical systems is examined and the ways of its solution is showed by choosing the target resources in accordance with the offered criterion. The way of elimination of a contradiction between the target resources aimed to maintain the checkability, productivity and low complexity of digital components is described. This way is based on parallelization of computations with the use of the serial codes.

Volodymyr A. Luzhetsky, Yuriy V. Baryshev

THE GENERALIZED CONSTRUCTION OF PSEUDONONDETERMINISTIC HASHING

This article is devoted to the development of hash constructions, which are based on the pseudonondeterministic hash conception. The conception allows to design hash functions with improved infeasibility to cryptanalysis. Both proposed constructions and known ones are generalized as pseudonondeterministic constructions.

Ashok Kumar, Vinay Goyal
EARCT – СЕРЕДОВИЩЕ ДЛЯ АВТОМАТИЗОВАНОГО ТЕСТУВАННЯ ПОСТІЙНИХ АГЕНТІВ НА ОСНОВІ РАНЖУВАННЯ

Тестування мультиагентних систем (MAC) є складним завданням, оскільки ці системи є дистрибутивними, складними та автономними за своєю природою. Агенти існують у відкритому середовищі, мають своє розміщення та вимагають контекстної компетентності. Таким чином, у зв'язку з цими характеристиками агентів, тестування MAC за допомогою існуючих методів є дуже рутинною роботою. Агенти також створюють проблеми щодо комунікаційних повідомлень та семантичної сумісності, а також синхронізації з іншими агентами, існуючими в навколишньому середовищі. Всі ці особливості вносять складність не тільки в проектування та програмування, але й тестування. У цій статті ми пропонуємо створення унікального EARCT середовища для тестування MAC з урахуванням необхідних парадигм розробки програмного забезпечення, таких як затрачені зусилля, виявлені помилки тощо.

Антон Михайлюк, Олена Михайлюк, Олексій Пилипчук, Володимир Тарасенко
ФОРМУВАННЯ ЛІНГВІСТИЧНОЇ ОНТОЛОГІЇ НА БАЗІ СТРУКТУРОВАНОГО ЕЛЕКТРОННОГО ЕНЦИКЛОПЕДИЧНОГО РЕСУРСУ

Вирішення задачі інтелектуалізації інформаційно-пошукової діяльності вимагає застосування допоміжних спеціалізованих лінгвістичних ресурсів. Одним із таких ресурсів може бути лінгвістична онтологія предметної галузі. Стаття розглядає підхід до організації програмних засобів для автоматизованого формування онтологічної бази знань шляхом конвертації структурованого енциклопедичного ресурсу у відповідні об'єкти онтології. Розглядаються процедури створення понятєвої бази онтології, ієрархії понять та мережі асоціативних зв'язків. Проводиться дослідження якісного та кількісного складу сформованої експериментальної онтології на основі українського сегменту Вікіпедії.

Марина Полякова, Віктор Крилов, Наталія Волкова
МЕТОД ПОБУДОВИ ПОКРАЩЕНИХ ВЕЙВЛЕТІВ ШЛЯХОМ ПЕРЕТВОРЕННЯ ГРАФІКА СТЕПЕНЕВОЇ ФУНКЦІЇ ДЛЯ ЗАДАЧІ КОНТУРНОЇ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

В даній роботі розроблено метод побудови покращених вейвлетів для задачі контурної сегментації зображень шляхом перетворення графіка степеневої функції.

Dipak V. Patil, Rajankumar S. Bichkar
ІНТЕГРАЛЬНИЙ ЕФЕКТ ПІДГОТОВКИ ДАНИХ І ВИБІРКИ ДЛЯ МЕТОДУ НАВЧАННЯ НА ОСНОВІ ДЕРЕВА РІШЕНЬ НА ПРИКЛАДІ ВЕЛИКИХ НАБОРІВ ДАНИХ

Розвиток і використання технологій у всіх сферах життя призводить до величезного росту даних, доступних для аналізу. Великий обсяг доступних знань може бути використаний для покращення процесу прийняття рішень. Дані, що містять шуми або викиди, знижують точність класифікації при побудові класифікаторів на основі таких навчальних даних. Процес навчання на великій вибірці даних стає дуже повільним, так як він повинен виконуватися послідовно на наявних великих вибірках даних. Було доведено, що зменшення об'єму даних випадковим чином може бути використано для побудови оптимального дерева рішень. Таким чином, ми можемо інтегрувати процеси підготовки даних та методику вибору даних для подолання проблем в обробці великих масивів даних. У запропонованій методиці спочатку фільтруються викиди, щоб отримати «чисті» дані покращеної якості, а потім застосовується випадковий вибір на цій «чистій» вибірці даних, щоб зменшити актуальну вибірку даних. Ця скорочена вибірка даних використовується для побудови оптимального дерева рішень.

Досліди, проведені на декількох вибірках даних довели, що запропонована методика будує дерево рішень з підвищеною точністю класифікації у порівнянні з точністю класифікації на повній вибірці даних. Завдяки використанню фільтрів класифікації якість даних покращується і відбір даних з вибірки зменшує її розмір. Таким чином, запропонований метод створює точніші та оптимальніші розміри дерев рішень, і це також дозволяє уникнути таких проблем, як перевантаження оперативної пам'яті та процесора великими вибірками даних. Крім того, час, необхідний для побудови моделі на чистих даних, значно скорочується, що забезпечує значне прискорення.

Антон Кабиш, Володимир Головко
УЗАГАЛЬНЕНА МОДЕЛЬ ОРГАНІЗАЦІЇ ВЗАЄМОДІЙ В МУЛЬТИАГЕНТНІЙ СИСТЕМІ

У даній роботі описується модель для знаходження оптимальної поведінки мультиагентної структури через організацію в ній оптимальних взаємодій між агентами. Модель включає дві основні техніки. Модель графів координації дозволяє явно виразити залежність між агентами, що дає можливість розбити цільову функцію поведінки в лінійну суму індивідуальних цільових функцій. Модель оцінки впливів дає можливість оцінити вплив інших агентів на дії один одного і як результат дає їм координувати свої дії. В роботі наведена реалізація даної моделі на основі навчання з підкріпленням і експериментальні результати застосування даної моделі.

Артур Воронич, Ярослав Николайчук, Володимир Гладюк
ФОРМУВАННЯ ТА ОПРАЦЮВАННЯ ЕНТРОПІЙНО-МАНІПУЛЬОВАНИХ СИГНАЛЬНИХ КОРЕКТУЮЧИХ КОДІВ У БЕЗПРОВІДНИХ СЕНСОРНИХ МЕРЕЖАХ

Викладена систематизація теоретичних основ та аналітичних виразів оцінки мір ентропії. Наведено характеристику сигнальних коректуючих кодів поля Галуа. Описано можливість використання ентропійної маніпуляції та сигнальних коректуючих кодів у безпровідних сенсорних мережах. Встановлено, що енторопійний підхід з застосування коректуючих кодів поля Галуа є найбільш ефективним та перспективним для формування і опрацювання сигналів при передачі інформації в безпровідних сенсорних мережах.

Світлана Антощук, Дмитро Маєвський
СТРУКТУРНИЙ СИНТЕЗ ДИНАМІЧНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

Розглянуто теоретичні основи та створено метод синтезу структури динамічних інформаційних систем із здатністю до адаптування при змінах предметної області. Здатність до адаптування забезпечується за рахунок виділення варіативної частини алгоритмів системи, що змінюються при змінах предметної області. Варіативна частина алгоритмів зберігається в базі даних системи, що забезпечує їх швидку заміну та підвищує надійність системи на етапі експлуатації.

S.P.Khandait, R.C.Thool, P.D.Khandait
АДАПТИВНЕ НЕЙРО-НЕЧІТКЕ ТА НЕЙРОМЕРЕЖЕВЕ ШВИДКЕ РОЗПІЗНАВАННЯ ВИРАЗУ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ CURVELET-ОСОБЛИВОСТЕЙ

Curvelet-перетворення є перспективним інструментом для мульти-роздільного аналізу зображень. Ця стаття представляє новий підхід до розпізнавання виразу обличчя на основі curvelet-особливостей, видобутих за допомогою curvelet-перетворень. Curvelet-перетворення застосовується в базі даних зображень і curvelet-коефіцієнти були отримані для вибраного масштабу для аналізу зображень. Риси обличчя у вигляді curvelet-коефіцієнтів стиснуті з використанням підходу сингулярного розкладання (SVD). Нейронні мережі із зворотнім поширенням помилки навчання (BPNN) і адаптивна нейро-нечітка система (ANFIS) використовуються в якості класифікаторів для класифікації виразів в одній із семи категорій, таких як гнів, відраза, страх, щастя, нейтральність, сум і здивування. Експеримент проводився на базі даних JAFFE. Результати експерименту показали, що новий підхід є кращим варіантом для видобутку значень ознак та класифікації виразу обличчя.

Богдан Волочій, Леонід Озірковський, Ігор Кулик
МЕТОД ПОБУДОВИ МОДЕЛЕЙ ПОВЕДІНКИ СКЛАДНИХ СИСТЕМ НЕМАРКОВСЬКОГО ТИПУ У ВИГЛЯДІ ГРАФА СТАНІВ І ПЕРЕХОДІВ

Об'єктом розгляду є процес побудови моделей поведінки дискретно-неперервних стохастичних систем немарковського типу за допомогою методу фаз Ерланга. На базі цього методу та з використанням удосконаленої технології моделювання дискретно-неперервних стохастичних систем марковського типу розроблено метод побудови моделей складних систем немарковського типу у вигляді графа станів і переходів. Розроблений метод проілюстровано прикладом аналізу системи масового обслуговування.

Юрій Семчишин, Іван Кульпа, Ігор Колосовський, Олександр Гречніков, Петро Гайда
РОЗРОБКА СИСТЕМИ КЕРУВАННЯ СЕМАНТИЧНИМИ МЕРЕЖАМИ

Робота присвячена семантичним мережам — способу представлення знань у виді набору вузлів-сутностей об'єднаних за допомогою ребер-зв'язків. Здійснено огляд історії розвитку та сучасного стану семантичних мереж, обґрунтовано актуальність проблеми. Описано процес проектування та

розробки системи керування семантичними мережами: проектування структур даних, розробку алгоритмів семантичного аналізу текстів українською мовою, вибір алгоритмів візуалізації графових структур даних, реалізацію користувацького інтерфейсу у вигляді Web-додатку. Розроблена система є повністю функціональним засобом побудови та дослідження семантичних мереж, що підтвердив свою ефективність та послужить базовою платформою для подальших досліджень.

Віталій Дейбук, Іван Юрійчук, Роман Юрійчук
СПІНОВА МОДЕЛЬ ПОВНОГО ОДНОРОЗРЯДНОГО СУМАТОРА НА ЕЛЕМЕНТАХ ПЕРЕСА

Побудована чотириспінова модель однорозрядного суматора для твердотільного ЯМР квантового комп'ютера та проаналізована коректність його роботи в залежності від параметрів системи. Встановлено, що для чистих та суперпозиційних станів, суматор на елементах Переса коректно спрацьовує за чотири π-імпульси. Отримано оптимальні значення параметрів обмінної взаємодії на основі аналізу функції чіткості та визначена максимальна відстань між кубітами для коректної роботи суматора.

Юлія Дрозд, Олександр Дрозд, Юліан Сулима
ПРИРОДНІ РЕСУРСИ ТА ЇХ ВИКОРИСТАННЯ ДЛЯ ПІДВИЩЕННЯ КОНТРОЛЕПРИДАТНОСТІ ЦИФРОВИХ КОМПОНЕНТІВ СИСТЕМ КРИТИЧНОГО ЗАСТОСУВАННЯ

Розглянуто моделі, методи та засоби як цільові ресурси для вирішення задач проектування та діагностування комп'ютерних систем та їх компонентів. Визначається критерій вибору цільових ресурсів, що активують природні ресурси, спрямовані на підвищення ефективності вирішення завдань. Розглянуто проблему низької контролепридатності цифрових компонентів систем критичного застосування, та показані шляхи її вирішення при виборі цільових ресурсів відповідно до запропонованого критерію. Описано шлях усунення суперечності між цільовими ресурсами, спрямованими на забезпечення контролепридатності, продуктивності та малої складності цифрових компонентів. Цей шлях ґрунтується на розпаралелюванні обчислень з використанням послідовних кодів.

Володимир Лужецький, Юрій Баришев
УЗАГАЛЬНЕНА КОНСТРУКЦІЯ ПСЕВДОНЕДЕТЕРМІНОВАНОГО ХЕШУВАННЯ

Дана стаття присвячена розробці конструкцій хешування, які базуються на концепції псевдонедетермінованого хешування. Дана концепція дозволяє розробляти хеш-функції з підвищеною стійкістю до криптоаналізу. І запропоновані конструкції, і відомі узагальнені як псевдонедетерміновані конструкції.

Ashok Kumar, Vinay Goyal
EARCT – СРЕДА ДЛЯ АВТОМАТИЗИРОВАННОГО ПОСТОЯННОГО КОНТРОЛЯ АГЕНТОВ НА ОСНОВЕ РАНЖИРОВАНИЯ

Тестирование мультиагентных систем (MAC) сложная задача, поскольку эти системы являются дистрибутивными, сложными и автономными по своей сущности. Агенты существуют в открытой среде, имеют свое место и требуют контекстной компетентности. Таким образом, в связи с этими характеристиками агентов, тестирование MAC с помощью существующих методов является очень рутинной работой. Агенты также создают проблемы относительно коммуникационных сообщений и семантической совместимости, а также синхронизации с другими агентами, существующими в окружающей среде. Все эти особенности вносят сложность не только в проектирование и программирование, но и в тестирование. В этой статье мы предлагаем создание уникальной EARCT среды для тестирования MAC с учетом необходимых парадигм разработки программного обеспечения, таких как затраченные усилия, обнаруженные ошибки и т.д.

Антон Михайлюк, Елена Михайлюк, Алексей Пилипчук, Владимир Тарасенко
ФОРМИРОВАНИЕ ЛИНГВИСТИЧЕСКОЙ ОНТОЛОГИИ НА БАЗЕ СТРУКТУРИРОВАННОГО ЭЛЕКТРОННОГО ЭНЦИКЛОПЕДИЧЕСКОГО РЕСУРСА

Решение задачи интеллектуализации информационно-поисковой деятельности требует применения вспомогательных специализированных лингвистических ресурсов. Одним из таких ресурсов может быть лингвистическая онтология предметной области. Статья рассматривает подход к организации программных средств для автоматизированного формирования онтологической базы знаний путем конвертации структурированного энциклопедического ресурса в соответствующие объекты онтологии. Рассматриваются процедуры создания понятийной базы онтологии, иерархии понятий и сети ассоциативных связей. Проводится исследование качественного и количественного состава сложившейся экспериментальной онтологии на основе украинского сегмента Википедии.

Марина Полякова, Виктор Крылов, Наталья Волкова
МЕТОД ПОСТРОЕНИЯ УЛУЧШЕННЫХ ВЕЙВЛЕТОВ ПУТЕМ ПРЕОБРАЗОВАНИЯ ГРАФИКА СТЕПЕННОЙ ФУНКЦИИ ДЛЯ ЗАДАЧИ КОНТУРНОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

В данной работе разработан метод построения улучшенных вейвлетов для задачи контурной сегментации изображений путем преобразования графика степенной функции.

Dipak V. Patil, Rajankumar S. Bichkar
ИНТЕГРАЛЬНЫЙ ЭФФЕКТ ПОДГОТОВКИ ДАННЫХ И ВЫБОРКИ ДЛЯ МЕТОДА ОБУЧЕНИЯ НА ОСНОВЕ ДЕРЕВА РЕШЕНИЙ НА ПРИМЕРЕ БОЛЬШИХ НАБОРОВ ДАННЫХ

Развитие и использование технологий во всех сферах жизни приводит к огромному росту данных, доступных для анализа. Большой объем доступных знаний может быть использован для улучшения процесса принятия решений. Данные, содержащие шумы или выбросы снижают точность классификаторов на основе таких учебных данных. Процесс обучения на большой выборке данных становится очень медленным, как он должен выполняться последовательно на имеющихся больших выборках данных. Было доказано, что уменьшение объема данных случайным образом, может быт использовано для построения оптимального дерева решений. Таким образом, мы можем интегрировать процессы подготовки данных и методику выбора данных для преодоления проблем в обработке больших массивов данных. В предложенной методике сначала фильтруются выбросы, чтобы получить «чистые» данные улучшенного качества, а затем применяется случайный выбор на этой чистой выборке данных, чтобы уменьшить актуальную выборку данных. Эта сокращенная выборка данных используется для построения оптимального дерева решений.

Опыты, проведенные на нескольких выборках данных доказали, что предложенная методика строит дерево решений с повышенной точностью классификации по сравнению с точностью классификации на полной выборке данных. Благодаря использованию фильтров классификации качество данных улучшается и отбор данных из выборки уменьшает ее размер. Таким образом, предложенный метод создает более точные и более оптимальные размеры деревьев решений, и это также позволяет избежать таких проблем, как перегрузка оперативной памяти и процессора большими выборками данных. Кроме того, время, необходимое для построения модели на чистых данных значительно сокращается, что обеспечивает значительное ускорение.

Антон Кабыш, Владимир Головко
ОБОБЩЕННАЯ МОДЕЛЬ ОРГАНИЗАЦИИ ВЗАИМОДЕЙСТВИЙ В МУЛЬТИГАГЕНТНОЙ СИСТЕМЕ

В данной работе описывается модель для нахождения оптимального поведения многоагентной структуры через организацию в ней оптимальных взаимодействий между агентами. Модель включает две основные техники. Модель графов координации позволяет явно выразить зависимость между агентами, что позволяет разбить целевую функцию поведения в линейную сумму индивидуальных целевых функций. Модель оценки влияний позволяет оценить влияния других агентов на действия друг друга и как результат позволяет им координировать свои действия. В работе приведена реализация данной модели на основе обучения с подкреплением и экспериментальные результаты применения данной модели.

Артур Воронич, Ярослав Николайчук, Владимир Гладюк
ФОРМИРОВАНИЕ И ОБРАБОТКА ЭНТРОПИЙНО-МАНИПУЛИРУЕМЫХ СИГНАЛЬНЫХ КОРРЕКТИРУЮЩИХ КОДОВ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ

Изложенная систематизация теоретических основ и аналитических выражений оценки мер энтропии. Приведена характеристика сигнальных корректирующих кодов поля Галуа. Описана возможность использования энтропийной манипуляции и сигнальных корректирующих кодов в беспроводных сенсорных сетях. Установлено, что энторопийный подход по использованию корректирующих кодов поля Галуа является наиболее эффективным и перспективным для формирования и обработки сигналов при передаче информации в беспроводных сенсорных сетях.

Светлана Антощук, Дмитрий Маевский
СТРУКТУРНЫЙ СИНТЕЗ ДИНАМИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ

Рассмотрены теоретические основы и создан метод синтеза структуры динамических информационных систем со способностью к адаптации при изменениях предметной области. Способность к адаптации обеспечивается за счет выделения вариативной части алгоритмов системы, которые изменяются при изменениях предметной области. Вариативная часть алгоритмов хранится в базе данных системы, что обеспечивает их быструю замену и повышает надежность системы на этапе эксплуатации.

S.P.Khandait, R.C.Thool, P.D.Khandait
АДАПТИВНОЕ НЕЙРО-НЕЧЕТКОЕ И НЕЙРОСЕТЕВОЕ БЫСТРОЕ РАСПОЗНАВАНИЕ ВЫРАЖЕНИЯ ЛИЦА С ИСПОЛЬЗОВАНИЕМ CURVELET ОСОБЕННОСТЕЙ

Curvelet-преобразования является перспективным инструментом для мульти-разрешающего анализа изображений. Эта статья представляет новый подход к распознаванию выражения лица на основе curvelet-особенностей добытых с помощью curvelet-преобразований. Curvelet-преобразования применяются в базе данных изображений, и curvelet-коэффициенты были получены для выбранного масштаба для анализа изображений. Черты лица в виде curvelet-коэффициентов сжатые с использованием подхода сингулярного разложения (SVD). Нейронные сети с обратным распространением ошибки обучения (BPNN) и адаптивная нейро-нечеткая система (ANFIS) используются в качестве классификаторов для классификации выражений в одной из семи категорий, таких как гнев, отвращение, страх, счастье, нейтральность, печаль и недоумение. Эксперимент проводился на базе данных JAFFE. Результаты эксперимента показали, что новый подход является лучшим вариантом для добычи значений особенностей и классификации выражения лица.

Богдан Волочий, Леонид Озирковский, Игорь Кулик
МЕТОД ПОСТРОЕНИЯ МОДЕЛЕЙ ПОВЕДЕНИЯ СЛОЖНЫХ СИСТЕМ НЕМАРКОВСКОГО ТИПА В ВИДЕ ГРАФА СОСТОЯНИЙ И ПЕРЕХОДОВ

Объектом рассмотрения является процесс построения моделей поведения дискретно-непрерывных стохастических систем немарковского типа с помощью метода фаз Эрланга. На базе этого метода и с использованием усовершенствованной технологии моделирования дискретно-непрерывных стохастических систем Марковского типа разработан метод построения моделей сложных систем немарковского типа в виде графа состояний и переходов. Разработанный метод проиллюстрирован примером анализа системы массового обслуживания.

Юрий Семчишин, Иван Кульпа, Игорь Колосовский, Александр Гречников, Петр Гайда
РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ СЕМАНТИЧЕСКИМИ СЕТЯМИ

Работа посвящена семантическим сетям – способу представления знаний в виде набора узлов-сущностей объединенных с помощью ребер-связей. Осуществлен обзор истории развития и современного состояния семантических сетей, обоснована актуальность проблемы. Описано процесс проектирования и разработки системы управления семантическими сетями: проектирование структур данных, разработку алгоритмов семантического анализа текстов на украинском языке, выбор алгоритмов визуализации графовых структур данных, реализацию пользовательского интерфейса в виде Web-приложения. Разработанная система является полностью функциональным средством построения и исследования семантических сетей, подтвердила свою эффективность и послужит базовой платформой для дальнейших исследований.

Виталий Дейбук, Иван Юрийчук, Роман Юрийчук
СПИНОВАЯ МОДЕЛЬ ПОЛНОГО ОДНОРАЗРЯДНОГО СУММАТОРА НА ЭЛЕМЕНТАХ ПЕРЕСА

Построенная четырехспиновая модель одноразрядного сумматора для твердотельного ЯМР квантового компьютера и проанализирована корректность его работы в зависимости от параметров системы. Установлено, что для чистых и суперпозиционных состояний, сумматор на элементах Переса корректно срабатывает за четыре р-импульса. Получены оптимальные значения параметров обменного взаимодействия на основе анализа функции четкости и определено максимальное расстояние между кубитами для корректной работы сумматора.

Юлия Дрозд, Александр Дрозд, Юлиан Сулима
ЕСТЕСТВЕННЫЕ РЕСУРСЫ И ИХ ИСПОЛЬЗОВАНИЕ ДЛЯ ПОВЫШЕНИЯ КОНТРОЛЕПРИГОДНОСТИ ЦИФРОВЫХ КОМПОНЕНТОВ СИСТЕМ КРИТИЧЕСКОГО ПРИМЕНЕНИЯ

Рассмотрены модели, методы и средства как целевые ресурсы для решения задач проектирования и диагностирования компьютерных систем и их компонентов. Определяется критерий выбора целевых ресурсов, активирующих естественные ресурсы, направленные на повышение эффективности решения задач. Рассмотрена проблема низкой контролепригодности цифровых компонентов систем критического применения, и показаны пути ее решения при выборе целевых ресурсов в соответствии с предложенным критерием. Описан путь устранения противоречия между целевыми ресурсами, направленными на обеспечение контролепригодности, производительности и малой сложности цифровых компонентов. Этот путь основывается на распараллеливании вычислений с использованием последовательных кодов.

Владимир Лужецкий, Юрий Барышев
ОБОБЩЕНАЯ КОНСТРУКЦИЯ ПСЕВДОНЕДЕТЕРМИНИРОВАННОГО ХЕШИРОВАНИЯ

Даная статья посвящена разработке конструкций хеширования, которые основываются на концепции псевдонедетерминированного хеширования. Эта концепция позволяет создавать хеш-функции с повышенной стойкостью к криптоанализу. И предложенные конструкции, и известные обобщены как псевдонедетерминированные конструкции.

Prepare your paper according to the following requirements:

Unconditional requirement - paper should not be published earlier.

(i) Suggested composition (frame) of paper:
- Issue formulation stressing its urgent solving; evaluation of recent publications in the explored issue
- short formulation of paper's purpose
- description of proposed method (algorithm)
- implementation and testing (verification)
- conclusion.

(ii) Use A4 (210 x 297 mm) paper. Size of paper has to be extended up to 6-8 pages.

(iii) Please use main text two column formatting;

(iv) A paper must have an abstract and some keywords;

(v) Place a full list of references at the end of the paper. Please place the references according to their order of appearance in the text.

(vi) An affiliation of each author is wanted.

(vii) The text should be single-spaced. Use Times New Roman (11 points, regular) typeface throughout the paper.

(viii) Equations should be placed in separate lines and numbered. The numbers should be within brackets and right aligned.

(ix) The figures and tables must be numbered, have a self-contained caption. Figure captions should be below the figures; table captions should be above the tables. Also, avoid placing figures and tables before their first mention in the text.

(x) As soon as you have the complete materials, the final versions should come electronically in MS Word'97 of MS Word 2000 format to the address computing@computingonline.net.

(xi) A hardcopy of your article is needed to be sent by regular mail for our publishing house.

(xii) Please send short CVs (up to 20 lines) and photos of every author.

(xiii) There is no other formatting required. The publishing department makes all rest formatting according to the publisher's rules.

Journal Topics:
- Algorithms and Data Structure
- Bio-Informatics
- Cluster and Parallel Computing, Software Tools and Environments
- Computational Intelligence
- Computer Modeling and Simulation
- Cyber and Homeland Security
- Data Communications and Networking
- Data Mining, Knowledge Bases and Ontology
- Digital Signal Processing
- Distributed Systems and Remote Control
- Education in Computing
- Embedded Systems
- High Performance Computing and GRIDS
- Image Processing and Pattern Recognition
- Intelligent Robotics Systems
- Internet of Things
- Standardization of Computer Systems
- Wireless Systems

Основні вимоги до подання і оформлення публікацій наукового журналу "Комп'ютинг":

Безумовною вимогою є те, щоб стаття не була опублікована раніше!

(i) Наукові статті повинні мати такі необхідні елементи:
- постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями;
- аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми і на які спирається автор, виділення невирішених раніше частин загальної проблеми, котрим присвячується означена стаття;
- формулювання цілей статті (постановка завдання);
- виклад основного матеріалу дослідження з повним обгрунтуванням отриманих наукових результатів;
- висновки з даного дослідження і перспективи подальших розвідок у даному напрямку.

(ii) Використовуйте A4 (210 x 297 mm) формат сторінки. Загальний розмір статті має містити 6-8 сторінок.

(iii) Використовуйте двоколонкове форматування основного тексту;

(iv) Стаття повинна обов'язково містити основний текст українською мовою, анотацію (написану на Англійській і Українській мовах) і список ключових слів;

(v) В кінці статті розмістіть список літератури. Розміщуйте список літератури в порядку її цитування.

(vi) Необхідною є інформація про наукові звання, титули та посади авторів.

(vii) Текст повинен бути набраним одинарним інтервалом із використанням шрифту Times New Roman (11 points, regular).

(viii) Формули повинні відділятись від основного тексту пустими стрічками а також пронумеровані у круглих дужках та відцентровані по правому краю.

(ix) Таблиці і рисунки повинні бути пронумерованими. Заголовки рисунків розміщують під рисунком по центру. Заголовки таблиць розміщують по центру зверху таблиці.

(x) Завершені версії статей повинні бути надісланими в електронному MS Word'97 або MS Word 2000 форматі за адресою computing@computingonline.net.

(xi) Просимо надсилати поштою роздруковані копії статей.

(xii) В кінці кожної статті потрібно подати її назву, резюме (абстракт) і ключові слова англійською мовою.

(xiii) Просимо надсилати нам короткі біографічні дані (до 20 рядків) і скановані фотографії кожного із авторів.

(xiv) Видавництво здійснює остаточне форматування тексту згідно із вимогами друку.

(xv) У закордонних читачів можуть виникнути проблеми при ознайомленні з працями на російській та українській мовах. В зв'язку з цим редакційна колегія просить авторів додатково прислати розширений реферат (резюме), щоб б містило дві сторінки тексту англійською мовою, і супроводжувалось заголовком, прізвищами та адресами авторів. Авторам рекомендується використовувати у рисунках статті позначення переважно англійською мовою, або давати переклад у дужках. Тоді у розширеному резюме можна буде посилатися на рисунки у основному тексті.

Тематика журналу:

- Алгоритми та структури даних
- Біо-інформатика
- Кластерні та паралельні обчислення, програмні засоби та середовище
- Обчислювальний інтелект
- Комп'ютерне та імітаційне моделювання
- Кібернетична безпека та захист від тероризму
- Обмін даними та організація мереж
- Видобування даних, бази знань та онтології
- Цифрова обробка сигналів
- Розподілені системи та дистанційне управління
- Освіта в комп'ютингу
- Вбудовувані системи
- Високопродуктивні обчислення та ГРІД
- Обробка зображень та розпізнавання шаблонів
- Інтелектуальні робототехнічні системи
- Інтернет речей
- Стандартизація комп'ютерних систем
- Безпровідні системи

Основные требования к подаче и оформлению публикаций научного журнала “Компьютинг”:

Безусловное требование – чтобы статья не была опубликована ранее!

(i) Научные статьи должны иметь такие необходимые элементы:
- постановка проблемы в общем виде и ее связь с важными научными или практическими задачами;
- анализ последних исследований и публикаций, в которых начаты решения данной проблемы и на которые опирается автор, выделение нерешенных прежде частей общей проблемы, которым посвящается обозначенная статья;
- формулирование целей статьи (постановка задача);
- изложение основного материала исследования с полным обоснованием полученных научных результатов;
- выводы из данного исследования и перспективы дальнейших изысканий в данном направлении.

(ii) Используйте A4 (210 x 297 mm) формат страницы. Общий размер статьи 6-8 страниц.

(iii) Используйте двухколоночное форматирование основного текста;

(iv) Статья должна обязательно содержать основной текст на Русском языке, аннотацию (написанную на Английском и Русском языках) и список ключевых слов;

(v) В конце статьи разместите список литературы. Размещайте список литературы в порядке ее цитирования.

(vi) Необходима информация о научных званиях, титулах и должностях авторов.

(vii) Текст должен быть набранным одинарным интервалом с использованием шрифта Times New Roman (11 points, regular).

(viii) Формулы должны отделяться от основного текста пустыми строками, а также пронумерованные в круглых скобках и отцентрованные по правому краю.

(ix) Таблицы и рисунки должны быть пронумерованными. Заголовки рисунков размещают под рисунком по центру. Заголовки таблиц размещают по центру сверху таблицы.

(x) Завершенные версии статей должны быть присланы в электронном MS Word'97 или MS Word 2000 формате по адресу computing@computingonline.net.

(xi) Просим присылать распечатанные копии статей по почте.

(xii) В конце каждой статьи необходимо предоставить ее название, резюме (абстракт) и ключевые слова на английском языке.

(xiii) Просим присылать нам короткие биографические данные (до 20 строчек) и сканированные фотографии каждого из авторов.

(xiv) Издательство осуществляет окончательное форматирование текста в соответствии с требованиями печати.

(xv) У зарубежных читателей могут возникнуть проблемы при ознакомлении с трудами на русском и украинском языках. В связи с этим редакционная коллегия просит авторов дополнительно прислать расширенный реферат (резюме), который содержал бы две страницы текста на английском языке, и сопровождался заголовком, фамилиями и адресами авторов. Авторам рекомендуется использовать в рисунках статьи обозначения преимущественно на английском языке, или давать перевод в скобках. Тогда в расширенном резюме можно будет посылаться на рисунки в основном тексте.

Тематика журнала:

- Алгоритмы и структуры данных
- Био-информатика
- Кластерные и параллельные вычисления, программные средства и среды
- Вычислительный интеллект
- Компьютерное и имитационное моделирование
- Кибернетическая безопасность и защита от терроризма
- Обмен данными и организация сетей
- Добыча данных, базы знаний и онтологии
- Цифровая обработка сигналов
- Распределенные системы и дистанционное управление
- Образование в компьютинге
- Встраиваемые системы
- Высокопроизводительные вычисления и ГРИД
- Обработка изображений и распознавание шаблонов
- Интеллектуальные робототехнические системы
- Интернет вещей
- Стандартизация компьютерных систем
- Беспроводные системы

Підписано до друку 19. 11. 2012 р.
Формат 60х84 $^{1}/_{8}$. Гарнітура Times.
Папір офсетний. Друк на дублікаторі.
Умов. друк. арк. 16,5. Облік.-вид. арк. 19,7.
Зам. № М003-12П. Наклад 300 прим.

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців ДК № 3467 від 23.04.2009 р.

Віддруковано у Видавничо-поліграфічному центрі «Економічна думка ТНЕУ»
46004 м. Тернопіль, вул. Львівська, 11
тел. (0352) 47-58-72
E-mail: edition@tneu.edu.ua