Ministry of Education and Science of Ukraine
National Aviation University
Software Engineering Department

In collaboration with
The Institute for Information Theories and Applications ITHEA
ITHEA® International Scientific Society (ISS)
The Association of Developers and Users of Intelligent Systems

ITHEA®
International Scientific Society

International Conference on Software Engineering

June 03 – 06, 2019

Proceedings

Kyiv 2019

Conference organizers and partners

	National Aviation University www.nau.edu.ua
	Software engineering department www.sed.nau.edu.ua
ITHEA®	**The Institute for Information Theories and Applications FOI ITHEA®** http://www.ithea.org/index.html
ITHEA® International Scientific Society	**International Scientific Society ITHEA® ISS** http://www.ithea.org/iss/iss.html
	The Association of Developers and Users of Intelligent Systems http://www.aduis.com.ua/

NATIONAL AVIATION UNIVERSITY

(www.nau.edu.ua)

For almost 80 years of its history more than 200,000 highly skilled professionals have been trained in this higher aviation educational institution. Among them there are well-known scientists, heads of aviation companies, enterprises, organizations and institutions that provide aircraft flights, their maintenance and repair, transportation of passengers and cargo.

Strong scientific and pedagogical schools enable to train not only engineers but also economists, lawyers, environmentalists, translators, psychologists, sociologists and others.

SOFTWARE ENGINEERING DEPARTMENT

(www.sed.nau.edu.ua)

Software engineering department studies more than 900 students. Students obtain fundamental knowledge and practical skills in software development area as well as theoretical basic in software development life-cycle management.

INSTITUTE FOR INFORMATION THEORIES AND APPLICATIONS FOI ITHEA® (BULGARIA)

(http://www.ithea.org/index.html)

The Institute for Information Theories and Applications FOI ITHEA® is an international nongovernmental organization functioning since 2002 year. ITHEA® is aimed to support international scientific research through international scientific projects, workshops, conferences, journals, book series, etc.

INTERNATIONAL SCIENTIFIC SOCIETY

ITHEA® ISS (INTERNATIONAL SCIENTIFIC SOCIETY)

(http://www.ithea.org/iss/iss.html)

ITHEA® International Scientific Society (**ITHEA® ISS**) is aimed to support growing collaboration between scientists from all over the world.

Till now, the **ITHEA® International Scientific Society** has been joined by more than **4000** members from **53** countries from all over the world:

ITHEA® Publishing House (**ITHEA® PH**) is the official publisher of the works of the ITHEA®ISS. The scope of the issues of the ITHEA® ISS covers the area of Informatics and Computer Science. ITHEA® PH welcomes scientific papers and books connected with any information theory or its application.

ITHEA® ISS has four International Journals, established as independent scientific printed and electronic media, published by ITHEA® Publishing House:

- International Journal “Information Theories and Applications” (IJ ITA), since 1993;
- International Journal “Information Technologies and Knowledge” (IJ ITK), since 2007;
- International Journal “Information Models and Analyses” (IJ IMA), since 2012;
- International Journal “Information Content and Processing” (IJ ICP), since 2014.

Every year, ITHEA® ISS organizes many scientific events called **“ITHEA® ISS Joint International Events of Informatics" (ITA**). ITHEA® ISS has organized 135 International Conferences and workshops and has published 45 scientific books*:* monographs, thematic collections and proceedings.

All publications of ITHEA® ISS are accessible freely via ITHEA® main site www.ithea.org as well as via ITHEA® Digital Library (IDR) http://idr.ithea.org.

The great success of ITHEA® International Journals, Books, and Conferences belongs to the whole of the **ITHEA® International Scientific Society.** Membership in ITHEA® ISS is free.

THE ASSOCIATION OF DEVELOPERS AND USERS OF INTELLIGENT SYSTEMS

(http://www.aduis.com.ua/)

Association of Developers and Users of Intelligent Systems (ADUIS) consists of about one hundred members including ten collective members. The Association was founded in Ukraine in 1992. The main aim of **ADUIS** is to contribute to the development and application of the artificial intelligence methods and techniques.

Association has long-term experience in collaboration with teams, working in different fields of research and development. Methods and programs created in Association were used for revealing regularities, which characterize chemical compounds and materials with desired properties.

Conference topics

Theoretical Foundations of Software Engineering

Software Engineering Standards;

Software Development Technologies;

Empirical Software Engineering;

Approaches and Models to Estimate Expenses of Future Software Project;

Software Development Life Cycle Management.

Applied aspects of software engineering

Software Quality;

Approaches to Software Development Life Cycle Processes Improvement;

Requirements Analysis;

Software Designing;

Software maintains and evolution;

AGILE and Model-Driven approaches to Software Development.

Application of software engineering methods and tools

Security and Software Engineering;

Internet and Software Engineering;

Data Bases, Knowledge Bases and Software Engineering;

Software Development for Mobile Operation Systems.

Steering committee

- Abdel-Badeeh M. Salem (Ain Shams University, Cairo, Egypt);
- Albert Voronin (National Aviation University, Ukraine);
- Krassimir Markov (Institute of Information Theories and Applications, Bulgaria);
- Oleksii Pysarchuk (National Aviation University, Ukraine);
- Olena Chebanyuk (National Aviation University, Ukraine);
- Vitalii Velychko (V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine).
- Vladimir Talalaev (National Aviation University, Ukraine)

Program committee

- Antinisca Di Marco (University of L'Aquila, Italy);
- Dmytro Progonov (National Technical University of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Ukraine);
- George Totkov (University of Plovdiv, Bulgaria);
- Juan Bautista Castellanos Peñuela (Universidad Politécnica de Madrid, Spain);
- Lyudmyla Kirichenko (Kharkiv National University of Radio Electronics, Ukraine);
- Michael Gr. Voskoglou (Graduate Technological Educational Institute of Western Greece School of Technological Applications, Greece);
- Mohammad M. Banat (General Chairperson of Mosharaka International Conferences, Jordan);
- Mahmoud Hammad (Jordan University of Science and Technology, Jordan);
- Mykola Guzii (National Aviation University, Ukraine);
- Nagwa Badr (Ain Shams University, Cairo, Egypt);
- Oleh Ilarionov (Taras Shevchenko National University of Kyiv, Ukraine);
- Yuh Wen Chen (Da Yeh University, Taiwan).

Organization committee

- Tetiana Konrad (National Aviation University, Ukraine);
- Valentyna Skalova (National Aviation University, Ukraine):
- Yana Byelozorova (National Aviation University, Ukraine).

TABLE OF CONTENTS

Section: Theoretical Basics of Software Engineering

APPROXIMATE SOFTWARE MODELS' COMPARISON: A FOUNDATION FOR REFINEMENT OPERATION

Olena Chebanyuk, Juan Bautista Castellanos Peñuela

***Abstract**: Software models, represented as UML diagrams are central artifacts in AGILE approach. Approximate comparison of software models is the first step for performing software models refinement. The essence of the approximate comparison of software models proposed in this paper. In order to perform comparison analytical representation of software model is used. Software models represented as graphs*

***Keywords**: Software Engineering, AGILE, UML, Graph.*

Technique of approximate software models' comparison

The analytical foundations of software models' comparison introduced in (Brambilla and Wimmer, 2012) and (Chebanyuk and Abdel-Badeeh M. Salem, 2018). According to (Kolovos et. al., 2006) software models' comparison performed in three stages namely calculation, representation, and visualization. Consider these stages applicable for approximate comparison operation.

Calculation and representation steps

As refinement is a procedure of changing the structure of software model, the comparison should define the preliminary similarities of the both structures. Patterns for matching graphical and analytical representation represented in the Table 1. Comparison of different fragments of an analytical representation (namely $proj$ and ref) allows to substitute new sub-fragment $proj'$ instead of $proj$ fragment. All analytical foundations to perform operations of software models processing introduced in (Chebanyuk, 2018).

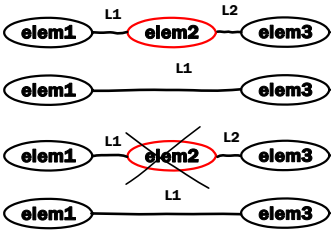
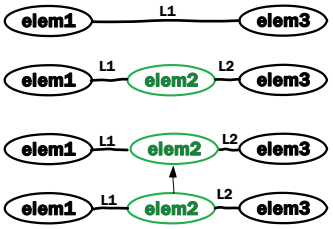
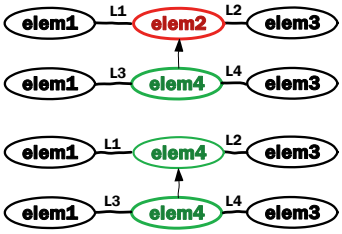
Visualization stage

After comparison lists for different refinement operations for every elementary sub-graph is prepared by means of modifying XMI in which UML diagram is stored. Consider the situation of modifying XMI representation when wrong type element defined.

For modifying XMI corresponding software tool designed. To accomplish this task regularities of XMI file investigated (for XMI produced in JS2UML) and templates for modification of XMI file parts were

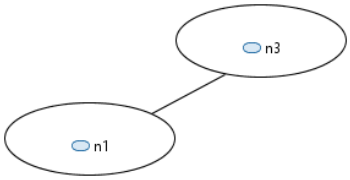
designed. One more feature that make convenient of modifying XMI created namely in JS2UML is that all information about UML diagram elements and their locations and color is stored in the same XMI file.

Table 3. Matching of UML diagrams fragments for its graphical and analytical representation

Type of refinement operation	Graphical representation	Analytical representation
Removing element of project software model		$proj = (elem_1, l_1, elem_2), (elem_2, l_2, elem_3)$ $ref = (elem_1, l_2, elem_2)$
		$proj^{/} = (elem_1, l_1, elem_3)$
Adding element to project software model		$proj = (elem_1, l_1, elem_3)$ $ref = (elem_1, l_2, elem_2), (elem_2, l_2, elem_3)$
		$proj^{/} = (elem_1, l_1, elem_2), (elem_2, l_2, elem_3)$
Changing wrong typed element in project software model		$proj = (elem_1, l_1, elem_3)$ $ref = (elem_1, l_2, elem_2), (elem_2, l_2, elem_3)$
		$proj^{/} = (elem_1, l_2, elem_4), (elem_4, l_2, elem_3)$

Usually in order to store UML diagram as a text two format are used, namely DTD file that contains “metaInformation” about UML diagram. This “metainformation” represents such features as position of UML diagram element, visual characteristics (color of element or color of font, etc.). Information about UML diagram elements and their interconnections is stored in separate XMI file.

Table 2. Analytical, graphical and XMI templates for changing name of object

Analytical representation	Graphical representation	Fragment of XMI for changing
$proj = (n_1, l, n_2)$ $ref = (n_1, l, n_3)$		<UMLUseCase id="0.9950891288870622:UMLUseCase_8" x="45" y="182" width="100" height="40" backgroundColor="#FFFFBB" lineColor="#0CF400" lineWidth="4" fontFamily="monospace" fontSize="12" fontStyle="normal" fontWeight="normal" tagValues=""> <item id="name" fontColor="undefined" value="N3"/> </UMLUseCase>

Consider examples performing refinement operations by means of editing “surface” of JS2UML XMI file in which software model is stored. Example considering the situation when it is necessary to change the name of precedent in Use Case diagram.

Fragments of XMI marked in yellow (Table 2) point what is necessary to change in XMI file to obtain modified fragment of XMI file matching of UML diagrams fragments for graphical and analytical representation.

Conclusion

The essence of the proposed refinement approach is to change fragments of UML diagram, performing operations of changing of its elements that are corresponds to specific domain processes. Project UML diagram compared with domain UML diagrams according to proposed comparison patterns. These comparison patterns allow setting given level of comparison precision.

Refinement is performed by means of consistent changing of objects and links those part of UML diagram fragment that match to some domain process. Firstly, such changing performed in analytical form. Then XMI file, storing UML diagram is modified.

Bibliography

Brambilla, M., Cabot, J., & Wimmer, M. 2012. *Model-driven software engineering in practice.* Synthesis Lectures on Software Engineering, 1(1), 1-182.

Chebanyuk, O. 2018. *An Approach of Text to Model Transformation of Software Models. In ENASE (pp. 432-439).*

Chebanyuk O., Abdel-Badeeh M. Salem, 2018. Formal Foundation, Approach, and Smart Tool for Software Models’ Comparison. *Egyptian Computer Science Journal Vol. 42 No.4., p. 89-102*

Kolovos, D., Paige, R., and Polack, F. 2006. Model comparison: a foundation for model composition and model transformation testing. In *IWGIMM*, pages 13–20.

Authors' Information

Olena Chebanyuk – Assoc. Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine.

E-mail: chebanyuk.elena@ithea.org (chebanyuk.elena@gmail.com)

M**ajor Fields of Scientific Research**: Model-Driven Architecture, Model-Driven Development, Software Architecture, Mobile Development, Software Development

Juan Castellanos *– Artificial Intelligence DPT Universidad Politécnica de Madrid (UPM); Campus Montegancedo, UPM; e- mail:jcastellanos@fi.upm.es.*

M**ajor Fields of Scientific Research**: Artificial intelligence, Natural computation.

APPLICATION OF POSSIBILITY THEORY TO DATA ASSESSMENT

Michael Gr. Voskoglou

Abstract: In large and complex systems that frequently appear nowadays in science, technology and real life problems many different and constantly changing factors are usually involved, the relationships among which are indeterminate. As a result the data obtained from their operation mechanisms cannot be easily determined precisely and in practice estimates of them are used. In the article at hands a method is developed for evaluating a system's fuzzy data in terms of the corresponding fuzzy possibilities.

Keywords*: Fuzzy Data, Fuzzy Set (FS), Fuzzy Relation, Possibility Theory.*

ITHEA Keywords*: G Mathematics and Computing.*

Introduction

While 50-60 years ago the unique tool in hands of the scientists for handling approximate or uncertain data used to be the theory of *Probability*, nowadays the *Fuzzy Set (FS)* theory initiated by Zadeh [Zadeh, 1965] and its extensions and generalizations that followed in the recent years [Voskoglou, 2019]], have given a new dynamic to this field. The article at hands proposes a method for handling fuzzy data with the help of the corresponding membership degrees and possibilities.

That method is developed in the next section, an example on market's research is presented in the third section illustrating it and the article closes with the conclusion obtained.

The reader is supposed to be familiar to the basics of the FS theory and the book [Klir & Folger, 1988] is proposed as a general reference on the subject.

A Method for Handling Fuzzy Data

Assume that one wants to study a system's behavior consisting of n components (objects), $n \geq 2$, during a process involving vagueness and/or uncertainty. Denote by S_i, i=1,2,3 the main steps of that process and by *a, b, c, d, e* the linguistic labels of very low, low, intermediate, high and very high success respectively of the system components in each step. For reasons of simplicity we consider here three steps in the process, but the method holds in general for a finite number of steps.

Set $U = \{a, b, c, d, e\}$. A FS A_i in U will be associated to each step S_i, i = 1, 2, 3. For this, if n_{ia}, n_{ib}, n_{ic}, n_{id}, n_{ie} denote the numbers of the system components that faced very low, low, intermediate, high

and very high success respectively at stage S_i, we define the *membership degree* $m_{A_i}(x)$ of each x in U by $m_{A_i}(x) = \frac{n_{ix}}{n}$ (1) .Then the FS A_i in U associated to S_i is of the form:

$A_i = \{(x, m_{A_i}(x)): x \in U\}, i=1, 2, 3$ (2).

In order to represent all possible *profiles (overall states)* of the system components during the corresponding process a *fuzzy relation*, say R, in U^3 (i.e. a FS in U^3) is considered of the form:

$R= \{(s, m_R(s)): s=(x, y, z) \in U^3\}$ (3).

Assuming that the performance of each component at each step does not depend on its performance in the previous steps we define the membership degrees of the several profiles by the equation

$m_R(s) = m_{A_1}(x) \cdot m_{A_2}(y) \cdot m_{A_3}(z)$ (4).

According to the British economist Shackle [3] and many other researchers after him, the behavior of a system involving the human presence can be better studied by using the possibilities rather of the several profiles, than their probabilities. For reasons of brevity, we shall write here m_s instead of $m_R(s)$. Then the *possibility* r_s of the profile s is defined by $r_s = \frac{m_s}{\max\{m_s\}}$ (5).

In equation (5) $max\ \{m_s\}$ denotes the greatest value of m_s for all s in U^3. In other words the possibility of s expresses the "relative membership degree" of s with respect to $max\ \{m_s\}$.

Frequently in practical applications it becomes necessary to study the *combined results* of the behaviour of k different groups of a system's components, $k \geq 2$, during the same process (e.g. the combined performance of two or more student classes in solving the same problems). For measuring the degree of evidence of the combined results of the k groups, it becomes necessary to define the *combined possibility* $r(s)$ of each profile s with respect to the membership degrees of s in all the groups involved. The values of $r(s)$ can be defined in this case with respect to the corresponding *pseudo-frequencies* $f(s) = \sum_{t=1}^{k} m_s(t)$ (6) and they are equal to $r(s) = \frac{f(s)}{\max\{f(s)\}}$ (7) , where $max\{f(s)\}$ denotes the maximal pseudo-frequency.

Application to Market's Research

A company performed a market's research about the degree of the consumer preference for its negotiable products, which was characterized by the fuzzy linguistic labels *a, b, c, d, e.* The research was performed separately for men and women and for three different categories of age, namely C_1: 18-30 years, C_2: 31-50 years and C_3: over 50 years old.

Denote by A_1 (t), A_2 (t) and A_3 (t) respectively the FSs representing the consumers' degree of preference for each of the above three categories of age, where the variable t takes the values t = 1 for

men and t = 2 for women. FSs of such type, whose entries depend on the values of a variable, are usually referred as *fuzzy variables.*

Table 1: Customers' profiles with non zero pseudo-frequencies

A_1	A_2	A_3	$m_s(1)$	$m_s(2)$	$f(s)$	$r(s)$
b	*b*	*b*	0	0.016	0.016	0.092
b	*a*	*b*	0	0.012	0.012	0.069
b	*c*	*b*	0	0.032	0.032	0.184
b	*b*	*a*	0	0.021	0.021	0.121
b	*b*	*c*	0	0.016	0.016	0.092
b	*a*	*a*	0	0.016	0.016	0.092
b	*a*	*c*	0	0.012	0.012	0.069
b	*c*	*a*	0	0.042	0.042	0.241
b	*c*	*c*	0	0.032	0.032	0.184
c	*c*	*c*	0.072	0.080	0.152	0.874
c	*a*	*c*	0.082	0.030	0.112	0.644
c	*b*	*c*	0.031	0.040	0.071	0.408
c	*d*	*c*	0.046	0	0.046	0.264
c	*c*	*a*	0.067	0.107	0.174	1
c	*c*	*b*	0.056	0.008	0.064	0.368
c	*a*	*a*	0.028	0.040	0.068	0.391
c	*a*	*b*	0.024	0.030	0.054	0.310
c	*b*	*a*	0.028	0.053	0.081	0.466
c	*b*	*b*	0.024	0.040	0.064	0.368
c	*d*	*a*	0.043	0	0.043	0.247
c	*d*	*b*	0.036	0	0.036	0.207
d	*d*	*a*	0.020	0	0.020	0.115
d	*d*	*b*	0.017	0	0.017	0.098
d	*d*	*c*	0.022	0	0.022	0.126
d	*a*	*a*	0.013	0.024	0.037	0.213
d	*a*	*b*	0.011	0.018	0.029	0.167
d	*a*	*c*	0.015	0.018	0.033	0.190
d	*b*	*a*	0.013	0.032	0.045	0.259
d	*b*	*b*	0.011	0.024	0.035	0.201
d	*b*	*c*	0.014	0.024	0.038	0.218
d	*c*	*a*	0.031	0.064	0.095	0.546
d	*c*	*b*	0.026	0.048	0.074	0.425
d	*c*	*c*	0.034	0.048	0.082	0.471
e	*a*	*a*	0.017	0	0.017	0.098
e	*a*	*b*	0.014	0	0.014	0.080
e	*a*	*c*	0.018	0	0.018	0.103
e	*b*	*a*	0.017	0	0.017	0.098
e	*b*	*b*	0.014	0	0.014	0.080
e	*b*	*c*	0.018	0	0.018	0.103
e	*c*	*a*	0.039	0	0.039	0.224
e	*c*	*b*	0.033	0	0.033	0.190
e	*c*	*c*	0.042	0	0.042	0.241
e	*d*	*a*	0.025	0	0.025	0.144
e	*d*	*b*	0.021	0	0.021	0.121
e	*d*	*c*	0.027	0	0.027	0.155

According to the collected data the FSs A_i (t), for i = 1, 2, 3 and t = 1, 2 were found to be the following:

A_1 (1) = {(*a*, 0), (*b*, 0), (*c*, 0.486), (*d*, 0.228), (*e*, 0.286)}, A_2 (1) = {(*a*, 0.171), (*b*, 0.171), (*c*, 0.4), (*d*, 0.257), (*e*, 0)}, A_3 (1) = {(*a*, 0.343), (*b*, 0.0286), (*c*, 0.371), (*d*, 0), (*e*, 0)}.

A_1 (2) = {(*a*, 0), (*b*, 0.2), (*c*, 0.5), (*d*, 0.3), (*e*, 0)}, A_2 (2) = {(*a*, 0.2), (*b*, 0.267), (*c*, 0.533), (*d*, 0), (*e*, 0)}

A_3 (2) = {(*a*, 0.4), (*b*, 0.3), (*c*, 0.3), (*d*, 0), (*e*, 0)}.

The calculation of the membership degrees of all the customer profiles is done by the product law of equation (4). For example, for the profile *s* = (*c, c, a*) one finds that

$m_s(1) = 0.486 \times 0.4 \times 0.343 \approx 0.67$ and $m_s(2) = 0.5 \times 0.5 \times 0.33 \approx 0.107$ and similarly $m_s(2) = 0.107$.

It turns out that the above profile has the greater pseudo-frequency $f(s) = 0.67 + 0.107 = 0.174$ and therefore its combined possibility is equal to 1, while the combined possibilities of all the other profiles are calculated by $r(s) = \frac{f(s)}{0.174}$.

The membership degrees, the pseudo-frequencies and the combined possibilities of all the customer profiles with nonzero pseudo-frequencies are presented in Table 1, where all calculations have been made with accuracy up to the third decimal point. The fuzzy data of Table 1 give to the company a detailed view about the consumers' preferences for its products.

Conclusion

The general character of the proposed in this work method for tackling fuzzy data enables its application to a variety of other human and machine activities and this is one of our main targets for future research.

References

[Klir, G. & Folger, 1988] Klir, G. J. & Folger, T.A., Fuzzy Sets, Uncertainty and Information, London, Prentice Hall Int., 1988.

[Shackle, 1961] Shackle, G.L.S., Decision, Order and Time in Human Affairs, Cambridge, Cambridge Univ. Press, 1961.

[Voskoglou, 2019] Voskoglou, M. Gr., Multi-Valued Logics: A Review, International Journal of Applications of Fuzzy Sets and Artificial Intelligence, 9, 5-12, 2019.

[Zadeh, L.A]., Fuzzy Sets, *Information and Control*, 8, 338-353, 1965.

Author's Information

Michael Gr. Voskoglou *– Graduate Technological Educational Institute (TEI) of Western Greece; Professor Emeritus of Mathematical Sciences; 26334 Patras, Greece e-mail: mvosk@hol.gr*

Major Fields of Scientific Research: Algebra, Artificial Intelligence, Fuzzy Logic, Markov Chains, Mathematics Education

SOFTWARE QUALITY: THEORETICAL FOUNDATIONS AND TESTING TOOLS

Vladimir Talalaev

***Abstract**: Paper touches questing related to measuring software quality from different point of view starting from the theoretical foundations to possibilities of software testing tools that are developed following principles to family of ISO/IEC/IEEE 29119 standards.*

***Keywords**: Software Engineering, AGILE, Testing, ISO/IEC/IEEE 29119.*

Introduction

To understand the landscape of software quality it is central to answer the so often asked question: *what is quality*?

1) **Conformance to specification**: Quality that is defined as a matter of products and services whose measurable characteristics satisfy a fixed specification – that is, conformance to an in beforehand defined specification.

2) **Meeting customer needs**: Quality that is identified independent of any measurable characteristics. That is, quality is defined as the products or services capability to meet customer expectations – explicit or not (Al-Badareen, 2011).

Software Quality Testing Standards

ISO/IEC/IEEE 29119 Software Testing is an internationally agreed set of standards for software testing that can be used within any software development life cycle and by any organization. Table given below matches standards to different software development lifecycle processes in AGILE approach.

Table 1. Matching ISO/IEC standards to software development life cycle process

ISO/IEC standard	Software development life cycle process	Reason of matching
ISO/IEC 29119-1:2013 - Concepts & Definitions	Testing	Defining main concepts and defining testing processes
ISO/IEC 29119-2:2013 - Test Processes	Testing	Managing Test-Driven Development processes and different types of testing

ISO/IEC standard	Software development life cycle process	Reason of matching
ISO/IEC 29119-3:2013 - Test Documentation	Requirement Analysis Testing	Software models replace documentation that are main artifacts in AGILE approach
ISO/IEC 29119-4:2015 - Test Techniques	Requirement analysis Development Testing	Suitable for Test-Driven Development approach
ISO/IEC 29119-5:2016 - Keyword Driven Testing	Requirement Analysis Testing	Keywords serves as tags for tracing procedure

To support static testing, the standard Work Product Reviews ISO/IEC 20246:2017 has been defined. To support adoption of ISO/IEC/IEEE 29119-2, the following process assessment standard has been defined:

ISO/IEC 33063:2015 - Process Assessment Model for Software Testing

The ISO/IEC/IEEE 29119 standards replace a number of existing software testing standards, which are now retired: IEEE 829 Test Documentation, IEEE 1008 Unit Testing, BS 7925-1 Vocabulary of Terms in Software Testing, BS 7925-2 Software Component Testing Standard.

Software testing tools analysis

The next point of software quality investigation is to analyzing of theoretical principles that become a basic for testing tools designing.

Table 2. Criterion for testing tools effectiveness

FEATURES	**SELENIUM**	**TEST COMPLETE**	**QUCIK TEST PROFESSIONAL**
Sphere of usage	Web Applications, and plug-ins	All types of applications.	A client server application, add-ons, (purchase license).
Licensing	Open source tool (no licensing).	$2K Enterprise Seat License	Licensed and very Expensive, Ten user license costs approx. 60L.
Support for operating system/platforms	Windows PC/MAC/ UNIX.	Windows operating systems.	Windows XP, Windows 7
Required programming skills	Beginner or intermedia.	From beginner to high qualified programmer.	Beginner modifying the scripts

FEATURES	SELENIUM	TEST COMPLETE	QUCIK TEST PROFESSIONAL
Object Oriented Language support and Scalability	Supports Java, .Net, Perl, PHP, Python and Ruby.	Test Complete supports scripting in VBScript, JSScript, DelphiScript, C++Script and C#Script, so you can create scripts in the language.	Scripts can be developed only in VBScript or JavaScript.
Database applications testing	Supports	Supports	Supports
Report Generation	Practically no	Supports report generation as well as dump file generation	Supports

Conclusion

The analysis of factors influencing to software quality performed in this paper. Analysis of different software quality standards allows formulating criterions of software testing tools effective working. According to these criterions, different software testing tools were analyzed. Comparative table allows choosing the best software tool for different variants of adopting software development process as well as estimate new software testing tool for using it in AGILE approach.

Bibliography

Al-Badareen, A. B., Selamat, M. H., Jabar, M. A., Din, J., & Turaev, S. (2011, June). Software quality models: A comparative study. In International Conference on Software Engineering and Computer Systems (pp. 46-55). Springer, Berlin, Heidelberg.

Kaur, Harpreet, and Gagan Gupta. "Comparative study of automated testing tools: selenium, quick test professional and testcomplete." Harpreet kaur et al Int. Journal of Engineering Research and Applications ISSN (2013): 2248-9622.

Kaur, Manpreet, and Rupinder Singh. "A Review of software testing techniques." International Journal of Electronic and Electrical Engineering 7.5 (2014): 463-474.

Authors' Information

Vladimir Talalaev – Assoc. Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine. **E-mail:** VATalalaev@gmail.com)

M**ajor Fields of Scientific Research**: Ontology-Driven Architecture, Risk management, Software Architecture, Software Project management

A WEIGHT-BASED APPROACH FOR ENHANCING OF THE QUALITY OF AUTOMATICALLY GENERATED TEST SCENARIOS FOR SOFTWARE SYSTEMS

Alexander Kolchin, Stepan Potiyenko

V.M. Glushkov Institute of cybernetics NAS of Ukraine, Kyiv, Ukraine

***Abstract**: This paper describes a method of quality scores improvement of each individual test scenario as well as whole test suite. Weight-based heuristic method for automatic test suite generation is proposed. In many practical cases the method allows to achieve better results in same time budget and test suite size.*

***Keywords**: model-based testing, data flow, tests quality.*

Introduction

Testing is the most used way of assessing quality in software industry [Myers 2004]. The industry moves toward model-based development, and automated test generation from the model is often considered as a form of requirements-based testing. This approach is cost effective because it can be used to automate the generation of test cases and dramatically reduce the cost of test maintenance.

Generating complete set of possible paths is too labor-expensive and unrealizable: number of paths grows up exponentially with number of branches and becomes even infinite in programs with loops, so the majority of test generation approaches use some structural coverage criterion based on a behavioral model of the system under test (SUT) to guide the selection of test cases. However, there is some evidence [Rushby 2004, Staats 2012, Inozemtseva 2014] that simply using the coverage provision as a target for automated test generation may be a flawed strategy: coverage metrics are intended to measure the thoroughness of human-produced tests, and do not necessarily lead to good test sets when used in an inverted role as a specification for the tests required. Common drawback of the automatic approach is that the test scenario generated is not necessarily a good exercise regarding the verified property: a counterexample path is an artifact of search strategy, so it may terminate at the middle of some interesting behavior, may include a lot of redundancy, cause-effect relations are obscure and intricate, such tests often lack semantic content, their purpose is unclear.

In this paper, the test quality metrics will be considered in addition to level of coverage – scores of data cohesion and coverage overlapping. The main goal is to achieve higher scores of quality staying at

same level of computational complexity and resulting test suite size, making it a complementary optimization to existing test generation approaches.

Background

Software industry uses some structural criterion to estimate coverage. For example, Modified Condition/Decision coverage (MC/DC) is required by standard DO-178B. However, many empirical studies [Staats 2012, Inozemtseva 2014, Palomba 2016] suggest that coverage, while useful for underlying under-tested parts of a program, should not be used as a quality target because it is not a good indicator of test suite effectiveness.

Data-flow testing is less popular, however is very important because it augments control-flow testing criteria and concentrates on how a variable is defined and used in the model, which could lead to more efficient and targeted test suites. The occurrences of a variable in a model can be associated with its definition (def – location of a statement, where a value is assigned to a variable), and usage (use – location of a statement, where the value of a variable is read). A def-clear path with respect to a variable if it contains no new definition of that variable. [Rapps and Weyuker 1982] propose criteria for data flow testing. Test suite satisfies the all-uses criterion (most practical), meaning that each computation and condition effected by each definition will be tested.

Fault sensitivity analysis is often based on the idea of fault-based testing (or mutation analysis), which assumes that the intended implementation closely matches the actual implementation. Mutation analysis is widely used in order to compare the fault detection success of test sets. Mutants are formed by making specific changes introducing a “mistake” on the original model. A test set is said to kill a mutant, if it fails during execution simulation on model with that mutant.

Methods of test suites generation

Many different approaches were invented: stochastic and combinatorial methods are easy to implement and have high performance, but result in poor coverage and high redundancy; search-based methods (e.g., genetic algorithms) [Vivanti 2013, Palomba 2016] use approximation function for measuring the "distance" between the generated test population and the desired test coverage objectives, and tend to minimize it; many systems implement hybrid and heuristic approaches. Systematic methods (like model checking, constraint solving, symbolic execution [Cadar 2013, Kolchin 2018] etc.) extract constraints for executing program paths based on some code analysis, and obtain test data that direct the execution of the program along these paths. Advantage of the systematic methods is a possibility to detect behaviors associated with very small set of input values, the drawback is inefficient state space traversal – algorithms often stuck in enumerating of paths without detection of new coverage. Conventional approaches to test suite generation mainly focus on coverage as a unique goal to achieve, which result

in typical shortcomings: lack of cohesion (tests often lack semantic content) and overlapping (some parts will be exercised more than thoroughly, but another – just in a single test case) problem.

Weight-based test suite quality improving heuristic

Search heuristics is the most popular mechanism to guide program exploration to the most promising paths in the program [Neetu 2019]. Level of data cohesion inside of each test case and score of overlapping of test cases inside whole suite are preferred practical parameters for heuristic weighting factor [Palomba 2016]. We compute average data cohesion score as a mean number of def-use pairs each test case covers; overlapping score is computed as follows: overlapping score(tc_a) = Σ sim(tc_a, tc_k) / (n-1), $tc_a \neq tc_k$, where tc_a is the test case under analysis, T = {tc_1, …, tc_n} is the test suite, n is size of T, and sim (tc_a, tc_k) denotes cosine textual similarity between tc_a and another test case $tc_k \in$ T. The resulting value of the score lies in [0,1], the higher score means the higher the overlapping between tc_a and other tests. Key insight for increasing cohesion inside individual tests is the observation of def-use coverage level of inherent (w.r.t. particular coverage item) variables in hand-made tests: the variables most likely to be in a backward slice of the coverage item, and test designed to cover it will also cover appropriate du-pair around the item, e.g., it will cover some branch in cause-effect graph of the SUT.

```
For each coverage item c in Coverage Goal,
  Construct Backward Slice(c);
  Identify set of defs(c), which belongs
  to Backward Slice(c).

During search for coverage item c,
  At each location with non-deterministic choice
  of transitions alternatives {t1, …, tn},
    For each ti ∈ {t1, …, tn},
      Compute DU-priority(ti) := number of reachable
        uses of active definitions from def(c);
      Compute Overlap Level(ti) := number of
       occurrences of transition ti in whole test set
       produced so far;
    Re-order the set of transitions w.r.t DU-priority
and then apply Overlap Level inside subsets with equal
score of DU-priority.
```

Fig. 1. Data-flow–based prioritization principle

```
For each coverage item c in Coverage Goal,
  Construct Forward Slice(c);
  Identify set of Observation Points, which belongs
  to Forward Slice(c).

During prolongation of path to coverage item c,
  At each location with non-deterministic choice
  of transitions alternatives {t1, …, tn},
    For each ti ∈ {t1, …, tn},
      Compute Observation Proximity := number of
        transitions on control flow graph to
         the nearest Observation Point.
      Compute DU-priority(ti) := number of reachable
        uses of active definitions from def(c);
      Compute Overlap Level(ti) := number of
       occurrences of transition ti in whole test set
       produced so far;
    Re-order the set of transitions w.r.t Observation
Proximity, and then apply DU-priority and Overlap Level
inside subsets with equal score of Observation
Proximity.
```

Fig. 2. Data- and Control-flow–based prolongation principle

After the coverage item reached, the scenario-under-construction is prolonged up to some observation point. In order to observe effect of the item, the set of such points will be chosen from ones belonging to forward slice of the covered item. Figure 2 presents the prolongation principle.

Test suite may include a lot of redundancy, as each test likely covers several coverage items. Special techniques which select a subset of tests from the available suite that exercise the given set of items are applied to minimize the test suite size. Each test scenario produced will include information about coverage items it covers together with overall def-use score. In order to increase test suite efficiency, mutation-based score is used to assess test scenario during test suite minimization.

Based on the suggested methods an experimental prototype was implemented. It is a largely complementary approach, which extends an ordinary state space traversal routine with weight-based prioritization heuristic. For the study, 7 different medium-sized models were used, which are prototypes of real systems of different domains – telecom, automotive and finance.

Conclusion

Empirical results indicate that prioritization during the test suite generation search makes positive effect on level of data cohesion of each test and level of overlapping in whole test suite; mutation analysis scores improvement is achieved mainly due to the post-generation minimization procedure: in more than 30% of tests the mutation-based assessment plays decisive role in comparison. Overall test suite size is not affected by the prioritization while it's ability to reveal faults is improved.

References

[Myers 2004] G.J. Myers The Art Of Software Testing. N.Y. John Wiley & Sons, Inc. 2004.

[Rushby 2004] Rushby J. Automated test generation and verified software // Verified Software: Theories, Tools, Experiments. 2008. pp. 161-172.

[Staats 2012] Staats M. and oth. On the danger of coverage directed test case generation // LNCS 7212. pp. 409-424. 2012.

[Inozemtseva 2014] Inozemtseva L., Holmes R. Coverage is not strongly correlated with test suite effectiveness // In Proc. of ACM ICSE'14. 2014. pp.435-445.

[Rapps and Weyuker 1982] Rapps S., Weyuker E. Data flow analysis techniques for test data selection // In proc. of Int. Conf. of Softw. Eng. 1982. pp. 272-277.

[Palomba 2016] Palomba F., Panichella A., and oth. Automatic Test Case Generation: What if Test Code Quality Matters? // In proc. of Int. Symp. on Software Testing and Analysis. 2016. pp.130-141.

[Vivanti 2013] Vivanti M., Mis A. and oth. Search-based data-flow test generation // In IEEE Int. Symp. on Software Reliability Engineering. 2013. pp.10.

[Cadar 2013] Cadar C., Sen K. Symbolic execution for software testing: three decades later // Commun. ACM. 2013. Vol. 56(2). pp. 82-90.

[Kolchin 2018] Kolchin A. A novel algorithm for attacking path explosion in model-based test generation for data flow coverage // Proc. of IEEE 1st Int. Conf. on System Analysis and Intelligent Computing, SAIC. 2018.

[Neetu 2019] Neetu J., Rabins P. Automated Test Data Generation Applying Heuristic Approaches—A Survey // Software Engineering. 2019. pp.699-708.

COMPARISON OF WEB SERVICES CREATION TECHNOLOGIES USING TESTING APPLICATION SOAPUI

Yurii Milovidov

Abstract. *A comparative analysis of the technologies of Web services and multiple platforms - Technology MS Visual Studio.NET and Java-technology was provided. For web services testing was used application with open source SoapUI. Each of these technologies has its advantages and disadvantages. The use of specific technology is determined by specific conditions.*

Key words: *web service, web technology, MS Visual Studio, Apache Axis, SOAP, WSDL, SoapUI.*

ITHEA keywords: *D.2 Software engineering, D.2.1 Requirement/Spesification, D.2.13 Reusable Software.*

Introduction

There are many of non-commercial and commercial Web service development tools. All of them differ in degree of completeness, possibilities, conditions of licensing, requirements to software and hardware, etc.

The main advantages of web-services:

• Web services provide interoperability of software systems regardless of platform.

• Web services are based on open standards and protocols. Thanks to the use of XML, it is easy to develop and configure web services.

The purpose of the study: based on the findings of the comparison of various technologies for the creation and use of web-services, provide useful information to developers who want to implement and deploy their web-services.

Theoretical foundations

The term "web service" describes a standardized way of integrating web-based applications using the XML, SOAP, WSDL and UDDI open standards over an Internet Protocol backbone. XML is the data format used to contain the data and provide metadata around it, SOAP is used to transfer the data, WSDL is used for describing the services available and UDDI lists what services are available.

A web service is a method of communication between two electronic devices over a network.

Different software may use different programming languages, and hence there is a need for a method of data exchange that doesn't depend upon a particular programming language. Most types of software can, however, interpret XML tags. Thus, web services can use XML files for data exchange.

Functional and non-functional testing of web services is done with the help of WSDL parsing. Regression testing is performed by identifying the changes made to upgrade a software. Web service regression testing needs can be categorized in three different ways, namely, changes in WSDL, changes in code, and selective re-testing of operations. We can capture the above three needs in three intermediate forms of Subset WSDL, namely, Difference WSDL (DWSDL), Unit WSDL (UWSDL), and Reduced WSDL (RWSDL), respectively. These three Subset WSDLs are then combined to form Combined WSDL (CWSDL) that is further used for regression testing of the web service. This will help in Automated Web Service Change Management (AWSCM), by performing the selection of the relevant test cases to construct a reduced test suite from the old test suite.

Web services testing can also be automated using several test automation tools like SOAP UI, Oracle Application Testing Suite (OATS), Unified Functional Testing, Selenium, etc.

Related works

Testing of Web Services using Behavior-Driven Development [Oruç, A. and Ovatman, T. 2016. The task of this paper is to study the way the proposed approach of using Gherkin language and JMeter tool to define and run test cases for Web Services improves the productivity of these operations compared to commonly used methods.

A Linear Logic based Synchronization Rule for Deadlock Prevention in Web Service Composition [Oliveira, V. 2017] The task of this paper is to propose an approach based on Linear Logic to deal with deadlock situations in the WS composition modeled by Petri nets, where the workflow modules are deadlock-free but not necessarily the composed system.

Semantic Networks – Based Approach for Web Services Management [Nacer, H. and Iggui, T. 2016] The task of this paper is to propose an approach of Web services management in order to facilitate the Web service composition. This work uses an ontology-based approach to organize the UDDI into domains.

Task

By comparing different technologies for creating and using web-services, provide useful information to developers who want to implement and deploy their web-services. By using this information, you can understand which technology is best for use.

Reaearch methods

The MS Visual Studio .NET technology and Java technology were chosen to conduct a comparative analysis of technologies for creating web-services.

The implementation of the creation of web services by means of MS Visual Studio .NET is the ASP.NET Web Service technology and the deployment of web services on the IIS server.

The implementation of the creation of web-services by means of Java is the Apache Software Foundation project called AXIS (Apache eXtensible Interaction System).

For web services testing we used application with open source SoapUI SoapUI is an open-source web service testing application for service-oriented architectures (SOA) and representational state transfers (REST). Its functionality covers web service inspection, invoking, development, simulation and mocking, functional testing, load and compliance testing.

It is free software. Since the initial release, SoapUI has been downloaded more than 2,000,000 times.

Results

The choice of development tools are influenced by their prevalence, accessibility, ease of deployment, programmer's passion.

MS Visual Studio

• This is a paid product.

• Long installation of the product.

• It takes a lot of disk space.

• Provides quick and easy deployment of web services

AXIS and Apache Tomcat

• Apache Tomcat is widespread

• Available for free (http://tomcat.apache.org/)

• Quickly put.

• Works on any platform where Java is located

• Famous for its reliability

We tested the time of web services, conducted 100 tests and calculated the average time. Test results are presented in Table 1.

Table 1 Comparison of work time and stability of web-services

	MS Visual Studio, IIS	Axis, Apache
Response time	0.015 sec	0.016 sec
Time of work, returning records from one table of DB (about 30,000 records)	0,41 sec	0,88 sec
Time of work, accepting as a parameter a large array of values.	0,9 sec	4,3 sec
Time of data transfer from web-service to client	0,004 sec	0,01 sec
Stability of web-service to many connections	287 connections	568 connections

From the comparisons, we see that the web service created with .Net and deployed on the web server IIS works faster, in less time returns data than a web-service implemented with Apache Axis. Moreover, the time difference increases with the increase in the amount of transferred values from the client to the service. But the web service on IIS loses on the stability of the set of connections.

Conclusions

By comparing .Net and Java technology to create and use web services, you can draw the following conclusions:

Each of these technologies has its advantages and disadvantages. It can be concluded that the application of specific technology is determined by specific conditions.

If the developer is not familiar with any of these technologies, then it is naturally better to develop in technology already familiar. If the developer is not familiar with any of these technologies, then in my opinion it's easier and quicker to figure out and create a web-service using MS Visual Studio.Net.

References

[Oruç, A. and Ovatman, T. 2016] Testing of Web Services using Behavior-Driven Development, Proceedings of the 6th International Conference on Cloud Computing and Services Science (CLOSER 2016) - Volume 2, pages 85-92 ISBN: 978-989-758-182-3 – Access mode: https://www.scitepress.org/Papers/2016/59114/59114.pdf

[Oliveira, V., Julia, S., Passos, L. and Oliveira, K. 2017] A Linear Logic based Synchronization Rule for Deadlock Prevention in Web Service Composition, Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017) - Volume 2, pages 316-323 ISBN: 978-989-758-248-6 – Access mode: https://pdfs.semanticscholar.org/05a2/0725cbb574b6391ba9257730ab32803ced53.pdf

[Nacer, H. and Iggui, T. 2016] Semantic Networks – Based Approach for Web Services Management, Proceedings of the 11th International Joint Conference on Software Technologies (ICSOFT 2016) - Volume 2: ICSOFT-PT, pages 95-99 ISBN: 978-989-758-194-6 – Access mode: https://pdfs.semanticscholar.org/7832/5eb6408fecbf5c0eb65386c942b75f96212f.pdf

[Irum Rauf ; Inna Vistbakka ; Elena Troubitsyna 2018] Formal Verification of Stateful Services with REST APIs Using Event-B, Published in: 2018 IEEE International Conference on Web Services (ICWS) ISBN: 978-1-5386-7247-1 – Access mode: https://ieeexplore.ieee.org/document/8456341

Authors' Information

Yurii Milovidov, National University of Life and Enironmental Sciences of Ukraine, Kyiv, scientific Department of Programming Technologies. Senior Lecturer.

Email: milovidov.email.ua

Programming technologies, web-design, Internet – technologies.

ABOUT TWO STEREOTYPES OF UML PROFILE FOR MULTIDIMENSIONAL NUMBERED DATA BASES

Olena Chebanyuk, Krassimira Ivanova, Vitalii Velychko, Krassimir Markov

***Abstract**: Two stereotypes for Numbered Space (NSpace) and Space with Natural Language Addressing (NLASpace) from an UML profile for modeling Multi-Dimensional Numbered Data Bases (MDNDB™) are shortly outlined in this paper. MDNDB™ is a tool for storing Big Data locally as well as in the cloud. Because of this, it is important to have appropriate modeling language to support the design of practical applications.*

***Keywords**: MDNDB, UML profile for modeling multi-dimensional numbered data bases.*

Introduction

The independence of dimensionality limitations is very important for developing new software systems aimed to process large volumes of high-dimensional data. To achieve this, we need information models and corresponding access methods to cross the boundary of the dimensional limitations and to obtain the possibility to work with large information spaces with variable and practically unlimited number of dimensions. A step in developing such methods is the **Multi-Domain Information Model (MDIM)** introduced in [Markov, 2004]. Below we will remember main idea of MDIM and its using for so called "Natural Language Addressing". After that, two UML stereotypes from an UML profile for numbered information bases are outlined. Finally, a short example and conclusion are given.

Basic structures of Multi-domain information model (MDIM)

Main structures of MDIM are *basic information elements, information spaces, indexes, meta-indexes* and *aggregates*. The definitions of information elements and spaces structures are remembered below [Markov, 2004].

The basic information element (***BIE***) of MDIM is an arbitrary long string of machine codes (bytes). When it is necessary, the string may be parceled out by lines. The length of the lines may be variable.

Let the universal set ***UBIE*** be the set of all ***BIE***. Let $\boldsymbol{E}_1$ be a set of basic information elements. Let μ_1 be a function, which defines a biunique correspondence between elements of the set $\boldsymbol{E}_1$ E_1 and elements of the set $\boldsymbol{C}_1$ $C_1$of positive integer numbers, i.e.: $\boldsymbol{E}_1 = \{e_i \mid e_i \in \boldsymbol{UBIE}\ ,\ i=1,\ldots, m_1\}$, $\boldsymbol{C}_1 = \{c_1 \mid c_i \in N,$

i=1,…,m_1}, μ_1 $E_1 \leftrightarrow C_1$. The elements of $\boldsymbol{C}_1$ are said to be numbers (**co-ordinates**) of the elements of E_1. The triple $\boldsymbol{S_1 = (E_1, \mu_1, C_1)}$ is said to be a ***numbered information space of level 1*** (one-dimensional or one-domain information space). The triple $\boldsymbol{S_2 = (E_2, \mu_2, C_2)}$ is said to be a ***numbered information space of level 2*** (two-dimensional or multi-domain information space of level two) iff the elements of $\boldsymbol{E}_2$ are numbered information spaces of level one (i.e. belong to the set $\boldsymbol{NIS}_1$) and μ_2 is a function which defines a biunique correspondence between elements of $\boldsymbol{E}_2$ and elements of the set C_2 of positive integer numbers, i.e.: $\boldsymbol{E}_2 = \{e_i \mid e_i \in \boldsymbol{NIS_1}$, i=1,…, $m_2\}$, $C_2 = \{c_i \mid c_i \in N$, i=1,…,$m_2\}$, $\mu_2 : E_2 \leftrightarrow C_2$.

The triple $\boldsymbol{S_n = (E_n, \mu_n, C_n)}$ is said to be a ***numbered information space of level n*** (n-dimensional or multi-domain information space) iff the elements of $\boldsymbol{E}_n$ are numbered information spaces of level n-1 (set $\boldsymbol{NIS}_{n-1}$) and μ_n is a function which defines a biunique correspondence between elements of $\boldsymbol{E}_n$ and elements of the set C_n of positive integer numbers, i.e.: $\boldsymbol{E}_n = \{e_j \mid e_j \in \boldsymbol{NIS_{n-1}}$, j=1,…, $m_n\}$, $C_n = \{c_j \mid c_j \in N$, j=1,…,$m_n\}$. $\mu_n : E_n \leftrightarrow C_n$. The information space $\boldsymbol{S}_n$, which contains all information spaces of a given application is called ***information base*** of level **n**.

Operations in the MDIM: After defining the information structures, we need to present the operations, which are admissible in the model. In MDIM, we assume that **all** information elements of **all** information spaces **exist**. If for any $S_i : E_i = \emptyset \wedge C_i = \emptyset$, than it is called ***empty***. Usually, most of the information elements and spaces are empty. This is very important for practical realizations.

Operations with basic information elements: Because of the rule that all structures exist, we need only two operations with a BIE: (1) Updating; (2) Getting the value. Updating, or simply – ***writing*** the element, has several modifications with obvious meaning: Writing as a whole; Appending/inserting; Cutting/replacing a part; Deleting. There is only one operation for getting the value of a BIE, i.e. ***read*** a portion from a BIE starting from given position. We may receive the whole BIE if the starting position is the beginning of BIE and the length of the portion is equal to the BIE length. Special kind of operations concerns the navigation in a space. We may receive the space address of the **next** or **previous, empty** or **non-empty** elements of the space starting from any given co-ordinates. The possibility to **count** the number of non empty elements of a given space is useful for practical realizations.

Natural Language Addressing (NLA)

In this research we follow a proposition of Krassimir Markov to use the computer encoding of name (concept) letters as logical address of connected to it information. This way no indexes are needed and high speed direct access to the text elements is available. This approach us called “Natural Language Addressing” [Markov et al, 2015]. Let remember a short example: let have the next definition: “London: The capital city of England and the United Kingdom, and the largest city, urban zone and metropolitan area in the United Kingdom, and the European Union by most measures”. In the computer memory, for

example, it may be stored in a file at relative address “00084920” and the index couple is: (“London”, “00084920”). At the memory address “00084920” the main text, “The capital … measures.” will be stored. To read/write the main text, firstly we need to find name “London” in the index and after that to access memory address “00084920” to read/write the definition.

If we assume that name “London” in the computer memory is encoded by six numbers (letter codes), for instance by using ASCII encoding system London is encoded as (76, 111, 110, 100, 111, 110), than we may use these codes for direct address to memory, i.e. (“London”, “76, 111, 110, 100, 111, 110”).

For human this address will be shown as “London”, but for the computer it will be “76, 111, 110, 100, 111, 110”. The tool for storing and accessing information using Natural Language Addressing is just the Multidimensional Numbered Information Spaces and corresponded Data Bases.

UML stereotypes for multidimensional information spaces

Two stereotypes for Numbered Space (NSpace) and Space with Natural Language Addressing (NLASpace) are shortly outlined below. Extended explanation will be given in further publications.

NLASpace is very simple class. No attributes are needed. NSpace stereotype is more complex - we need attributes for level of the space and some constraints if needed.

The mainy two functions are update and read using NL address (Figure 1) or coordinates’ address (Figure 2). Additional functions for navigation in the space are available for both kind of spaces. For NSpace some constraints to structure the space may be used.

<<NLASpace>>
NLAWrite NLARead NLANext non-empty NLAPrevious non-empty NLANext empty NLAPrevious empty

Figure 1

<<NSpace>>
Level {Constraints}
NWrite NRead NNext non-empty NPrevious non-empty NNext empty NPrevious empty

Figure 2

Example

To model the structure of the NSpace for storing data for salaries of a company (Figure 3) we may use NSpace stereotype (Figure 4). as it is shown, the space of level 3 with data for year salaries contains spaces of level 2 with data for month salaries. The space of level 1 may contain the individual salaries of the employees.

Conclusion

Two stereotypes for Numbered Space (NSpace) and Space with Natural Language Addressing (NLASpace) from an UML profile for Multi-Dimensional Numbered Data Bases (MDNDB™) were shortly outlined in this paper. MDNDB™ is a tool for storing Big Data locally as well as in the cloud. Because of this, it is important to have appropriate modeling language to support the design of practical applications.

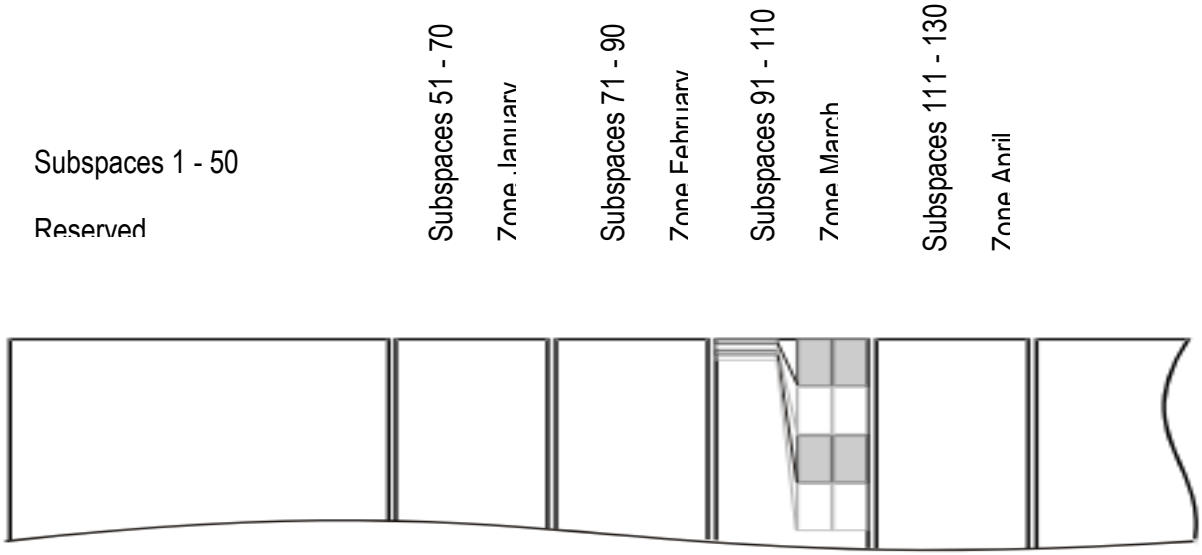


Figure 3

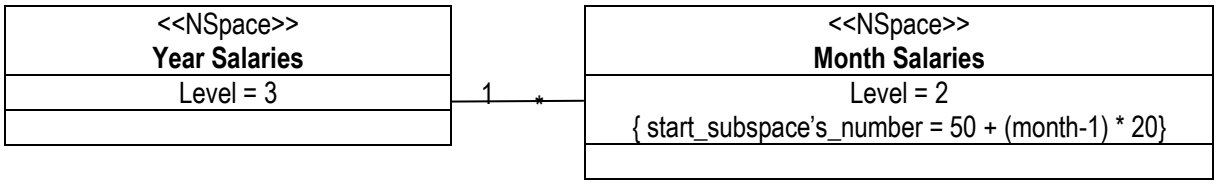


Figure 4

Bibliography

[Markov, 2004] Kr. Markov. Multi-Domain Information Model. Int. Journal "Information Theories and Applications", 2004, Vol. 11, No. 4, pp. 303-308.

[Markov et al, 2015] Krassimir Markov, Krassimira Ivanova, Koen Vanhoof, Vitalii Velychko, Juan Castellanos, „Natural Language Addressing", ITHEA® Hasselt, Kyiv, Madrid, Sofia, IBS ISC No.: 33, 2015, ISBN: 978-954-16-0070-2 (printed), ISBN: 978-954-16-0071-9 (online), 315 p.

Authors' Information

Olena Chebanyuk *– PhD, Professor associate of Software Engineering Department, National Aviation University, Kyiv, Ukraine. e-mail: chebanyuk.elena@ithea.org ; chebanyuk.elena@gmail.com*

Krassimira Ivanova *– PhD, Professor associate, Head of Information Technologies Department, University of Telecommunications and Post, Sofia, Bulgaria. e-mail: ivanova.kr@ithea.org ; krasy78@mail.bg*

Vitalii Velychko *– PhD, Professor associate at V.M.Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine. e-mail: aduisukr@gmail.com ; Velychko.Vitalii@gmail.com*

Krassimir Markov *– PhD, Honorary Professor, Information Technologies Department, University of Telecommunications and Post, Sofia, Bulgaria. e-mail: markov@ithea.org*

DOMAIN ANALYSIS OF PROCESSES GPS TECHNOLOGY POSITIONING SYSTEMS

Mykola Guzii

***Abstract**: Paper touches the next questions: general principle of GPS operation, algorithm of positioning using GPS technology, and Tasks of GPS components*

***Keywords**: Global Position System, GPS-receiver*

1. General principle of GPS operation

Global positioning system (GPS) is a satellite navigation system that precisely determines the position and velocity of objects in any point of the Earth surface at any time and any weather (Олиферб 2001).

The near-Earth space is evenly covered with a network of satellites. The orbits of satellites are calculated with very high accuracy, so the coordinates of each satellite are known at any time. Sattelite transmitters continuously emit signals towards the Earth. These signals are received by the GPS-receiver, which is located somewhere on the Earth and this place is to be determined.

The receiver measures the time needed to transmit the signal from the satellite and calculates the distance "satellite-receiver" (radio signal is known to travel with the speed of light). As for determining the location of the point you need to know three coordinates (flat coordinates X, Y and height H), the receiver must measure distances to three different satellites. Obviously, this method of navigation (without request) implies that the precise timing of the signal propagation is possible only when the timelines of the satellite and the receiver are synchronized. That is why the satellite and the receiver include reference clocks (frequency standards), and the accuracy of the satellite time standard is extremely high (long relative frequency stability is ensured at the level of 10-13- 10-15 per day). Onboard clocks of all satellites are syncronized and comply the so-called "system time". The time standard of the GPS-receiver is less accurate not to raise its cost. This etalone should provide only short-term stability of frequency – while taking measurements (Ватаманюк, 2008).

Taking into account the error of measuring time, allows to calculate distance from the GPS-receiver not only to three but to four satellites. Processing of these measurements results in calculating the coordinates (X, Y and H) and the precise time by the receiver. If the receiver is mounted on a moving object the speed of the object can be calculated too. Thus, to make the required navigational calclations

constant visibility of at least four satellites is to be provided. Modern GPS-receivers have from 5 to 12 channels that can simultaneously receive signals from the same number of satellites. Redundant measurements (more than four) can improve the accuracy of the coordinates and ensure continuity in solving the navigation problem.

2. The Algorithm of Positioning Using GPS Technology

If there is a request for positioning (simultaneously for all satellites fig. 1):

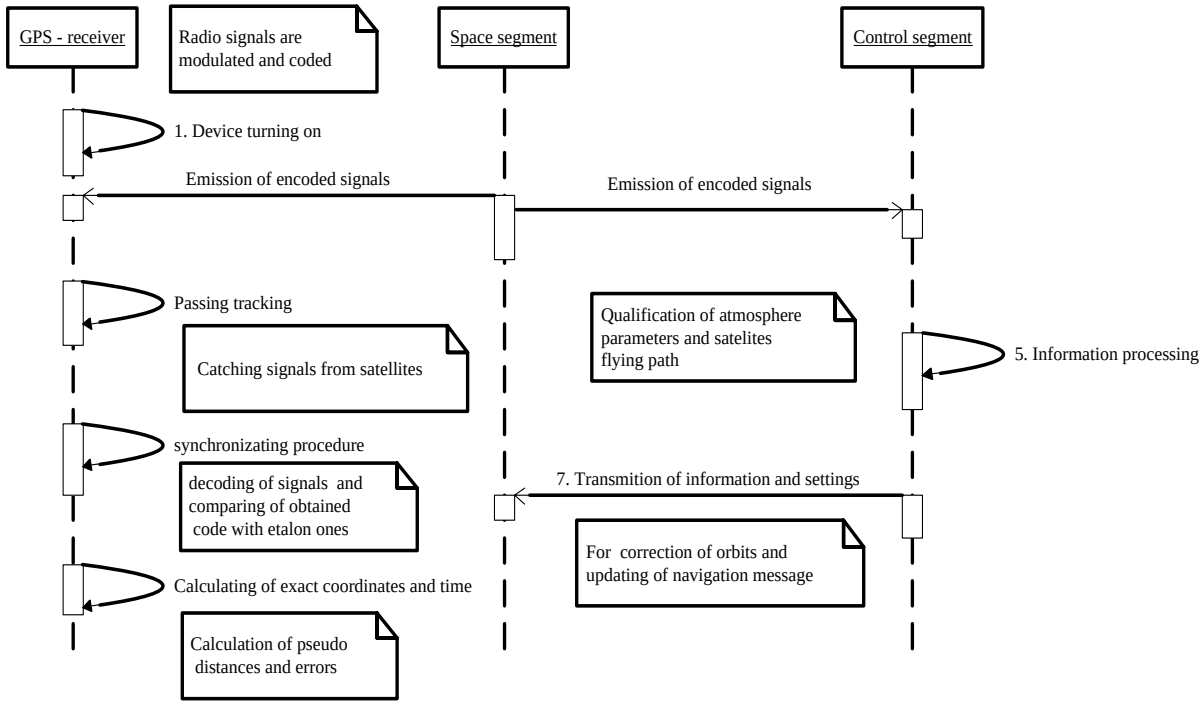


Fig 1 *Sequence diagram*: GPS-receiver operation

1. GPS-receiver accepts encrypted signal from satellite transmitters
2. Consumer equipment is transformed into the tracking mode (procedure of syncronizing is performed)

3. The receiver decodes the received signals

4. Compares the obtained code with the reference code (the reference code is generated by the GPS-receiver)

5. Analyzes obtained values - defines the delay of the signal transmitted from the satellites to the user equipment

6. Calculates pseudorange and precise time

7. If the object is moving:

7a. simultaneously measures the doppler shift of the radio signal frequency
7b. calculates the velocity of the object

8. Define position of the object (Гайер, 2006)

3. The tasks of GPS components.

The Global positioning system includes a constellation of satellites (space segment), the network of ground tracking and control stations (control segment) and GPS-receivers (consumer equipment). The space segment is composed of 26 satellites (21 primary and 5 spare) that rotate on 6 orbits, whose planes are tilted at an angle of about 55° to the plane of the equator and shifted to each other by 60° in longitude. Orbital radius is about 26 thousand km, and the rotation period is half a sidereal day (approximately 11 hours 58 min.). Each satellite has 4 frequency standards (two cesium and two rubidium - for the purpose of reservation), a solar battery, engines of correcting orbits, receiving and transmitting equipment, a computer on board. Encoding of the radio signal emitted by the sattelite is aimed at:

- ✓ ensuring synchronization of the satellite and the receiver signals;
- ✓ providing the best conditions for distinguishing the signal from the background noise by the receiving equipment (pseudorandom codes are proved to possess such properties);
- ✓ implementation of restricted access to the GPS mode implying that precise measuring is possible only when the use of the system is authorized.

The control segment contains the main control station management (Falcon airbase in Colorado), five tracking stations and three laying stations. In addition, there is a network of public and private stations of tracking satellites that perform observations to refine the parameters of the atmosphere and the flight path of the satellites. The collected information is processed in supercomputers and periodically transmitted to satellites to correct the orbits and update navigation reports.

The received signal is decoded in the consumer equipment (GPS-receiver). The resulting code is compared with the same code generated by the GPS-receiver that allows to determine the delay in transmitting the signal from the satellite and to calculate pseudodistance. After picking up the satellite signal the equipment of the receiver operates in the tracking mode, i.e. received and reference signals are synchronized.

Bibliography

1. *Олифер Н.А.* Компьютерные сети. Принципы, технологии, протоколы / Н.А. Олифер, В.Г. Олифер. – СПб. : Питер, 2001. – 672 с.
2. *Ватаманюк А.И.* Популярный самоучитель Домашние и офисные сети под Vista и XP. / А.И. Ватаманю – М. : Питер, 2005. – СПб. : 2008. – 560 с.

3. *Гайер Дж.* Самоучитель Беспроводные сети. Установление и устранение неполадок за 5 минут. / Дж. Гайер, Р. Кинг. – М. : НТ Пресс, 2006. – 340 с.

Authors' Information

Mykola Guzii *–PhD., Prof. of Computer Systems and Networks Department, National Aviation University, Kyiv, Ukraine. E-mail: mykola.guzii@gmail.com*

Major Fields of Scientific Research*: Computer Networks, Software Engineering*

Section: Applied Aspects of Software Engineering

SPECIFICATION BASED VERIFICATION OF CONTROL SYSTEMS WITH CONCURRENT BEHAVIOR

Liudmila Cheremisinova

Abstract*: The problem of specification model based verification of control systems with concurrent behavior is considered. A specification model of the desired behavior of the system implementation is in the form of parallel automaton that describes concurrent control algorithm. The proposed test case generation approach is based on constructing a reduced parallel automaton reachability graph. To avoid the state explosion problem the technique of partial order reduction is suggested.*

Keywords*: Conformance testing, concurrent system, reachability graph*

Introduction

The complexity of modern chip designs and the quality requirements have led to increased requirements for the ability of the verification techniques. Verification is fulfilled in different design stages. In particular, when an implementation of a system is received, the aim is to check whether it functions as we desired. Here we will focus on the problem of conformance testing [Tretmans, 2008] that aims to check whether the implementation of a system under test conforms to its functional specification. Usually the verification process is achieved by means of carrying out series of experiments with implementation under test where test cases supply inputs to the implementation and the outputs are observed. Generating test vectors is very labor-consuming and one cannot state which cases such a testing has missed. One of the new technologies for solving the problem of conformance testing uses model-based testing (MBT) [Tretmans, 2008] where the starting point for testing is a specification of the desired behavior of the system under test (SUT) in some formal language.

In the paper, we consider the problem of implementation testing for the case of reactive control systems with concurrent behavior. Their peculiarity consists in continuously signal interchanging with an external

environment to accomplish a task. The most important property of such control systems is their inherent concurrence of the processes occurring in them.

Input-Output Conformance Testing

Conformance testing evaluates the functional behavior of a SUT with respect to its specification. A SUT implementation is a real, physical object that is accessible only through its external interfaces and is seen in the paper as black-box. The specification is assumed to be correct and valid and is expressed in some formal language, for example, finite state machine (FSM), which is widely used for describing control protocols. At the same time, there are a number of systems in which the expressive means of the FSMs are not enough. The most important property of such control systems is their inherent concurrence of the processes occurring in them. In the paper the test-case generation problem is considered for concurrent systems that are specified by true-concurrency models – parallel automata [Zakrevskij, 1999], which represent a subclass of interpreted Petri nets.

In the paper, conformance testing reactive systems is done on the base of input-output conformance (ioco) relation [Tretmans, 2008]. A SUT conforms to its specification with respect to the conformance relation ioco if after any trace of the specification states the observable outputs of the SUT are possible outputs in the specification. There is a number of evident assumptions in order for the test to be possible: a) both specification and SUT have the same input and output alphabets; b) specification algorithm is strongly connected; c) a SUT should be input-enabled, i.e. in any its state, it accepts all inputs (that is the assumption of the correct ioco implementation [Tretmans, 2008]); d) at any time an input that is possible in the specification should be possible also in the implementation; e) any output produced by the implementation should be produced by the specification too.

Model specification

To define the behavior of reactive control systems, the formal model of a parallel automaton in PRALU language [Zakrevskij, 1999] is used. The parallel automaton may simultaneously be in several internal partial states, which are called parallel. A set of all such states, the automaton is in some moment, defines a global state that can be interpreted as a marking of the corresponding Petri net.

A parallel automaton is specified by: 1) a set S of partial internal states; 2) input and output alphabets X and Y of Boolean variables; 3) a set of transitions τ_i in the form $\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2)$, where γ_i^1 and γ_i^2 are subsets of partial states from S, k_i^1 and are elementary conjunctions of variables from X and Y. k_i^1 defines a condition of firing the transition τ_i, k_i^2 defines output signals accompanying it. A transition $\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2)$ is enabled when two conditions take place: 1) reached automaton global state N_t

contains all partial states from γ_i^1 ($\gamma_i^1 \subseteq N_t$) and 2) $k_i^1 \wedge W_t = 1$, where W_t is a state on the set of logical signals from X, which has been reached to the time moment t. After a transition is executed, the variables from k_i^2 are assigned values that reverse k_i^2 to 1, and N_t is replaced by $N_{t+1} = (N_t \setminus \gamma_i^1) \cup \gamma_i^2$. In the case of a correct parallel automaton for any pair of transitions τ_i and τ_j, from the condition $(i \neq j) \wedge (\gamma_i^1 \cap \gamma_j^1 \neq \varnothing)$ follows that $\gamma_i^1 = \gamma_j^1$, and if γ_i^1 and $\gamma_j^1$are concurrent, then $\gamma_i^1 \neq \gamma_j^1$.

Thus, the parallel automaton dynamics is described in the space of reachable states and values of automaton logical signals. To calculate the next automaton global state, it is necessary to determine the set T of transitions τ_i, which are fired by the marking N_t and the state W_t. The following states are specified such a manner: 1) $N_{t+1} = (N_t \setminus \gamma^1) \cup (\gamma^2)$, where $\gamma^1 = \cup_i \gamma_i^1$ and $\gamma^2 = \cup_i \gamma_i^2$ are unions of sets of partial states of all transitions $\tau_i \in T$; 2) all variables mentioned in conjunctions k_i^2 for all $\tau_i \in T$ change their values to result in $k_i^2 = 1$. The structure of a parallel automaton can be represented by a state transition diagram, that is similar to that of Petri nets. Here nodes are put into correspondence with partial states, and labeled arcs – with transitions.

Reachability graph reduction

One of the natural approaches to MBT of concurrent systems is one based on exploration on the state space of its model. The enumeration of state space requires the construction of the reachability graph and it is the main approach underlying the methods for analyzing the behavioral properties of concurrent systems. The reachability graph can serve as a specification for the SUT as well. The global states of the parallel automaton defined by the reachable markings N_t correspond to the vertices of the directed graph of reachable states. An arc is labeled with the symbol τ_i of the automaton transition and connects a pair of vertices with such markings, that the automaton change the first marking onto the second one as result of the transition τ_i firing. The reachability graph can be easily generated by computing all successor markings starting from the initial one.

Reachability graph contains all possible partial orderings of parallel transitions for a model with interleaving transitions and as result the reachability graph confronts with the state-explosion problem. Its size grows exponentially with the size of partial states and transitions and its generation suffers from exponential space and time complexity. The large state-space has a very negative impact on the performance of testing. Sometimes, the system is too large to construct reachability graph and receive even any covering path. More promising approach is partial order reduction [Lluch-Lafuente, 2002] to cure the complexity issues of reachability graph generation. Partial order reduction is successively applied in Petri net formal verification methods. However, these methods do not take into account the information interaction of the transitions, which is unacceptable when solving the task of generating a

test cases. Partial order reduction approach is based on the fact that there is independence between some concurrent transitions, that is, the order in which the transitions are executed is not important. In the process of the reachability graph reduction we try to explore only some successors for each vertex.

A transition $\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2)$ is enable in a global state N_t of parallel automaton if $\gamma_i^1 \subseteq N_t$ and current values of input signals draw conjunction k_i^1 into 1. For a pair of transitions τ_i and τ_j that are enable in a global state N_t the transition τ_i do not disable τ_j if after executing τ_i the transition τ_j remains enable. That takes place when the following holds: 1) the condition of transitions firing are compatible with each other, that is $k_i^1 \wedge k_j^1 \neq 0$; 2) output signals mentioned in k_i^2 are compatible with input condition of τ_j, that is $k_j^1 \wedge k_i^2 \neq 0$.

A pair of concurrent transitions τ_i and τ_j are independent if for each global state the properties hold: 1) transitions τ_i and τ_j are commutative in the sense that they may be executed in any order leading to same global state; 2) transitions τ_i and τ_j do not disable each other. For independent transitions, the order, in which they are executed, is not relevant. So we can change the order of their firing or even fire them simultaneously in some cases.

To reduce the reachability graph we propose to fire some independent transitions simultaneously. A pair of transitions τ_i and τ_j can be fired simultaneously in some global state N_t if 1) they are independent; 2) they are enable in N_t and current values of input signals. If for some achieved marking N_t the transitions τ_i and τ_j are enable and can be fired simultaneously then the marking of N_t is replaced by $(N_t \setminus (\gamma_i^1 \cup \gamma_j^1)) \cup \gamma_i^2 \cup \gamma_j^2$.

Conclusion

The paper deals with the task of analyzing a control system for the most difficult and insufficiently studied case when the specification model describes a concurrent system. The considered method is based on the construction of reduced reachability graph of a parallel control algorithm.

Bibliography

[Tretmans, 2008] J. Tretmans. Model based testing with labelled transition systems. In: Formal Methods and Testing: Lecture Notes in Computer Science, Vol. 4949, Springer, 2008, pp. 1–38.

[Zakrevskij, 1999] A. D. Zakrevskij. Parallel Algorithms of Logical Control. Minsk: Institute of Engineering Cybernetics of the National Academy of Sciences of Belarus, 1999. (in Russian).

[Lluch-Lafuente, 2002] A. Lluch-Lafuente. Partial Order Reduction in Directed Model Checking. In: Proceedings of the 9th International SPIN Workshop on Model Checking of Software, Springer-Verlag Berlin, Heidelberg, April 11–13, 2002,pp. 112–127.

Authors' Information

- ***Liudmila Cheremisinova*** *– The United Institute of Informatics Problems of National Academy of Sciences of Belarus, principal, Surganov str., 6, Minsk, 220012, Belarus; e-mail: cld@newman.bas-net.by*

Major Fields of Scientific Research: Discrete mathematics, Logic design automation

UDC 004.78

INTERACTION MODEL OF INTERNET-CONFIGURABLE DEVICES FOR THE WEB OF THINGS ECOSYSTEM

Kyrylo Malakhov*, Oleksandr Shchurov**

Introduction

In the last few years, the Internet of Things (IoT) has become one of the most promising and exciting developments in technology and business. The vision of a world where tiny computers with sensors and communication interfaces that are embedded in the infrastructure of our cities or in cars, offices, or clothes is likely to revolutionize every area of our lives – how we play, how we work and do business, and how we live. Until recently, IoT projects mostly focused on building small-scale, closed, and isolated deployments where devices were not designed to be easily accessible or reprogrammable. The bespoke coupling between devices and applications in a given use case means that any change to an existing deployment is complex and expensive. These limits both the maintenance and evolution of the Internet of Things because considerable resources (time, money, technical skills) are required each time a new function is added. In contrast, the web has become widely successful in the last two decades because it's simple to learn and use and it also emphasizes loose coupling between servers, browsers, and applications. The simple and clearly defined programming model of HTTP makes it possible for anyone to change pieces of the system without breaking the whole system. Therefore, building new web applications has been relatively inexpensive and accessible to a much larger group of technology enthusiasts.

The Web of Things (WoT) is a specialization of the Internet of Things that uses what made the web so successful and applies it to embedded devices in order to make the latest developments in the Internet of Things accessible to as many developers as possible. On the Web of Things – just like on the web – anyone with a text editor and the basic understanding of web standards (HTML and HTTP) can quickly start connecting devices and objects to the web. But it also enables going to the next level and helps you to effectively build interactive and innovative real-world applications that blend the physical and digital worlds.

The given paper describes general Web Thing interaction model (with semantic description to be a part of the Semantic Web ecosystem.

The Internet of Things

Capturing the essence of the Internet of Things in one sentence is nearly impossible. The concepts have been around for decades and there are no clear boundaries to what the IoT is or isn't. Nevertheless, the broad definition of the Internet of Things vision is a world where the internet is much more than the collection of multimedia content it is today: it extends into the physical, real-time world using a myriad of small or even tiny computers. In short, the simplest definition for the Internet of Things is the following [Guinard and Trifa, 2016] – the Internet of Things is a system of physical objects that can be discovered, monitored, controlled, or interacted with by electronic devices that communicate over various networking interfaces and eventually can be connected to the wider internet.

Today thanks to the colossal progress in embedded devices that brought into the world a new class of objects: smart things. A smart thing (which we'll refer to as a Thing in the rest of this paper) is a physical object that's digitally augmented with one or more of the following: *sensors* (temperature, light, motion, and so on), *actuators* (displays, sound, motors, and so on), *computation* (can run programs and logic), *communication* interfaces (wired or wireless). Things extend the world we live in by enabling a whole new range of applications. Concretely, the Things in the Internet of Things can range from simple tagged products such as your Nova Poshta package with

an Auto-ID tag (Automatic Identification methods such as bar codes, QR codes, NFC or RFID tags) attached to it so it can be tracked from the shipping center to your door; to more elaborate, complex, and wirelessly connected products, devices, or machines such as security systems, your car, or a factory assembly line; and all the way up to a building or even a city. The internet part of the term means that the Thing can be accessed and processed by other applications through the existing internet infrastructure. Note that this does not imply that the Thing itself must be directly connected to the internet. The communication network used can be an Auto-ID method, short-range radio (Bluetooth, ZigBee), or the Wi-Fi network in a building.

For the Internet of Things to become real, we need a single universal application layer protocol (think language) for devices and applications to talk to each other, regardless of how they're physically connected. Rather than inventing yet another protocol from scratch, why not reuse something that's already widely used to build scalable and interactive applications, such as the web itself? This is what the Web of Things is all about: using and reusing readily available and widely popular web protocols, standards, and blueprints to make data and services offered by Things more accessible to a larger pool of (web) developers.

The Web Thing model

The Web of Things is composed of web Things. A Web Thing (or simply Thing) is a digital representation of a physical object accessible via a RESTful Web API. Examples of Web Things are: a single board computer, a TV, a building or any entity that can be processed by a control device hosted in the Cloud service, etc. The API of the Web Thing can be hosted on the Thing (entity) itself or on an intermediate host in the network such as a Gateway or a Cloud service (for Things that aren't directly accessible through the Internet).

Implementation strategy of our Web Thing interaction model represented on the figure 1, where: Clients – thin client such as smartphone, Chromebook etc.; E – SPARQL endpoint; Web Thing – NanoPi NEO Air SBC for knowledge processing; Knowledge sensor – applications or devices for data processing and knowledge extraction.

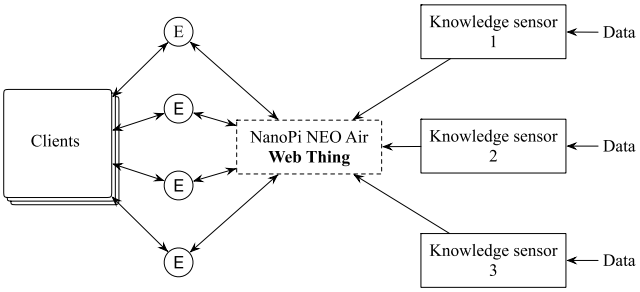


Figure 1. Web Thing interaction model of Internet-configurable device for knowledge processing (data processing and knowledge extraction)

Conclusion

The given paper describes general Web Thing interaction model (with semantic description to be a part of the Semantic Web ecosystem) of the SPARQL server engine (triplestore / RDF store) Apache Jena Fuseki [Fuseki, 2019] based on the Single board computer (SBC) NanoPi NEO Air [NanoPi, 2019] with ARM32v7 architecture for data processing and knowledge extraction.

References

[Fuseki, 2019] Apache Jena Fuseki SPARQL server. Access mode http://jena.apache.org/documentation/fuseki2/index.html

[NanoPi, 2019] NanoPi NEO Air Wiki documentation by Friendly ARM. Access mode http://wiki.friendlyarm.com/wiki/index.php/NanoPi_NEO_Air

[Guinard and Trifa, 2016] Guinard, D., Trifa, V. (2016). Building the web of things: with examples in node.js and raspberry pi. Manning Publications Co.

APPROACH TO DESIGNING A SYSTEM FOR SMART CINEMAS

Mariia Kryvoruchko, Mariya Shirokopetleva

Abstract*: The approach to designing a system for smart cinemas is given. Deployment diagram, sequence diagram and technologies are proposed.*

Keywords*: Software engineering, design, system, diagram, cinema.*

ITHEA Keywords*:Design Tools and Techniques.*

Introduction

In the modern world most developments are intended to improve people's lives, automate various day-to-day tasks, save money and time. Every day people spend a lot of time in queues, especially in cinemas and other public places. A system that wouldlet users to easily buy tickets in advance or on the spot, order meals in the food court from the app, move quickly among the cinema zones and most importantly help them not waste time in queues would be helpful.Also it would be a financial advantage for the cinema owners to have the opportunity to attract visitors effectively and monitor the activities to improve their service. Currently there are no popular systems that solve all these problems.

It is possible to solve the given problems by developing a hardware and software system that would automate the access to cinema zones therefore minimizing queues and providing owners with useful statistics.

Design

Figure 1 shows the system deployment diagram, which includes all system components:Web app,Android app, IoT device and the main server that is also connected to the database server. Client applications have REST API access to the server-side data, which is only available to authorized users.

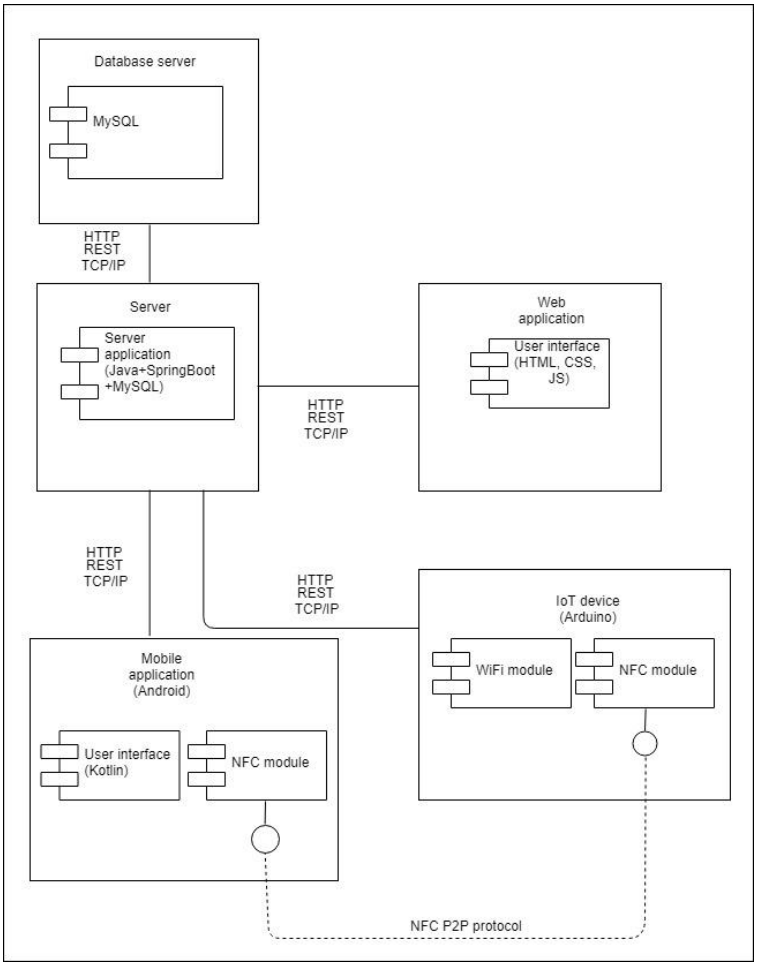


Figure 1. Deployment diagram

For the system to operate correctly a device for authorizing the tickets is required. The device would use NFC technology to read and process users entrance ticket id by sending it to the server. For solving such a task it is proposed to use the following components: a NodeMCU microcontroller with a Wi - Fi module ESP8266; NFC RFID PN532 for reading the NFC data and RGB WS2812 indicator for displaying the result.

Let's look at the usage of such an access control device at one of the cinema zones – the cinema hall for instance.

The sequence of actions performed on the device, as well as the interaction of the device with the mobile application is described as sequence diagram, which is shown in Figure 2.Initially the IoT device connects to Wi-Fi, logs on to the server and receives a token that will be used for further server requests as well as the cinema zone number where the device is located. When the visitor hovers the phone over the device, the ticket number is transmittedto the device, and sent to the server for verification.If the ticket is valid, the device receives the permission to let the visitor into the room.

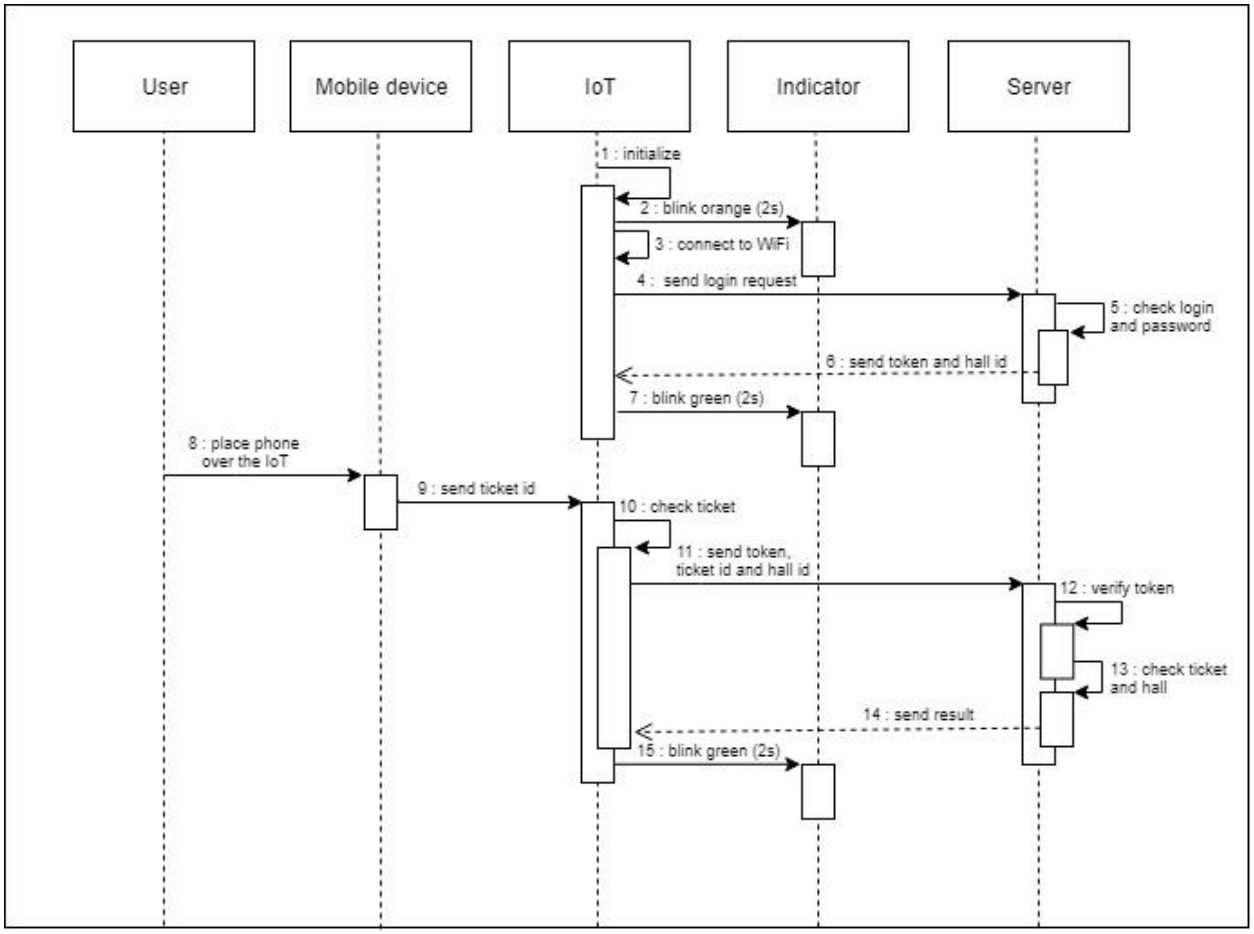


Figure 2. Sequence diagram

Such reliabletechnologies can be used for the development of the system: Kotlin Android for the mobile application;JavaSpringBootframework for the backend server;MySQL for the database server;JavaScript,HTML, CSS for the web app;Arduino platform and C++ for the IoT device.

Conclusion

The subject area of the smart cinema system was analyzed, the deployment and sequence diagrams were created and demonstrated, the system was designed and the main technologies for development were chosen.

Authors' Information

Mariia Kryvoruchko – *Kharkiv National University of Radio Electronics; student.E-mail: mariia.kryvoruchko@nure.ua.*

Mariya Shirokopetleva – *Kharkiv National University of Radio Electronics, Senior Lecturer. E-mail: marija.shirokopetleva@nure.ua*

A METHOD OF AUTOTUNING .NET PROGRAMS ON TARGET PLATFORMS WITH REWRITING RULES

Tural Mamedov, Anatoliy Doroshenko

***Abstract**: Software method for computational optimization that allows to semiautomatically increasing execution speed was developed. The plugin to analyze C#-programs for Termware system, which implements a concept of rewriting rules, was implemented. The plugin consists of analyzer for C# using Roslyn platform and interpreter for generating terms of Termware system from source code and vice versa. A TuningGenie framework was used for optimization purposes. The work of software tool was illustrated on famous academic algorithm "Conway`s Game of Life". The results were obtained by experiments on various data sizes. Speed of the program before and after modifications and the number of required modifications in the source code to use TuningGenie and another framework Eazfuscator.NET were measured on personal computer.*

***Keywords**: autotuning, speed measurement, C# analyzer, cellular automaton "Conway`s Game of Life".*

***ITHEA Keywords**: software, information systems*

Introduction

The solving of many technical tasks requires the usage of complex algorithms and/or processing large amount of data. Software engineers need a lot of time to implement application, then to optimize an existing solution.

Another feature of software development is to create a program for the characteristics of the target-computing environment, and optimization stage for the target platform requires a lot of effort.

Obviously, writing an optimal program for different operating systems is a very complex, or even impossible, task because of the too large variety of modern processors.

Therefore, to achieve maximum performance, at least in terms of the performance of the program, its additional configuration is required for the various operating systems, in which it will be executed. Consequently, a modern tuning software methodology allows automating this debugging. This methodology is called auto-tuning. An auto-tuning is a concept, which is to empirically evaluate several program options, select, and use the best of them.

The traditional approach to this methodology is to develop a separate tuner program that performs the required user selection and generates various modifications of the program. This approach is very convenient in terms of the distribution of tasks, that is, the source program calculates only the solution of the task.

Task of the work

The purpose of this work is to develop methods for automating the configuration of applications for the target platform using the rewriting rules for the .NET platform. As to the requirements for this work:

- Correctness of changes in the program being configured;
- Development of the analyzer for the C# language and the translation into the language of the terms;
- Development and usage of methods to improve the performance of an optimized program, implemented in C#;
- Application of the developed platform to academic example.

Implementation an analyzer for C#

In this work the TermWare [Shevchenko and Doroshenko, 2019] framework was used, which developed the concept of terminal systems of object structures and rewriting rules. TermWare is written in Java, but uses the Parser / Printer plugin separately. Therefore, despite the fact that the framework is written on the same technology, it provides the possibility of multiplatform. To do this, an instance of TermWareInstance was created - which is, figuratively speaking, a core that contains a hierarchical system of domains and dictionaries of parser names of languages and printers. To work with any other technology stack you must implement 4 interfaces: IParser, IParserFactory, IPrinter, IPrinterFactory

For parsing program analysis, open source software - Roslyn compiler was used. With the help of the compiler you can get the syntax tree of the implemented program, which is required for the parser.

Because Roslyn works on CLR, and the parser/printer should work on JVM, Javonet bridge was used to solve this problem, which allows you to register any dll file by having it in the same directory as the project. This allows you to use and create instances of classes from C# to Java, and vice versa from Java to C#.

Then TuningGenie [Ivanenko et al, 2014] framework was used to optimize the program in terms of performance.

Results

To determine the effectiveness of program transformations an experiment was conducted using two types of automatic program optimization. As an example, the well-known academic algorithm "Conway`s Game of Life" was taken. "Conway`s Game of Life" is an infinite, two-dimensional, orthogonal grid of square cells, each of which is in one of two possible states, living or dead (or settled and uninhabited, respectively). The experiment consisted of a comparison of the results of the execution of programs obtained using the described systems. The experiment was conducted using a computer equipped with an Intel Core i7-7700HQ 2.80GHz processor with 16 GB of RAM. Scaling the research problem is controlled by selecting the size of the plane on which the "Game of Life" is performed, as well as the number of generations that must pass before the program is completed. The experiment was conducted on different plane sizes, ranging from 300 and ending with dimensions of 1 200 000 cells. The result of the experiments is displayed on the Excel file and shown in Fig. 1. where you can see that TuningGenie is more efficient in terms of performance than the Eazfuscator.NET library.

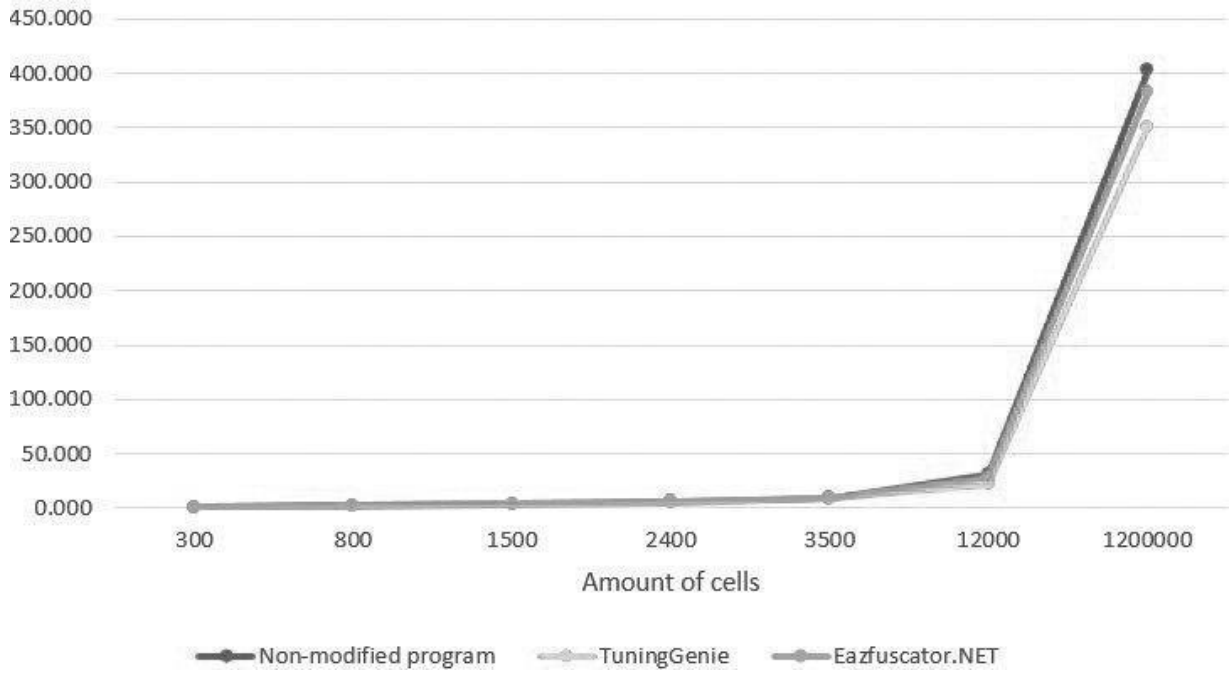


Fig. 1. Comparison of TuningGenie and Eazfuscator.NET

Conclusion

In this paper, the methods of auto-tuning C#-programs to the target platform with the help of rewriting rules were considered. An experiment was conducted that showed that the applied method is the best

indicator of software performance. For the experiment, the algorithm "Conway`s Game of Life" and the method of choosing the optimal strategy for bypassing the indexed data structures, which was implemented in the framework TuningGenie, was used. The result of the work was obtained and shown on the graph and compared with the work of the other framework Eazfuscator.NET.

References

[Shevchenko and Doroshenko, 2019] Shevchenko R.S., Doroshenko A.Y., Termware-3 – term rewriting system, based on context-term calculus, Problems in programming, Vol 1., 2019, pp. 48-56., ISSN: 1727-4907 (print version), doi: https://doi.org/10.15407/pp2019.01.048

[Ivanenko et al, 2014] Ivanenko P.A., Doroshenko A.Y., Zhereb K.A. (2014) TuningGenie: Auto-Tuning Framework Based on Rewriting Rules. Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2014. Communications in Computer and Information Science, vol 469. Springer, Cham

Authors' Information

Tural Mamedov – Taras Shevchenko National University of Kiev; e-mail: tural.mamedov1@gmail.com;

Anatoliy Doroshenko – Institute of Software Systems, National Academy of Sciences of Ukraine; e-mail: doroshenkoanatoliy2@gmail.com

Major Fields of Scientific Research: Parallel Computing, Computing in Mathematics, Natural Science and Engineering

PROVIDING STABILITY OF THE PROCESSES TO RISKS OF THE SOFTWARE SYSTEMS' LIFECYCLE BASED ON THE CONCEPTS OF THE NEW RISK MANAGEMENT PARADIGM

Vladimir Talalaev, Yuliia Bezkorovaina

***Abstract**: The theory and practice of creating and applying modern application systems have shown that their effectiveness significantly depends on the ability of such systems to withstand the influence of factors of uncertainty on the fundamental processes of their lifecycle.*

***Keywords**: Software Engineering, Management, Lifecycle, Risk, Paradigm.*

Introduction

Nowadays, there are various software system development technologies. Some are intended for agile development and implementation, and others are intended for sustainable development. But all of them pass the main stages of development. At these stages, under the influence of various factors, some risks can lead to undesirable results, such as delaying the release of the final product. So eventually emerged the concept of the new paradigm of risk management. This concept should help reduce the impact of factors on the timing and quality of the software system and ensure the stability of the processes to risks of the software systems' lifecycle.

Concepts of the new risk management paradigm

Systems that provide software production, implementation, and maintenance related to a specific class of application systems. They have high levels of uncertainty that characterize the internal and external environment. Therefore, their activity will be accompanied by the existence of potential manifestations of risks at all stages of the software lifecycle. It is no coincidence that the overwhelming majority of modern technologies and methodologies for software development, implementation, and maintenance as a critical component contain a variety of risk management mechanisms. These mechanisms typically involve the implementation of processes for identifying, identifying, assessing, and regulating the effects of risks on the processes of the software lifecycle. Most of these mechanisms were built based on best managerial risk management practices that were reflected in international standards.

An analysis of the content of the fundamental risk management mechanisms for the vast majority of commonly used risk management practices in the software lifecycle processes shows that their structure and performance are based on the application of classical principles and concepts regarding risk management.

The essence of such mechanisms is to create and implement management schemes that ensure implementation of the processes of the risk management cycle: "detection of risk factors - identification of risk - risk assessment - the selection of protection against risks measures - implementation of protection against risks measures." The resulting schemes were further applied to the factors of influence, the uncertainty of which can lead to negative consequences for the processes of the software life cycle. Nowadays, in the practice of risk management, a large number of models have been developed that allow the implementation of these risk management schemes. The vast majority of such models were targeted at specific types of risks and, therefore, are useful only concerning these types of risks. At a certain period of development of risk management, the application of such approaches allowed to solve the problem of risk management both in the life processes of application systems in general, as well as in the processes of the software life cycle, in particular.

However, the practice of risk management in recent decades has shown that due to changes in the processes of the life cycle of modern software systems, the application of traditional risk management schemes is complicated due to the following reasons:

1) Significant expansion of the processes of the life cycle of software systems;
2) Significant increase in external and internal interactions in systems is due to a significant increase in the openness of modern application systems;
3) Strengthening the levels of variability of the factors of the influence of the external and internal environment of the processes of the life cycle;
4) Absence of reliable and complete data on the factors of influence and the nature of their manifestation, which increases the uncertainty (entropy) of the external and internal environments;
5) The complexity of using analytical models to describe the uncertain effects of factors influencing the processes of the software lifecycle;
6) Significant levels of synergistic interaction uncertainties in their manifestation in the processes of the software lifecycle.

In recent years, many works have appeared in the theory and practice of risk management, the authors of which express an opinion on the need to change the traditional paradigm of risk management. A generalized conclusion about the necessity of such a change for the processes of the

life cycle of the software can be formulated as a transition from the "risk management" system to the system "risk management of processes," based on the deep integration of risk-management and management of the life cycle of software processes.

Conclusion

The paper analyzes the state of development of modern risk management systems in the processes of creating and applying software. Following the provisions of the new paradigm of risk management, a systematic presentation of the methodological approach to solving the essential tasks of creating and applying risk management system lifecycle software processes was given.

Bibliography

Karpuntsov, M.V. Upravlinnia ryzykostiikistiu pidpryiemstva. Ekonomika ta derzhava. 2008. №9, p. 72-74..

Kravchenko, O. S. Ryzykostiikist yak peredumova efektyvnoho rozvytku pidpryiemstva: indykatory vyznachennia ta metodyka diahnostyky. Visnyk Donetskoho natsionalnoho universytetu ekonomiky i torhivli im. Mykhaila Tuhan-Baranovskoho. Ser. : Ekonomichni nauky. 2013. № 4., p. 81-90.

Authors' Information

	**Vladimir Talalaev** – Assoc. Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine. **E-mail:** VATalalaev@gmail.com M**ajor Fields of Scientific Research**: computer networks, data networks, information systems, security, management
	**Yuliia Bezkorovaina** — Senior Lecturer of Software Engineering Department, National Aviation University, Kyiv, Ukraine **E-mail:** Yuliia.bezkorovaina@nau.edu.ua M**ajor Fields of Scientific Research**: software engineering, software development, software construction

Section: Software Engineering Application Domains

HYBRID METHOD OF RANKING QUERY RESULTS IN SEARCH ENGINES

L. Chala, D. Hrazhevskyi

Kharkiv National University of Radioelectronics, Ukraine

Introduction

With the advent of the Internet, the available information amount has greatly increased. At the same time, most information data are dispersed in a variety of nodes, both global and local, and is not divided into the necessary categories, meaning, purpose, which greatly complicates and slows down the search for the necessary information at a particular time.

Search engines have been established to resolve these issues. Such services, when creating a request for unstructured information to them, choose from all available documents those whose content most satisfies the content of the request itself, and return them as results – in the form of a list, links, individual files.

At the same time, a meaningful ordering of the results positions is required. Its absence similarly leads to an increase in the amount of time required for the user to find the required documents. It is also necessary to consider the size of the documents collections involved in the information search. In the case of large collections, it is not practical to view all the results and choose the most appropriate ones. Therefore, we can conclude about the openness of the question about the level of search efficiency, using search engines.

Ranking

The effectiveness of search [Langville, 2013], especially research, largely depends on special algorithms for sorting query results, based on the degree of compliance of each document directly to the query content – so-called relevance. This process is called ranking. The probability of the absence in the search results of a more appropriate user request document, or its high remoteness in an ordered list of results is inversely proportional to the level of search efficiency. Considering the constant increase in the amount of information, it can be considered relevant research and development in this area.

Obviously, the use of only one indicator will not be enough to ensure high relevance. There are many factors that affect the relevance, examples of which can be called the content of the document, its technical component [Turnbull, 2016], the frequency of its opening, location. For example, the well-known measure TF-IDF, used to determine the importance of a word in the context of a document, uses both the frequency of the term and the frequency of the document – the number of documents containing the term [Robertson, 2004]. Also, using this measure, it is possible to represent documents in the form of vectors (vector search model), compare and rank them using commonly known metrics - Euclidean distance, cosine measure [Shahzad and Ramsha, 2018].

Different from the vector model of search and ranking based on metrics is ranked search, focused on the Internet and based on links. Considering the reference structure of the Internet when ranking can give a significant improvement in search results [Langville and Meyer, 2004]. The reference structure consists of documents submitted by sites, their pages, and hyperlinks between them. It can be represented as a directed graph with web pages as nodes and hyperlinks as edges. The link analysis methods in the web graph is based on two assumptions that the text of the link that pointing to page B is a good description of page B. and a hyperlink from page A to page B is the recognition of the authority of page B by the creator of page A.

The main methods of reference ranking [Patel and Patel, 2015], based on which create new methods, or use modified, as, for example, global search engines Google, Baidu, are RankPage and Hypertext Induced Topic Search (HITS).

Hybrid method of ranking

One of the most effective ways to increase the productivity of ranking is the use of a hybrid ranking method [Чала, 2019]. When using this method, it is assumed that the collections of documents pass through several successive stages of ranking. As a result, the resulting lists become much more accurate, because with each such pass, the conformity assessment of the content, features of the documents and the request is clarified. Also, it is possible to cover more widely the concepts and purpose laid down directly in the request, to consider new factors.

The hybrid method proposed for practical research also involves the passage of several stages of ranking, but using not one autonomous, but several queries. The idea of this method is to rank the results of a sequence of queries that differ from each other by replacing, adding, or rearranging words. It is assumed that each change will be synonymous to its original expression to obtain more correct results.

A web application was created that allows the user to use a hybrid ranking method when searching. This application was written using .NET and ASP.NET MVC. The algorithm of the program, as well as the method, begins with sending a http request with the content in the form of a proposal to the server of

a remote search engine in order to receive an http-response. The user has the option to enter his query in the search bar of the application. In this application the library Google Custom Search API was used, which allows to use Google search servers, and with which you can send up to a hundred requests a day.

Most search engines use the PageRank reference ranking method to search for the required documents, often modified to consider some other factors as well as vector models. The resulting http response in the form of a ranked list of results is parsed and recorded in the NoSQL database in the form of links that were received first. After that, the links from the database are displayed to the user, and user enters a query similar to the previous one in meaning into the search bar of the application, but the content of which is slightly changed, in order to obtain a result list from both new and previously obtained elements.

Each new search output – http response in the form of a list – of similar consecutive queries will be compared with the previous one in the database. If the two lists have the same item, it will be recorded in the subsequent search results, and marked – it will be assigned a number or other labels at the request of the user, which is also reflected in the database. Each time a subsequent query is detected, the item will receive an additional label, or its number will be increased by one. Thus, the user will be able to view the history of responses to their queries on one screen. Such links will be given to the user first, and ordered by the number of their labels, number, which makes it possible to immediately see the most suitable result.

Conclusion

It was found that the accumulated results of successive synonymous queries can be considered indicators of increasing the productivity of the ranking, and enough factors of relevance. Sorting them by the number of labels allows to see which documents or links most correspond to a certain meaning, the essence of the content of the user's request, and at the same time reduce the likelihood of inappropriate documents or links to high places in the resulting list only because of some coincidence of the text content. We can also talk about the feasibility of this method in the researches, where the relevance of the results of the query is of great importance, which, in this case, depends on the number of labels.

Bibliography

[Langville, 2013] A.N. Langville. Who's #1?: The Science of Rating and Ranking. Ed. A.N. Langville and C.D. Meyer. Princeton University Press, 2013.

[Turnbull, 2016] D. Turnbull. Relevant Search: With applications for Solr and Elasticsearch. Ed. D. Turnbull and J. Berryman. Manning Publications, 2016.

[Robertson, 2004] Robertson S., Understanding Inverse Document Frequency: On theoretical arguments for IDF. Journal of Documentation, Vol.60 Issue 5. Emerald Group Publishing Limited, 2004. pp. 503-520. doi: 00220410410560582 https://doi.org/10.1108/00220410410560582

[Shahzad and Ramsha, 2018] Shahzad Q., Ramsha A. Text Mining: Use of TF-IDF to Examine the Relevance of Words to Document. International Journal of Computer Applications, Vol.181 Issue 1. Foundation of Computer Science, 2018. pp. 25-29. doi: 10.5120/ijca2018917395 https://www.ijcaonline.org/archives/volume181/number1/29681-2018917395

[Langville and Meyer, 2004] Langville A.N., Meyer C.D., Deeper Inside PageRank. Internet Mathematics, Vol.1 Issue 3. University of Chicago, 2004. pp. 335-380. doi: 10.1080/15427951.2004.10129091 https://projecteuclid.org/euclid.im/1109190965

[Patel and Patel, 2015] Patel P., Patel K. A Review of PageRank and HITS Algorithms. International Journal of Advance Research in Engineering, Science & Technology(IJAREST), Vol.2, Issue 1. Veerayatan Group of Institutions, 2015. pp. 1-4. doi: 10.13140/2.1.2719.8082 https://www.researchgate.net/publication/264235759_Research_of_Page_ranking_algorithm_on_Search_engine_using_Damping_factor

[Чала, 2019] Чала Л.Е., Гражевський Д.С. Гібридний метод ранжирування результатів запитів у пошукових системах, Міжнародний науковий симпозіум «ІНТЕЛЕКТУАЛЬНІ РІШЕННЯ». Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф., 15-20 квітня 2019 р., Ужгород / М-во освіти і науки України, ДВНЗ «Ужгородський національний університет», та [ін.]; наук. ред. В.Є. Снитюк. – С.135-137.

ANALYZE AND DEVELOP THE SOFTWARE OF AUTOMATIC SEARCH FOR AN ANONYMOUS PERSON IN THE VOICE DATABASE

Yana Bielozorova

***Abstract**: Objective of the proposed development is to investigate efficiency and develop a program system for formation of the voice databases of information messages involvants. Automatic search for involvants in the database by means of similarity of the voice characteristics.*

***Keywords**: Voice database, probability density, main tone, phonogram.*

Objectives

Sound recording materials are considerable part of evidences used during criminal investigations and proceedings related to corruption, bribery, extortion, racketeering, kidnapping etc. But there are special digital phonograms which contain messages being recorded, for example, in different monitoring and field services. They form a separate type of expert sound recording materials. It can be explained by the fact that the phonogram digital copy is usually used during the expert examination of the materials and examination of originality and authenticity of phonograms is not performed. Another feature of such phonograms is short-duration of the recorded messages.

Such messages are combined and stored in the voice databases. Very often the messages are anonymous. In many cases such messages are made by the same persons but there can be considerably long period between messages made by the same anonymous author. Very often these messages cause significant property and moral damage. On the other hand, sometimes anonymous messages can be related to terrorist activity or have serious social consequences. In view of these reasons sometimes it is required to identify anonymous person (involvant). And during operative investigations, if they are required, at initial stage it is very important to determine whether the messages of the involvant are of repeated nature. And in this case assistance for investigation and retrieval can be performed in expert way. The first task of such expert examination is to identify voices of involvants with similar parameters of signals of their oral speech, contained in the database. This can reveal both recidivism and potentially suspected persons and thus facilitate investigation process. The main tasks of such expert examination include:

- personal identification of the involvant, suspected in delivery of anonymous message or his available voice tokens, obtained during operation by means of parameters of his speech signals;

- identification of the involvants' voices with similar parameters of signals of their oral speech, contained in the database in order to identify persons, potentially suspected in delivery of information (or false information).

Appropriate expert departments can ensure completion of these tasks. However, the database of these messages even within the same country can contain thousands of records, and experts are not able to perform this task without automation of the process. In order to facilitate the involvants identification task it is required to develop automated system of quick search for suspected persons in the available voice database for further personal identification through the physical parameters of his (her) oral speech signals.

Experimental program modulus based on developed approach

Experimental model of the program product, created on the basis of investigations, for realization of these tasks was approved in the expert departments of the Ministry of Justice and the Ministry of Internal Affairs of Ukraine.

This approach was realized and approved using the developed experimental program modulus for search for involvants of information messages in the voice database [1].

Experimental program system (EPS) for search of involvants of information messages has a form of a search engine. Thus, the search results are not identical to identification of person by voice, because they are just results of ranking by the similarity degree of separate parameters of voice signals.

Using digital records of information messages the EPS performs an automatic calculation of parameters of voice characteristics and further ranking of these characteristics in the voice database.

The EPS uses a ranking method by four different criteria. They include:

- calculation of similarity of two-dimensional probability density functions curves for the main tone frequency (MTF) and arrangement in the spectrum of seven formants, extracted from the speech recorded in the phonogram;

- calculation of similarity of probability density functions curves for each of these signs separately;

- calculation of similarity degree for the absolute maxima of formats spectra, extracted from the speech recorded in the phonogram.

Peculiarity of the used method for spectrum calculation is that the calculation of spectrum characteristics is performed with resolution ability of 1 Hz (non-orthogonal transformations in a small time slot). Maxima of the first seven formants are distinguished in the spectrum of each time slot.

For calculation of similarity of two-dimensional probability density functions curves, the function Ff – voice characteristics function should be calculated. This function is a non-linear empirical function of formant maxima amplitude and frequency.

For all utterances the two-dimensional probability density is determined by means of two coordinates – MTF Fb and voice characteristics function Ff. Thus, a combined method for assessment based on

MTF and formant spectrum characteristics is the fundamental element of the system. Similarity degree of these characteristics of two voices is determined by absolute differences between the two-dimensional probability densities.

It is clear, that projections of two-dimensional probability densities per each of one-dimensional axes of coordinates provide distribution of MTF Fb and distribution of voice characteristics function Ff, which also can be presented in a form of separate dependences. It allows making calculation of similarity of probability density functions curves for each of these signs separately.

For ranking by the densities of distribution of the main tone Ft and voice characteristics functions Ff, the value equivalent to Kolmogorov fitting criterion is used as a similarity degree [2]:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \le x}, \quad (1)$$

where $I_{X_i \le x}$ indicates whether the value Xi $(-\infty, x]$ on condition

$$I_{X_i \leqslant x} = \begin{cases} 1, & X_i \leqslant x; \\ 0, & X_i > x. \end{cases} \quad (2)$$

For calculation of similarity degree of absolute maxima of formants spectra, extracted from the speech, recorded in the phonogram, the following criterion is used

$$S_i = abs\{\max[P_1(x_1, x_2) - \max P_2(x_1, x_2)]\}, \quad (3)$$

where $P_1(x_1, x_2), P_2(x_1, x_2)$ – are two-dimensional probability densities for two different records with arguments both as the main tone frequencies and voice characteristics functions, compared separately.

Results of search for voices with similar characteristics are presented as rank tables.

It should be mentioned, that all calculations of voice characteristics are made for the record fragments containing speech signals. Division into pauses and speech fragments in the system is made automatically based on two criteria – level of normalized audio signal and availability of record fragment which contains the speech sound [3-5]. System for sound extraction is additional and integrated into the EPS. It performs general division of the fragments in which the speech sounds are present, based on the special method, similar to the methodology of hidden Markov chains.

Conclusion

An experimental program system ensuring automatic search in the database of persons with voices which are the most similar to the involvant's voice, recorded in the phonogram pattern was created.

However, further improvement of the system in order to create the prototype requires carrying out of many investigations and approval of the program system. In particular, it requires:

- accumulation of large databases with different language groups presented;
- investigation of efficiency of voice characteristics ranking in the databases;
- correction and adaptation of the program product to conditions of different real tasks of recording of information messages involvants' voices;
- creation of multi-language versions of program product localization.

Bibliography

1. O. Rubalsky, V. Solovyov, A. Shablja, V. Zhuravel. New tools to identify the person for voice of Data (Russian) Protection and Security of Information Systems: Proceedings of the 3rd International Scientific Conference (sity. Lviv, 05 - June 6, 2014). - Lviv: - p.p.. 110 - 112.
2. Mallat S., A wavelet tour of signal processing, Courant Institute, New York University, 1999, 671 pp
3. V. Solovyov, Y. Byelozorova: Multifractal approach in pattern recognition of an announcer's voice. Polish Academy of Sciences University of Engineering and Economics in Rzeszów, Teka, Vol. 14, no 2, p.p. 164-170, 2014.
4. V. Solovyov. Using multifractals to study sound files (Russian). Visnik of the Volodymyr Dahl East Ukrainian national university. Vol 9 (151). – p. p. 281-287, 2013.
5. V. Solovyov. Using the fractal dimension of audio files in the problem of segmenting the audio file (Russian). Visnik of the Volodymyr Dahl East Ukrainian national university. Vol 5 (194). – p. p. 165-169, 2013

Authors' Information

Yana Bielozorova – Assistant of Prof. of Software Engineering Department, National Aviation University, Kyiv, Ukraine.

E-mail: bryukhanova.ya@gmail.com

M**ajor Fields of Scientific Research**: Speech Recognition Models, Wavelet analysis, Software Architecture

RESEARCH OF INFORMATION SECURITY THREATS IN MULTIMODAL TRANSPORT INFORMATION SYSTEMS

Tetiana Konrad

Abstract. *The information form of the interaction of types of transport and the main types of threats to information security in multimodal transport information systems is considered.*

Keywords: *Security Threat, Information Systems, Multimodal Transport.*

Intro. Problem statement

Unsanctioned impact in the information process of interacting modes of transportis a direct threat to functioning for multimodal transport information systems (MTIS).

It's can be cause to fail in work devices, machinery and equipment that even become the cause of emergencies. That is why the topical issue is the study of information security threats in multimodal transport information systems.

Presentation of the main research material

Information form of interconnection in transport in multimodal transport information systems will determine the amount of information per form, frequency, hour, classifier and automatization of control systems in transport facilities. Coordination of interacting modes of transport is practically impossible without informational interaction, since the speed and timeliness of decision-making is extremely important for the efficient and safe functioning of all levels of technological interaction and the timely elimination of all shortcomings.

The threat of information security is called the possibility of implementing an action on information processed in the system, leading to a reduction, destruction, copying, blocking access to information. As well as the possibility of affecting system components, resulting in the loss, destruction or failure of the operation of the media, the means of interaction with the carrier or the means of its management.

The main types of threats for multimodal intelligent transport systems are: the threat of confidentiality, the threat of integrity violation, the threat of service failure.

The threat of breach of confidentiality is that information becomes known to someone who does not have the authority to access it.

The threat of integrity violation includes any deliberate alteration of information stored in a computing system or transmitted from one system to another. If the change of information is intentional this means the integrity of the information is violated. This will also be broken if a random software or hardware error leads to a non-sanctioned change.

The threat of service failure occurs as a result of deliberate actions by other users or the attacker resulting in blocking access to some resource of the computing system.

These types of threats are primary as they consider the concept of a threat as some potential hazard, the implementation of which is damaging the system.

Also, a separate kind of threat to system security is the threat of disclosure of the parameters of the information system, including security systems. The threat of disclosure can be considered mediated. Implications of the implementation do not cause any kind of damage to the system, but give the opportunity to realize the primary or non-threatening threats listed above.

Conclusions

The paper considers the information form of interaction of types of transport in multimodal transport information systems. The concept of the threat of information security is disclosed and the main types of information security threats for multimodal transport information systems are analyzed.

Bibliography

Pravdin N.V., Negrey V.Y. *Vzaimodeyctvie razlichnykh vidov transporta v uzlakh.* [Interaction between different modes of transport in knots]. Minsk, Vysheyshaya shkola Pabl., 1997. 289 p.

Gapanovich V.A., Grachev A.A. *Sistemyi avtomatizatsii i informatsionnyie tehnologii upravleniya perevozkami na zheleznyih dorogah* [Automation systems and information technologies for transportation management on railways] M.: Marshrut Pabl., 2006. 544 p.

Jean-Paul Rodrigue. The Geography of Transport Systems, 3-rd edition, New York: Routledge, 2013, pp. 416.

Authors' Information

Tetiana Konrad – Assistant at the chair of of Software Engineering Department, National Aviation University, Kyiv, Ukraine.

E-mail: Konrad.t_i1@ukr.net

Major Fields of Scientific Research: Software engineering, Transport systems, Information systems, Intelligent transport systems

Тези доповідей

Міжнародної науково-практичної конференції

Інженерія програмного забезпечення

Технічні редактори – Markov K., Чебанюк О.В.

Матеріали друкуються за редакцією авторів.

Підписано до друку.

© Національний авіаційний університет